A Bayesian Approach for Estimating Hatchery Contribution in a Series of Salmon Fisheries

Harold J. Geiger

Reprinted from the
Alaska Fishery Research Bulletin
Vol. 1 No. 1, Summer 1994

Alaska Fishery Research Bulletin 1 (1):66–75. 1994.

A Bayesian Approach for Estimating Hatchery Contribution in a Series of Salmon Fisheries

Harold J. Geiger

ABSTRACT: Bayesian methods provide an under-appreciated way of analyzing tag or mark data for hatchery salmon stock identification. For example, the otolith bones can now be marked in captive juvenile salmon, these marks remaining visible in the returning adults. Fishery managers can summarize what is known and unknown about the underlying proportion of hatchery fish in these fisheries using Bayesian methods and the beta probability distribution. Close examination of Bayesian probability theory exposes a philosophy in close agreement with common sense and a form of inference that is direct and agrees with the way people use the notion of probability in everyday, colloquial speech. This theory also provides a straightforward means to allocate sampling resources, in a staged manner, based on information obtained from initial sampling.

INTRODUCTION

Bayesian statistics, although gaining acceptance (e.g., G.G. Thompson 1992; Walters and Ludwig 1993; Hilborn et al. 1993) is largely unused and misunderstood in the field of fisheries. The key features of this form of statistical analysis are (1) that unknown parameters are treated as random variables, (2) only the data that is actually observed is used in the analysis, and (3) the Bayesian algorithm is centered around the idea of using data to update the state of knowledge about the parameters.

Traditional statistical inference is based on the idea of a probability distribution for the data and fixed unknown parameters; both data that was and was not observed is important in this analysis. Tools such as confidence intervals from traditional, or sampling-based, analysis are based on the idea of sets of data that you are expected to observe if you repeat the experiment or study over and over again (Lindgren 1993; S.K. Thompson 1992; Seber 1982). Effron (1986) and Berger and Berry (1988) provided discussions of the merits of the underlying theory behind Bayesian and non-Bayesian inference.

Here, I will examine Bayesian methods for fisheries managers interested in a stock identification system for hatchery-produced pink *Oncorhynchus gorbuscha* and sockeye *O. nerka* salmon, when all of the hatchery population has been marked. This kind of mass-mark can now be applied to juvenile hatchery salmon by manipulation of water temperature in such a way as to create unique mark on the otolith bones (Volk et al. 1990; Brothers 1990).

Bayes' Theorem (e.g., Lindgren 1993) states that, if A is some event that can occur only if one of the mutually exclusive events, B_1 , ... , B_n, occurs, then the probability (Pr) of the event B_i, given that the event A occurred, can be described as follows (the vertical bar "|" denotes a that the latter event is "a given"):

$$Pr\{B_i|A\} = \frac{Pr\{A|B_i\} \cdot Pr\{B_i\}}{Pr\{A\}} \quad .$$

For example, the probability that a fisherman is Norwegian is proportional to the proportion of Norwegians that are fishermen, times the proportion of Norwegians in the world. Dividing these two factors by the proportion of the world's population that are fishermen would give the probability that a fisherman is Norwegian. His being a fisherman corresponds to event A, and the nationality corresponds to B_1 , ... , B_n .

Bayesian analysis of a series of independent success-failure trials, such as the presence or absence of a tag, dates back to Bayes' original work in the 18th century (Press 1989). Stroud (1994) discusses Bayesian analysis in the context of binary survey data, such as the recovery of a tag or other mark. Bayesian inference is unaffected by the rules that govern how the total sample size was generated in success-failure trials, as noted in almost every essay on the virtues of

---

Authors: HAL GEIGER is the Statewide Salmon Biometrician for the Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, P.O. Box 25526, Juneau, Alaska 99802.
Acknowledgments: Jim Blick and Jay Ver Hoef — provided critical reviews of the manuscript and several good ideas that were incorporated.

Bayesian methods (e.g., Box and Tiao 1973, Berger and Wolpert 1988). New features in computer spreadsheets have made Bayesian statistics much more workable for tag or mark data. In particular, the beta probability distribution, which is now in computer spreadsheets, can play a central role in this kind of Bayesian analysis.

In mixed-stock salmon fisheries containing hatchery stocks, the goal is to estimate the fraction of hatchery fish present. Coded wire tags have been used for such a purpose (Peltz and Miller 1990) in hatchery salmon, but these tags are now being phased out for pink salmon in Alaska. In this species coded wire tags are expensive and leave lingering questions about undetected tag loss and tag-related mortality (Geiger and Sharr 1990). Each individual fish must be anesthetized, have a tag injected into its nasal cartilage, and then be marked with a visible fin-clip. Because of the expense and labor involved, only a small fraction of some hatchery releases can be marked. In Prince William Sound, approximately 1 in 600 hatchery fish have been tagged. Sampling for coded wire tags consists of examining the harvest for fin clips, indicating a tag may be present. Fish with visible fin-clip marks are saved for later decoding. In contrast, the otolith bones of the entire production of a salmon hatchery can be microscopically marked by exposing either the embryo or emergent fry to water temperature changes (Volk et al. 1990; Brothers 1990). The cost of otolith marking is low, following initial capital costs for heating equipment. Because all fish are marked the statistical models are much simpler.

If the cost of decoding the marks is small, then a sensible rule might be to take a large sample in each fishery. S.K. Thompson (1987) explains this approach for two or more mark types, with advice on sample size selection. If the cost of mark decoding is very high, or there are a limited number of marks that can be decoded, managers might want to decode some marks and use this information to decide how to proceed with the rest of our sampling resources — that is, dynamically allocate the sampling resources based on sample results.

THE METHOD

Simple Example of Two Fisheries and Two Mark Types

For simplicity, consider only two fisheries and two mark types, although the generalization to more fisheries and multiple marks or tags is straightforward. Some of the details about these generalization are provided in Appendix A. For concreteness, call the first mark type hatchery otoliths and assume all hatchery fish are marked. The second mark type could represent the absence of a hatchery mark. Let N denote the number of fish caught in the first fishery and M the number caught in the second fishery. We assume they are known quantities calculated from some kind of harvest reporting system. Let x denote the number of marks of hatchery otoliths in a random sample of size n in the first fishery, and y the number hatchery otoliths in a random sample of size m in the second fishery. Assume that n is much smaller than N and m is much smaller than M to avoid worry about sampling with or without replacement. Finally, let π denote the true proportion of hatchery fish in the first fishery, and let λ denote the true proportion of hatchery fish in the second fishery. The sampling distributions of the number of hatchery otoliths in the sample are then given by the binomial distribution,

$$f(x|\pi) = \binom{n}{x}\pi^x(1-\pi)^{n-x} \quad , \quad x = 1,2,\ldots n \ \text{, and}$$
$$f(y|\lambda) = \binom{m}{y}\lambda^y(1-\lambda)^{m-y} \quad , \quad y = 1,2,\ldots m.$$

In the Bayesian setting the parameters are treated as random variables, and probability is used as a measure of the certainty at particular parameter values. In the binomial sampling situation, the *beta distribution* is often used to model the probability distribution of the parameters (Lee 1989), for reasons that will soon be clear. This distribution, developed prior to observing any data, is called the *prior distribution*.

Suppose we begin by just assuming that π and λ follow beta distributions with parameters c_1 ... d_2:

$$f(\pi) = \frac{1}{B(c_1,c_2)}\pi^{c_1-1}(1-\pi)^{c_2-1} \ \text{, and}$$

$$f(\lambda) = \frac{1}{B(d_1,d_2)}\lambda^{d_1-1}(1-\lambda)^{d_2-1} \ .$$

$B(c_1, c_2)$ denotes a function called the beta function (which is now in most computer spreadsheets). Note the similarities between the binomial distribution of the number of hatchery otoliths in a sample and the beta distribution of the proportion of hatchery otoliths in the population. Looking at the distribution of π from any statistical theory text (e.g., Lindgren

1993), the mean, denoted $E(\pi)$, and variance, denoted $V(\pi)$, is given by

$$E(\pi) = \frac{c_1}{c_1 + c_2} \quad , \tag{1}$$

$$V(\pi) = \frac{c_1 c_2}{(c_1 + c_2 + 1)(c_1 + c_2)^2} \quad . \tag{2}$$

By using Bayes' Theorem, we obtain the probability distribution for the unknown fraction of hatchery fish in a fishery after observing some data. A new probability distribution is formed by what is called in probability, *conditioning* on the data that was actually observed. We call the new distribution the *posterior distribution* — that is, the distribution of the parameters after, or posterior to, observing the data. More importantly this posterior distribution is also a beta distribution (Lee 1989). This distribution will be proportional to the sampling distribution of the number of marked fish in a sample from the fishery (i.e., $f(x|p)$) and the prior distribution for π (i.e., $f(\pi)$; Lee 1989). Recalling that x was the number of hatchery fish in a sample of n fish from the fishery, define $c_1' = x + c_1$ and $c_2' = n - x + c_2$, (and so forth with m and y for d_1' and d_2'). This emphasizes the recursive nature of the process by which the probability distribution that describes what we know, or don't know, is updated in the Bayesian setting. In mathematical form — which is not really needed to make this work in a computer spreadsheet — the posterior distribution is given by

$$f(\pi|x) = \frac{1}{B(x + c_1, n - x + c_2)} \pi^{x + c_1 - 1} (1 - \pi)^{n - x + c_2 - 1}$$

$$= \frac{1}{B(c'_1, c'_2)} \pi^{c'_1 - 1} (1 - \pi)^{c'_2 - 1}, \text{ and}$$

$$f(\lambda|y) = \frac{1}{B(y + d_1, m - y + d_2)} \lambda^{y + d_1 - 1} (1 - \lambda)^{m - y + d_2 - 1}$$

$$= \frac{1}{B(d'_1, d'_2)} \lambda^{d'_1 - 1} (1 - \lambda)^{d'_2 - 1} \quad .$$

When data have been collected, $c_1' = x + c_1$ and $c_2' = n - x + c_2$ (and d_1', and d_2') become constant, fixed numbers. These numbers are placed in a spreadsheet function to get useful descriptions about the unknown fraction of hatchery fish from the probability distributions. Specifically, the best single estimate of the unknown parameter is the mean of the posterior distribution (e.g., Lee 1989 describes why the mean of the posterior distribution is the best estimate by the squared-error loss criterion).

Using equations (1) and (2) find the posterior mean and variance of π. The mean is given by

$$E(\pi|x) = \frac{x + c_1}{n + c_1 + c_2} \quad ,$$

which will be near the usual estimate of x/n when the sample size is large. The posterior variance the proportion of hatchery fish is given by

$$V(\pi|x) = \frac{(x + c_1)(n - x + c_2)}{(n + c_1 + c_2 + 1)(n + c_1 + c_2)^2}$$

$$= \frac{E(\pi|x) E(1 - \pi|x)}{(n + c_1 + c_2 + 1)} \quad ,$$

which will be near the usual sample variance for the binomial distribution, $(x/n)[(n-x)/n]/n$, with a large n and relatively small c's. The mean and variance of λ are found by similar applications of equations (1) and (2). Notice, by using beta prior distributions, the parameters of the posterior distribution are found by combinations of the c_1', c_2', d_1', and d_2' constants. At each stage of the analysis the latest values of the constants c_1', c_2', d_1', and d_2' are all that needs to be supplied to the computer spreadsheet for a complete inventory of what we know about the fishery.

In the Bayesian setting, by conditioning on what we know at the moment, today's prior distribution will simply be yesterday's old posterior distribution. The values c_1', c_2', d_1', and d_2' are updated by adding the current data into the appropriate place (i.e., the new c_1' will be the old c_1' plus the new x). Notice also that with increasing data, the original constants c_1, c_2, d_1, and d_2 have increasingly less influence.

To look at a sensible interval for the unknown parameter, have your computer spreadsheet look up the $\alpha/2$ (100%) and $(1-\alpha/2)$(100%) percentiles of the beta distributions with parameters c_1' and c_2' (or d_1', and d_2'). In Bayesian inference these intervals are called *credible intervals*. Unlike confidence intervals, with credible intervals you can correctly make the direct statement that the probability is (1-α/2) that the true value is in this interval. For an interval with 95% probability, using either Excel™ or Quattro™ spreadsheets, these limits are found with the statements "@BETAINV(0.025, c_1', c_2')" for the lower limit and "@BETAINV(0.975, c_1', c_2')" for the upper limit, with c_1' and c_2' denoting the cells with the current beta parameters.

Dynamic Sample Sizes Between Two Fisheries

Suppose that the real source of our interest is the overall proportion of marked fish in both fisheries combined, which we will denote ρ. Then recalling that N and M are the total number of fish in the harvest of each fishery

$$\rho = \frac{N}{N+M}\pi + \frac{M}{N+M}\lambda \ .$$

Then,

$$E(\rho|\text{data}) = \frac{N}{N+M}E(\pi|x) + \frac{M}{N+M}E(\lambda|y) \ , \qquad (3)$$

and

$$V(\rho|\text{data}) = \left(\frac{N}{N+M}\right)^2 V(\pi|x) + \left(\frac{M}{N+M}\right)^2 V(\lambda|y) + 2\frac{NM}{(N+M)^2}Cov(\pi,\lambda|x,y) \ . \qquad (4)$$

Now to achieve a good distribution of sampling resources, we may decide that we want the standard deviation of ρ to be as small as possible. Assume for the moment that π and λ are independent; I will return to this subject below. Recall samples of size n and m have already been taken, but suppose one or both of these is to be increased by some amount. If $N\,[\pi(1-\pi)]^{1/2}/n >> M\,[\lambda(1-\lambda)]^{1/2}/m$, then additional samples should come from Fishery 1 because increasing n will cause the fastest decline in the standard deviation of ρ (Appendix B). Similarly, if $N\,[\pi(1-\pi)]^{1/2}/n \approx M\,[\lambda(1-\lambda)]^{1/2}/m$, then additional samples should come from both fisheries, and if $N\,[\pi(1-\pi)]^{1/2}/n << M\,[\lambda(1-\lambda)]^{1/2}/m$, then additional samples should come from Fishery 2. This rule is simply telling us to devote more sampling resources where there is more variability or where the catch is larger.

Choosing parameters for an initial prior distribution

Because the Bayesian algorithm is based on a flow of learning, the prior is needed to prime the inference pump. The prior distribution is the most misunderstood component of Bayesian analysis and is sometimes thought of as a way to subvert what the data are trying to say. A poorly thought out prior distribution can lead to poor inference, just as unreasonable α and β error rates can in a classical statistical hypothesis test. Berger (1985) provides a complete, but highly mathematical, discussion of how to choose prior distributions. As n gets bigger and bigger, x and n-x will increase too, x and n-x being the number of hatchery and non-hatchery marks in the sample. Eventually, at very large sample sizes, c_1 and c_2 will be insignificant compared with x and n-x. So, the first consideration might be a prior probability that will result in the sum of c_1 and c_2 being relatively small as a reasonable amount of data are accumulated .

By restricting the sum of c_1 and c_2 to a small value, say even less than one, the prior probability can model a wide range of reasonable and realistic representations of the state of knowledge for the manager. One alternative is to let c_1 and c_2 equal 1. This, in effect, lets all values between 0 and 1 be equally likely before looking at any data. Figure 1 shows an example of prior and posterior distributions with the sum of c_1 and c_2 restricted to 0.75, a sample size of 100, and outcomes of x = 10 and x = 90. In this example the final inference is almost completely controlled by the data and virtually unaffected by the choice of prior probability.

In the end, the choice of c_1 and c_2 has the same basis in judgement that the choice of α and β error rates in a statistical hypothesis testing. After experimenting with various values of c_1 and c_2, I recommend letting c_1 = 0.25 and c_2 = 0.5, although another analyst may have some reason for increasing these initial values. The most obvious reason is that data from fisheries adjacent in time or space is providing information that can somehow be incorporated.

Inseason analysis of serially ordered fisheries, and dependence between π and λ

Earlier we simplified by assuming that π and λ are independent. But in fisheries separated by only a couple of days, most managers will think that what happens in one period will be similar to what happens in the next. Adjacent fishing districts will also tend to have similar proportions of hatchery fish. In statistical terms, π and λ are dependent.

In traditional statistical sampling, the sampling events are independent in the two fisheries so that the estimates of the parameters are independent. In the Bayesian setting we are developing a probability distribution for the parameters themselves, and so the parameters are not completely independent.

Suppose Fishery 1 takes place first. The results from Fishery 1 should influence the prior distribution for Fishery 2, and this is how a dependence between π and λ is expressed. The question of how to set the prior probabilities for Fishery 2 has no automatic answer. One straightforward way to model dependence from one fishery to the next is to introduce the quantity h, which gives a proportional relationship between the posterior information from the earlier fishery and the prior probabilities in the latter fishery; i.e., either

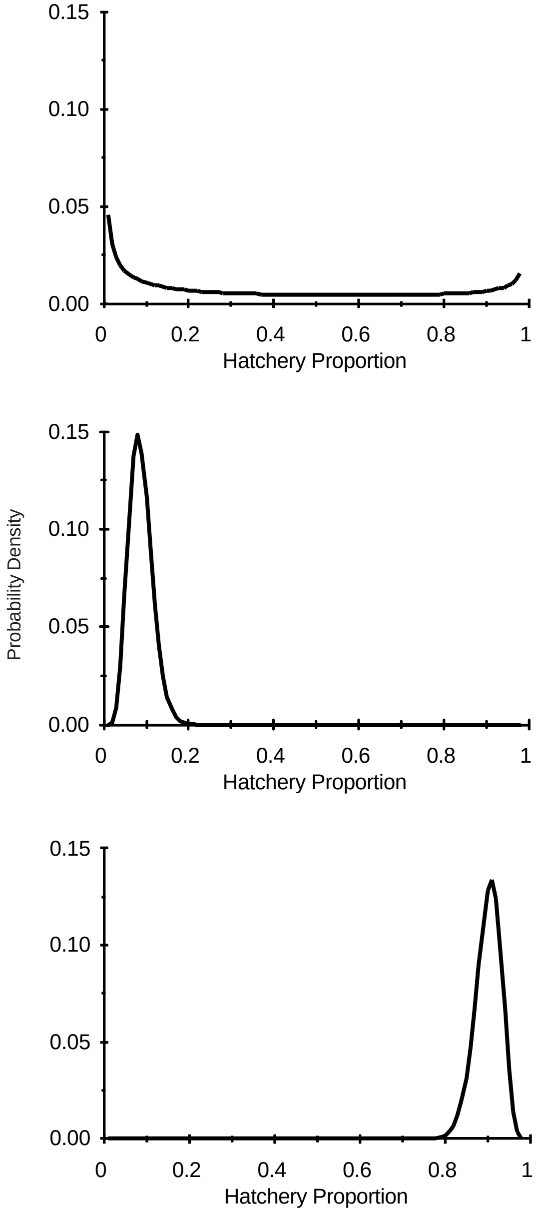

Figure 1. The top figure shows the prior distribution of the proportion of hatchery fish a fishery before examining any otoliths. This distribution is a beta distribution with parameters $c_1 = 0.25$ and $c_2 = 0.5$. The middle distribution shows the posterior distribution after decoding 100 otoliths and observing 10 of the hatchery type. The bottom figure shows the posterior distribution after decoding 100 otoliths and observing 90 of the hatchery type.

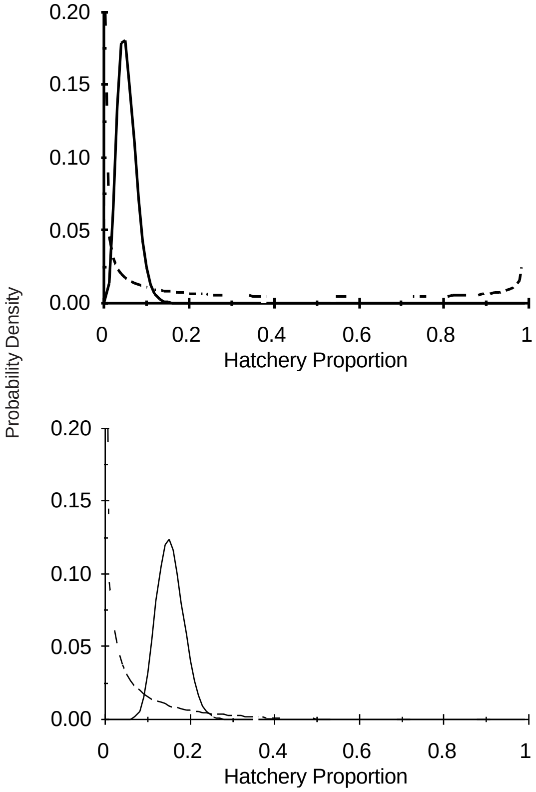

Figure 2. Prior and posterior distributions for the proportion of hatchery fish in two 1994 Southeast Alaskan pink salmon fisheries occurring in series. The prior distribution for the first fishery is a beta distribution with parameters $c_1 = 0.25$ and $c_2 = 0.5$, shown with a dashed line in the top graph. After observing $x = 6$ hatchery fish out of a sample size of 100, the posterior distribution is given by the solid line on the same graph. Based on the results of the first fishery, the prior distribution of the second fishery is a beta distribution with parameters $d_1 = 5\ (6/100)$ and $d_2 = 5\ (94/100)$, with mean 0.056, shown as the dotted line on the bottom graph. After observing $y = 20$ hatchery fish in a sample of 103 from the second fishery, the posterior distribution, with mean 0.158, shown with the solid line on the bottom graph.

Table 1. Parameters of the posterior distribution of proportions of hatchery fish in two fisheries. These proportions are denoted π and λ and are dependent through the prior distribution of the second fishery. The second fishery's prior is based on observed hatchery otoliths in the first fishery. The beta prior distribution in the first fishery is given by parameters $c_1 = 0.5$ and $c_2 = 1.0$. With n and m, the sample sizes in each of the two fisheries (fixed at 100), x and y, are the respective detected hatchery otoliths. The beta parameters in the second fishery are $d_1 = 5\ (x/n)$ and $d_2 = 5\ [(n - x)/n]$.

Level of	Level of	Mean of	Posterior Variance of		Covariance
x	y	λ	π	λ	*of* π *and* λ
10	10	0.098	$7.7\ 10^{-4}$	$7.1 \cdot 10^{-4}$	$6.2 \cdot 10^{-6}$
10	50	0.480	$7.7\ 10^{-4}$	0.002	$1.1 \cdot 10^{-6}$
10	90	0.853	$7.7\ 10^{-4}$	0.002	$2.1 \cdot 10^{-5}$
50	10	0.116	0.002	$9.7 \cdot 10^{-4}$	$2.8 \cdot 10^{-6}$
50	50	0.499	0.002	0.002	$1.4 \cdot 10^{-6}$
50	90	0.855	0.002	0.002	$3.0 \cdot 10^{-5}$
90	10	0.137	0.002	0.001	$1.9 \cdot 10^{-5}$
90	50	0.518	0.002	0.002	$6.6 \cdot 10^{-6}$
90	90	0.886	0.002	0.001	$3.3 \cdot 10^{-4}$

$d_1 = h\,c'_1$ and $d_2 = h\ c'_1$, or perhaps $d_1 = h\,x$ and $d_2 = h\ (n\text{-}x)$.

Based on trial and error, I suggest scaling the sum of d_1 and d_2 to be 5, and letting $h = 5/n$; i.e., $d_1 = 5\ (x/n)$ and $d_2 = 5\ [(n\text{-}x)/n]$. This seems to do a reasonable job of expressing prior beliefs about a fishery that has not yet happened and still results in the data dominating the posterior distribution when the sample sizes approach large values (e.g., near 100). Figure 2 shows an example of two fisheries with dependent parameters modeled using this rule.

The analyst is usually concerned inseason about the fraction of hatchery fish, fishery by fishery. Inseason, the analyst is not concerned about an overall estimate of hatchery fish in the two fisheries combined.

Postseason analysis and the posterior distribution of the overall estimate

For the postseason analysis, the covariance between parameters can safely be ignored if n and m exceed 100 and the rule of scaling the sum of d_1 and d_2 to 5 is followed. Table 1 shows calculated parameters of the posterior distributions of π and λ for high medium and low values of x and y. The covariance of π and λ is always less than two orders of magnitude less than the minimum of the variances with x and y between 10 and 90 and sample sizes of 100.

For the postseason analysis, the manager will want the posterior distribution of ρ. By the time hundreds of otoliths have been examined, unless virtually all of them were of one type, the posterior distribution of ρ is very near the normal distribution, with the mean and variance given by equations (3) and (4). The credible intervals for ρ can safely be constructed with the posterior mean, plus or minus posterior standard deviation times the desired factor from the standard normal distribution (e.g., 1.96 for a 95% credible interval).

Sample Size Determination

In planning for the postseason analysis, most managers will usually be familiar with requesting resources sufficient to reduce the standard error of an estimate of the proportion of hatchery fish in all fisheries. Gauging the size of the request can be done by looking at previous years and guessing at the total catch in each fishery, N_1 , N_2 , ... , then assuming that proportion of hatchery fish is close to 0.5 (the worst case for the variance) for all fisheries. The standard error of the estimate of overall hatchery contribution to all fisheries is controlled by sample size. This standard error is reduced by increasing n. To reach precision objectives, solve the following for n:

$$SE(\rho) \le \left\{ \sum \left[\left(\frac{N_i}{\sum N_j} \right)^2 \times \frac{(0.5)^2}{n} \right] \right\}^{1/2} ,$$

with the left side replaced by the desired standard error of the estimate. Assuming we take good guesses at N_1 , N_2 , ..., then with the n_i's dynamically allocated, as explained above in the section *Dynamic*

Sample Sizes Between Two Fisheries, the precision will be generally better than planned.

Usually the manager does not have a definite idea about precision and is looking for guidance. Figure 3 shows a graph of $1/\sqrt{n}$, the rate at which the standard error decreases for a fixed-population standard deviation. Notice this flattens out considerably by 30, and very little gain in precision is achieved past a sample size of 100. For this reason, a starting sample size of 100 makes sense for all fisheries, with increases to come in the dynamic allocation step in the postseason analysis.

In a complex, multifishery setting the sampling should proceed in a series of steps. First, a starting sample size is determined for each fishery based on the minimum needed for inseason analysis. Second, large numbers of fish are sampled from the fishery and inventoried. Third, the fixed number of marks are decoded from each fishery. Fourth, the preliminary information is fed into the decision rules laid out above in the section *Dynamic Sample Sizes Between Two Fisheries*. Fisheries where additional samples will cause the fastest decline in the standard deviation in the posterior distribution of the overall mark rate are noted. Fifth, another increment of marks are decoded based on the newest information.

The size of the increment will be determined by production considerations in the processing laboratory. For small batch sizes, I suggest first running 100 otoliths from each fishery. Next, take an additional batch from the fishery that will cause the largest increase in precision, based on the rule in the section *Dynamic Sample Size Between Two Fisheries*. Repeat the process of sampling and reevaluating until the sampling resources are exhausted.

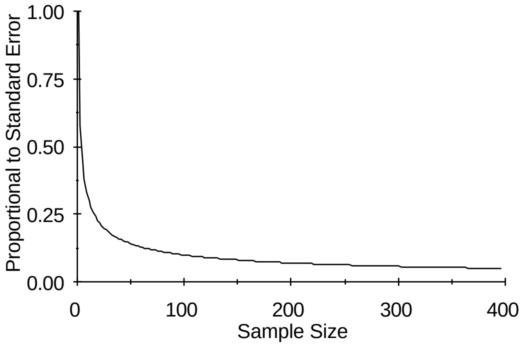

Figure 3. The rate at which the standard error declines as a function of sample size for a given population standard deviation. The rate is simply the function $1/\sqrt{n}$, where n denotes sample size.

Also see Fuchs et al. (1993) for a discussion of sample size for the kind of yes-no data fish marks present in the Bayesian setting, using prior information.

CONCLUSIONS

These methods provide a rational, staged means of estimating the contribution of hatchery fish in a series of fisheries. The use of stages makes the best possible use of sampling resources. The analyst proceeds logically from a point of relative ignorance to a point of relative certainty, with a cogent summary of the state of knowledge at each stage. The way the relative certainty is summarized — in the posterior distribution of the proportion of hatchery fish in the catch — is intuitive and easy to present and understand using graphs.

The original prior probabilities are to prime the inference pump. By starting with prior probabilities that reflect the fact that not much is known before collecting the data, the prior probabilities have little affect on the final inference. By using information to shape prior probabilities, when the information exists, the posterior distribution realistically reflects what the analyst knows at the end of the study.

Turning to the issue of how the results are used, in the Bayesian setting confidence intervals are not used because these are based on some idea of repeated sampling. Instead, a credible interval makes the direct statement people want from a confidence interval. If 95% is the desired level, we use the posterior probability distribution to directly state that the probability is 95% that the parameter is in the interval from *a* to *b*. Notice the difference between this simple statement and the complex, convoluted logic of the confidence interval.

The data user can think of a 95% credible interval as a 95% confidence interval with no real harm. The mean of the posterior distribution is so near to the classical point estimates as to be virtually indistinguishable when sample sizes are large. But the open-minded user gets the entire posterior distribution, which offers a richer, more sensible, and more direct summary of the available evidence.

LITERATURE CITED

Berger, J.O. 1985. Statistical decision theory and Bayesian analysis, 2nd edition. Springer-Verlag. New York.

Berger, J.O., and D.A. Berry. 1988. Statistical analysis and the illusion of objectivity. American Scientist 76(2):159–165.

Berger, J.O., and R.L. Wolpert. 1988. The likelihood principle, 2nd edition. Institute of Mathematical Statistics, Lecture Notes — Monograph Series Volume 6.

Brothers, E.B. 1990. Otolith marking. Pages 183–202 *in* Parker, N.C. and five coeditors. American Fisheries Society Symposium 7, Bethesda, Maryland.

Box, G.E.P., and G.C. Tiao. 1973. Bayesian inference in statistical analysis. Addison-Wesley, Reading, Massachusetts.

Efron. B. 1986. Why isn't everyone a Bayesian? The American Statistician. 40 (1):1–11.

Fuchs, K.D. Werncke, J. Golles, and J. Kofer. 1993. Calculating the sample size in the examination of animals for the presence of residues using prior information. Biomectrical Journal. 35(1993)7, 801–813.

Geiger, H.J., and S. Sharr. 1990. The 1988 tagging study of pink salmon from the Solomon Gulch Hatchery in Prince William Sound, Alaska. Pages 20–33 *in* H.J. Geiger, editor. Pilot Studies in tagging Prince William Sound hatchery pink salmon with coded-wire tags. Alaska Department of Fish and Game, Division of Commercial Fisheries, Fishery Research Bulletin 90-02. Juneau, Alaska.

Hilborn, R. E.K. Pikitch, and M.K. McAllister. 1994. A Bayesian estimation and decision analysis for an age-structured model using biomass survey data. Fisheries Research. 19:17-30

Lee, P. 1989. Bayesian statistics: an introduction. Charles Griffin. New York - Toronto.

Lindgren, B.W. 1993. Statistical theory, 4th edition. Chapman Hall. New York.

Peltz, L., and J. Miller. 1990. Performance of half-length coded wire tags implanted in emergent pink salmon. Pages 244–252 *in* Parker, N.C. and five coeditors. American Fisheries Society Symposium 7, Bethesda, Maryland.

Press, S. J.. 1989. Bayesian statistics: principles, models and applications. Wiley Interscience, New York.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters, 2nd edition. Macmillan Publishing Co. New York.

Stroud, T.W.F. 1994. Bayesian analysis of binary survey data. The Canadian Journal of Statistics 22(1):33–45.

Thompson, G.G. 1992. A Bayesian approach to management advice when stock-recruitment parameters are uncertain. Fishery Bulletin 90:561–573.

Thompson, S. K. 1987. Sample size for estimating multinominal proportions. The American Statistician 41:42–46.

Thompson, S. K. 1992. Sampling. Wiley-Interscience. New York.

Volk, E. C., S.L. Schroder, and K.L. Fresh. 1990. Inducement of unique banding patterns as a practical means to mass-mark Juvenile Pacific salmon. Pages 203-215 *in* Parker, N.C. and five coeditors. American Fisheries Society Symposium 7, Bethesda, Maryland.

Walters, C., and D.L. Ludwig. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Canadian Journal of Fisheries and Aquatic Sciences 51:713-722.

— *Appendix A.* —

Appendix A. Generalization to multiple tag codes and fisheries.

Two tag codes and two fisheries generalize in a straightforward way, although simultaneous analysis of more than one tag type is not possible in a computer spreadsheet at this time. Let $\xi_i(g)$ denote the true proportion of fish with tag-code g in fishery i, with g and i extending over all codes and fisheries. If $x_i(g)$ represents the number of g-type codes found in a sequence of n_i random tag decodings in fishery i, and $c_{i,j}$ are the small constants that define the prior distribution, then

$$f\left(\xi_i(A),\xi_i(B),\xi_i(C)\ldots\middle|x_i(A),x_i(B),x_i(C)\ldots\right)=\left[\sum x_i(j)\right]!\prod_{\mathrm{j}}\frac{\xi_{\mathrm{i}}(j)^{x_i(j)+c_{i,j}-1}}{x_i(j)!}\quad .$$

This distribution is called the Dirichlet distribution, and full descriptions can be found in appropriate statistical text books (e.g., Berger 1980). The constants that define the prior distribution are developed in the same way as in the two-tag, two-fishery example by using judgement and experience, just as with error rates in a hypothesis test. The initial constants should be very small and similar to express relative ignorance. After the first fishery, the constants should be small, but slightly larger and be in proportion to the sample results in adjacent fisheries.

Define $\rho(g)$ as the overall proportion of fish with the mark type g in all fisheries. Let N_i be the total catch of fish in fishery i. Then, for example, a straightforward way to summarize the best guess at the overall proportion of the g-type mark is the weighted average over all fisheries, just as before:

$$E\left(\rho(g)|\text{data}\right)=\sum_i\frac{N_i}{\sum_j N_j}E\left(\xi_i(g)|\text{data}\right)\quad .$$

The variance is found as before with covariances of proportions in each fishery found as

$$Cov\left(\xi_i(g),\xi_j(g)\ |\text{data}\right)=$$

$$\int_0^1\int_0^1\left[\xi_i(g)-E\left(\xi_i(g)|\text{data}\right)\right]\left[\xi_j(g)-E\left(\xi_j(g)|\xi_i(g),\text{data}\right)\right]f\left(\xi_i(g)|\text{data}\right)f\left(\xi_j(g)\ |\xi_i(g),\text{data}\right)d\xi_i(g)d\xi_j(g)\quad ,$$

where fishery i comes before fishery j and the prior of fishery j is assumed to be based on the outcome in fishery i. The notation $f(\xi_j(g)\,|\,\xi_i(g)$, data) denotes the posterior distribution for $\xi_j(g)$ after using information from fishery i.

— Appendix B. —

Appendix B. Dynamic sample size justification.

To find the combination of sample sizes that will cause the fastest decline in the standard deviation of ρ, consider the partial derivatives of $[V(\rho \mid \text{data})]^{1/2}$ with respect to n and m. The most negative of these is the direction with the fastest decreases in the standard deviation. Assume that the covariance between π and λ is negligible, as discussed above. Because x increases randomly in response to an increase in n, we cannot take the derivative of $E(\pi \mid x)$ and $E(\lambda \mid y)$ with respect to n or m. So, assume that the constants c_1, c_2, d_1, and d_2 are negligible, and let π and λ stand in place of $E(\pi \mid x)$ and $E(\lambda \mid y)$, which are functions of n and m, for the purposes of differentiating. Then the partial derivatives of the standard deviation of ρ are approximated as follows:

$$\partial \frac{V(\rho)^{1/2}}{\partial n} \approx -\left\{\left[\left(\frac{N}{N+M}\right)^2\right]\left(\frac{1}{n^2}\right)\left[\pi(1-\pi)\right]\right\} \times 1/2V(\rho)^{-1/2} \quad ,$$

$$\partial \frac{V(\rho)^{1/2}}{\partial m} \approx -\left\{\left[\left(\frac{M}{N+M}\right)^2\right]\left[\frac{1}{m^2}\right]\left[\lambda(1-\lambda)\right]\right\} \times 1/2V(\rho)^{-1/2} \quad .$$

Setting these equal and solving for either n or m gives the desired result.

Alternatively, the problem can be approached by minimizing the variance, subject to constraints. Define the Lagrange multiplier, γ, for the constraint $n + m = r$, where r is the total of the fixed resources. Now define

$$Q = \left[V(\rho)\right] + \gamma(n + m - r) \quad .$$

These partial derivatives are

$$\partial \frac{Q}{\partial n} \approx -\left[\left(\frac{N}{N+M}\right)^2\right]\left(\frac{1}{n^2}\right)\left[\pi(1-\pi)\right] + \gamma \ ,$$

$$\partial \frac{Q}{\partial m} \approx -\left[\left(\frac{M}{N+M}\right)^2\right]\left(\frac{1}{m^2}\right)\left[\lambda(1-\lambda)\right] + \gamma \quad ,$$

$$\partial \frac{Q}{\partial \gamma} = n + m - r \quad .$$

Setting them all equal to zero we get $\partial \frac{Q}{\partial n} = \partial \frac{Q}{\partial M}$, subject to $n + m = r$. So

$$n^2 = m^2 \, \frac{\pi(1-\pi)N^2}{\lambda(1-\lambda)M^2} \quad .$$

Replacing n^2 in the left side with $(m - r)^2$ reduces the problem to one variable that can be solved using the quadratic equation.

Using a spreadsheet, for more than two fisheries, the practical thing to do is use the resources one batch at a time until the resources are exhausted. Just recompute the approximate derivatives after each batch is processed and take the samples for the next batch from the fishery with the most negative derivative.