

Abstract—For most fisheries applications, the shape of a length-frequency distribution is much more important than its mean length or variance. This makes it difficult to evaluate at which point a sample size is adequate. By estimating the coefficient of variation of the counts in each length class and taking a weighted mean of these, a measure of precision was obtained that takes the precision in all length classes into account. The precision estimates were closely associated with the ratio of the sample size to the number of size classes in each sample. As a rule-of-thumb, a minimum sample size of 10 times the number of length classes in the sample is suggested because the precision deteriorates rapidly for smaller sample sizes. In absence of such a rule-of-thumb, samplers have previously under-estimated the required sample size for samples with large fish, while over-sampling small fish of the same species.

Precision estimates and suggested sample sizes for length-frequency data

Hans D. Gerritsen (contact author)¹

David McGrath²

Email address (for H. D. Gerritsen): hans.gerritsen@marine.ie

¹ Fisheries Science Service, Marine Institute
Rinville, Oranmore
County Galway, Ireland

² Commercial Fisheries Research Group
Galway-Mayo Institute of Technology
Dublin Road
Galway, Ireland.

Length measurements are fundamental to many aspects of fisheries science. However, there is little formal guidance on the appropriate size of a length sample. Such guidance is of particular relevance when the number of fish available exceeds the number that can be measured at a reasonable cost, and a subsample needs to be taken. Clearly, the required precision of a length sample depends on the purpose of sampling. In order to identify modes of individual year classes for a length-based assessment, the precision of the sample needs to be quite high. Sample sizes of more than 1000 are necessary to identify more than half the modes in a typical length distribution (Erzini, 1990). A sample size of at least 100 adult fish was recommended for age-based stock assessment purposes (Anderson and Neumann, 1996), although the authors did not mention how they arrived at this number.

Regardless of the type of assessment that is used, the shape of the length-frequency distribution is of interest, rather than simple summary statistics such as the mean or the variance. For this reason, it has proved difficult to quantify what constitutes a representative or adequately precise length distribution. Some studies have attempted to find minimum or optimum sample sizes by comparing samples to an expected distribution (e.g., Muller¹; Gomez-Buckley et al.²; Vokoun et al., 2001). However, the true distribution is usually unknown, and dissimilarity

from the expected distribution does not necessarily indicate an imprecise sample. In addition, these methods provide only indirect measures of precision that are difficult to evaluate objectively.

Thompson (1987) used the precision of a sample explicitly to establish an appropriate sample size. Thompson proved that a sample size of 510 is sufficient to be 95% confident that all estimated proportions in a multinomial distribution are no more than 5% from the true proportion. However, Thompson based this figure ($n=51$) on a worst-case scenario, which, in the present case, is a length-frequency distribution that is evenly apportioned over three size classes. Because this is not the typical shape of a length-frequency distribution used in fisheries science, Thompson's measure of precision is too conservative for the vast majority of cases.

For most fisheries applications, it would be more useful to define the

¹ Müller, H. 1996. Minimum sample sizes for length distributions of the catch estimated by an empirical approach. ICES CM 1996/J12, 18 p.

² Gomez-Buckley, M., L. Conquest, S. Zitzer, and B. Miller. 1999. Use of statistical bootstrapping for sample size determination to estimate length-frequency distributions for Pacific albacore tuna (*Thunnus alalunga*). Final report to National Marine Fisheries Services, FRI UW 9902, 7 p. Website: <http://www.fish.washington.edu/research/publications/pdfs/9902.pdf> (accessed 31 March 2006).

precision of a length-frequency sample as the mean precision over the entire size range. However, it appears that this approach has not been used to establish an optimum sample size. Such mean precision estimates over the entire size range might be used to obtain a rule-of-thumb for sample sizes that are required in order to obtain a certain precision level of the catch at each location. In the present study we aim 1) to determine a rule-of-thumb for obtaining an appropriate sample size when the number of fish available in a particular sample exceeds the number that can be measured at a reasonable cost, and 2) to examine the sample sizes that have been taken in the past, in absence of such guidance.

Materials and methods

Data were used from the Irish Groundfish survey, which was carried out on RV *Celtic Explorer* in the waters around Ireland during October and November 2005. The catch was sorted into species and, if appropriate, into size grades, each of which were treated as a separate length sample. Length measurements were taken from all fish and squid species that were caught. If the number of individuals in a sample was large, a subsample was taken by repeatedly transferring the sample from each fish box into two other boxes and discarding one of these. This method ensures that the entire catch is represented uniformly in the subsample. At the time of the survey, the samplers did not have any particular guidance on the appropriate size for a subsample; they used their own judgment to decide on the sample size.

The precision of the number of observations in each length class of a random sample can be estimated by assuming a multinomial distribution (Smith and Maguire, 1983). If the precision in each length class is expressed in the form of a coefficient of variation (CV), an overall measure of precision can be obtained by weighting each CV by the number of fish in each length class. This mean weighted CV (MWCV) provides a description of the precision over the entire range of size classes in a length frequency distribution.

Under the assumption of a multinomial distribution, the standard deviation (σ_i) of the number of fish in a sample that are length category i can be estimated by

$$\sigma_i = \sqrt{np_i(1-p_i)}, \quad (1)$$

where n = the total number of fish in the sample; and p_i = the proportion of the sample that is length i .

The coefficient of variation (CV) of the number of fish at length i , is given by

$$CV_i = \frac{\sigma_i}{np_i} \quad (2)$$

and the mean weighted coefficient of variation (MWCV) is given by

$$MWCV = \sum p_i CV_i = \frac{\sum \sigma_i}{n}. \quad (3)$$

The highest possible value of the MWCV results from a length-frequency distribution that is evenly distributed over a large number of size classes. The number of fish at each length class are then Poisson distributed with a standard deviation that equals the square root of the number at length (Zar, 1999). The theoretical maximum MWCV is therefore given by

$$MWCV = (n/c)^{-0.5}, \quad (4)$$

where c = the number of size classes in the sample.

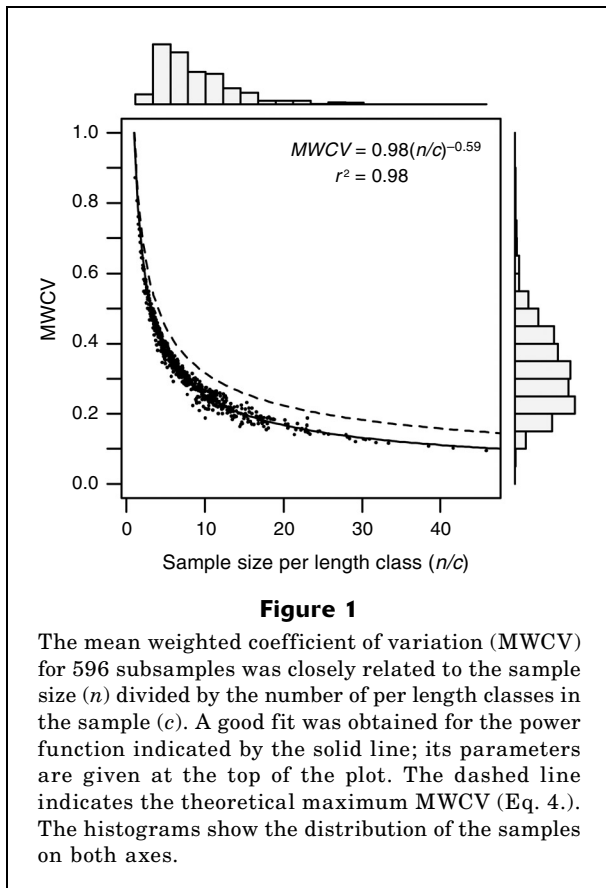
The minimum MWCV is zero and would result from a distribution where all observations fall within a single length category. Therefore, the MWCV estimates will always lie between zero and the curve described by Equation 4.

Results

During the 2005 survey, a total of 2332 length samples were taken for 80 different species of fish and squid. In most cases, the sample size was limited by the number of individuals in the catch. However, 596 samples were deemed too large to measure all individuals and subsamples were taken. The median subsample size was just under a quarter of the total catch (by weight), whereas 90% of the subsamples were smaller than half of the total catch. The four most common species that were subsampled were poor cod (*Trisopterus minutus*), blue whiting (*Micromesistius poutassou*), haddock (*Melanogrammus aeglefinus*), and Norway pout (*Trisopterus esmarkii*).

The estimated MWCV of the subsamples was closely associated with the ratio of the number of individuals measured to the number of length classes in the sample (Fig. 1). The MWCV appeared to follow an exponential curve that was close to the maximum MWCV given by Equation 4. The MWCV decreased very rapidly with increasing sample size up to sample sizes of around 10 times the number of length classes in the sample, after which the sample size would need to be increased considerably for a moderate further improvement in precision. If the sample size is taken as 10 times the number of length classes in the distribution, an MWCV of around 0.25 can be expected; a sample size of 48 times the number of length classes would result in an MWCV of 0.10 and a sample size of 155 times the number of length classes would be necessary to reduce the MWCV to 0.05.

The mean sample size in the subsamples taken on the survey was just under nine times the number of length classes per sample, resulting in a mean MWCV of 0.33. However, there was quite a large spread in the sample sizes (Fig. 1); therefore some samples were measured with very low precision, whereas others had



excessively large sample sizes. The range of sample sizes was between 2.2 and 24.7 times the number of length classes (2.5% and 97.5% quantiles), resulting in a range of MWCVs between 0.14 and 0.61. With a minor increase in effort, the sample size might be increased to 10 per length class for each subsample, resulting in an MWCV of around 0.25 for all samples. Considering that the precision deteriorates very rapidly for sample sizes of less than 10 per length class, a minimum sample size of 10 times the number length classes in the sample is suggested as a rule-of-thumb in the present case.

The previous analysis shows that, in order to obtain the same level of precision for all subsamples, the sample size should be directly proportional to the number of size classes. In absence of specific guidance on the sample size during the 2005 survey, the chosen sample size was only weakly correlated to the number of length classes in the sample of poor cod and haddock, whereas no significant correlation was found for blue whiting and Norway pout (Fig. 2). The same figure also shows that the MWCV in subsamples tended to increase with the mean length of the fish in the sample. This increase indicates that samples with a large mean size tended to be sampled with lower precision than samples of smaller fish of the same species.

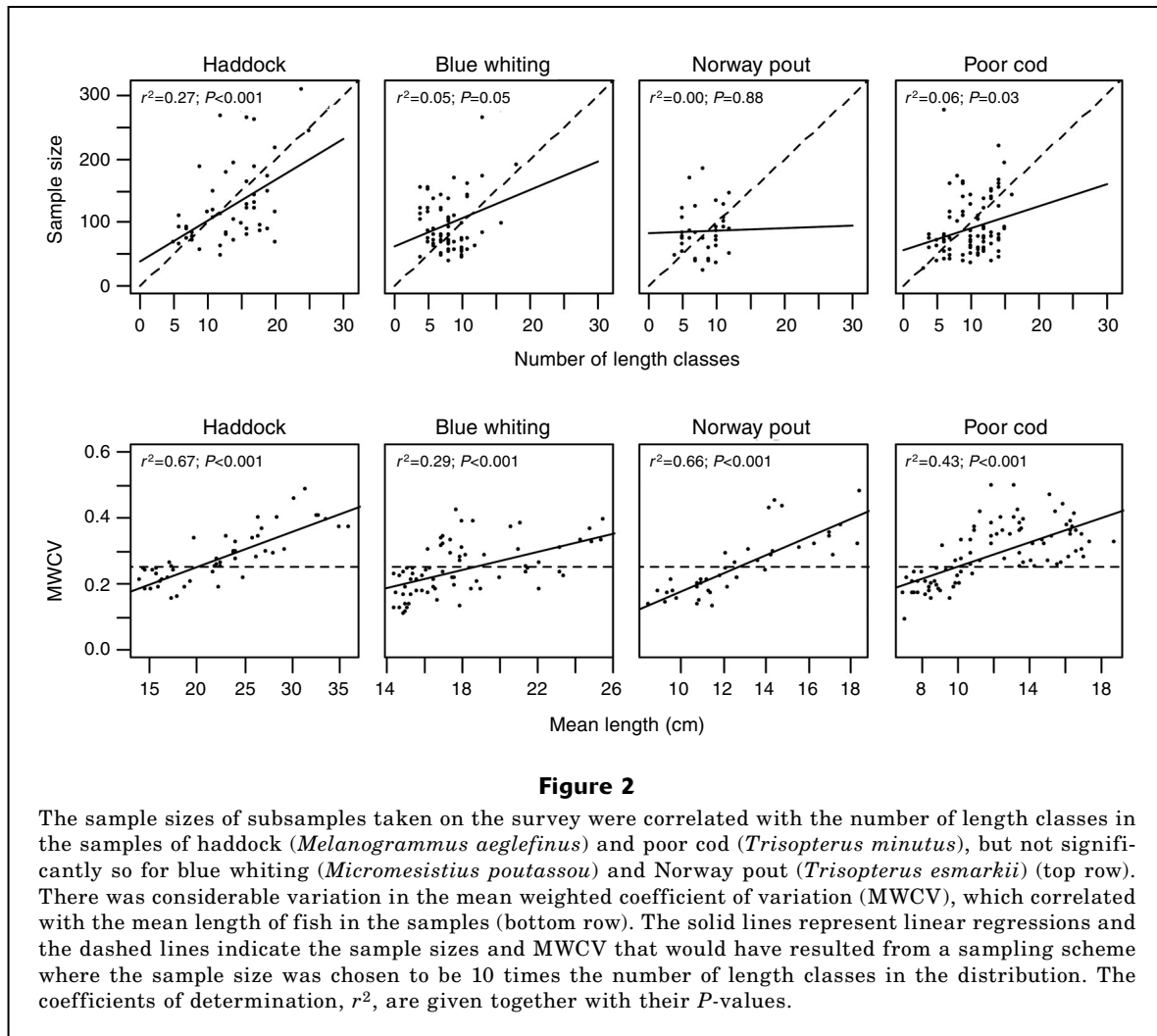
Discussion

Length distributions that result from combining a number of different samples exhibit greater variation than predicted under the multinomial model given in Equation 1 (Smith and Maguire, 1983). Fish populations are usually not uniformly mixed; therefore individual samples are not random samples from the population (Pennington et al., 2002). The simple multinomial model does not take account of the between-sample variability and will therefore underestimate the total variance. However, Equation 1 does provide an unbiased estimate of the variability within each sample, which is the variability that would occur if one could repeatedly take a random sample at the same location and time and measure these without error. This is the variability that is of interest when deciding whether the sample size is large enough to estimate the length distribution from a particular haul with a certain precision. Therefore, the MWCV is a suitable measure for this exercise.

In order to obtain a precise population estimate, it is important to maximize the number of sampling locations because of the considerable between-sample variability that is usually present (Pennington et al., 2002). Pennington et al. (2002) suggested maximizing the number of sampling locations at the expense of the number of fish measured. However, the number of hauls is often limited by practical considerations, and length measurements can be obtained quickly and cheaply. Therefore, it seems prudent to sample enough fish from each haul to obtain a length distribution that is representative of that catch at that particular location. Detailed information on the length distribution at each station can be valuable for exploratory data analysis, such as investigating the spatial structure in the data. Nevertheless, this level of sampling may not be strictly necessary for a precise population estimate of the length-frequency distribution for an age- or length-based assessment.

The samples in Figure 1 included a large range of species and size categories of fish, but the variability in the MWCV was small after taking account the sample sizes. This small amount of variability indicates that the MWCV is not very sensitive to the exact shape of the distribution and can be predicted with high precision, at least within the range of distributions encountered on the survey. A minimum sample size of 10 times the number of length classes in the sample appears to be a reasonable compromise between effort and precision in the present case.

The current analysis has focused on subsampling during surveys; however the same principles can be applied to any process of collecting data for which the shape of the distribution is of interest. The desired precision level for these cases will depend on a number of factors. For certain species that are of little commercial or scientific interest, but which may span across a large number of length classes, the suggested sample size of 10 per length class may be excessive. Likewise, as the MWCV is directly proportional to the number of length



classes in the sample, the choice of the interval of the length classes will determine the precision. Although increasing the size of length intervals will reduce the MWCV, this action will result in a loss of information which is undesirable. The cost of sampling, the detail required, and the purpose of the data collection need to be considered before the required precision level can be determined for applications other than the present example.

Without formal guidance on the appropriate sample size, the sample sizes chosen were, at best, weakly correlated with the number of size classes in the samples. It appears that the samplers under-estimated the required sample size for samples with large fish, whereas samples of smaller fish of the same species were over-sampled. This tendency to under-estimate the sample size may be related to the fact that the volume of a sample increases with the cube of its mean length; therefore a sample size of large fish may appear to be larger than the same number of small fish. In addition, samples with large fish tend to be spread out over a

larger number of size classes, thus requiring higher sample numbers.

In practice, it will be difficult for a sampler to estimate both the number of size classes and the number of fish in a sample. Therefore, the Marine Institute in Ireland is developing a software application that allows samplers to examine the length frequencies of the samples directly after they have been measured. The software estimates the weight of the suggested sample size for each distribution. Because size distributions tend to be similar on consecutive hauls, the sampler can gain an insight into the required weight of an appropriate sample for each species and size category.

The information contained in a length-frequency distribution is largely a function of sample size. The present method allows the amount of information contained in a length-frequency distribution to be quantified in terms of precision, allowing samplers to make informed decisions on the sample size that is required to obtain an adequate estimate of the length-frequency distribution of a particular catch.

Acknowledgments

Free R-software has been used for this work and the authors would like to thank the R development core team and all contributors to the R project (<http://www.R-project.org>). We also thank Colm Lordan and two anonymous reviewers for their valuable comments.

Literature cited

- Anderson, R. O., and R. M. Neumann.
1996. Length, weight, and associated structural indices. *In* Fisheries techniques, 2nd ed. (B. R. Murphy and D. W. Willis, eds.), p. 447–482. Am. Fish. Soc., Bethesda, MD.
- Erzini, K.
1990. Sample size and grouping of data for length-frequency analysis. *Fish. Res.* 9:355–366.
- Pennington, M., L. M. Burmeister, and V. Hjellvik.
2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. *Fish. Bull.* 100:74–80.
- Smith, S. J., and J. J. Maguire.
1983. Estimating the variance of length composition samples. *In* Sampling commercial catches of marine fish and invertebrates (W. G. Doubleday and D. Rivard, eds.), p. 165–170. Can Spec. Pub. Fish. Aquat. Sci. 66, Ottawa, Canada.
- Thompson, S. K.
1987. Sample size for estimating multinomial proportions. *Am. Statist.* 41:42–46.
- Vokoun, J. C., C. F. Rabeni, and J. S. Stanovick.
2001. Sample-size requirements for evaluating population size structure. *N. Am. J. Fish. Manage.* 21:660–665.
- Zar, J. H.
1999. *Biostatistical analysis*, 4th ed., 663 p. Prentice-Hall, Inc., Englewood cliffs, NJ.