Chapter 5
LENGTH FREQUENCY ANALYSIS

When performing length frequency analyses <u>within</u> a state/mode/wave stratum, reweighting of the data is not necessary assuming that the landed catch within a stratum was randomly sampled and that the true size distribution is represented. However, if length frequency analyses are performed <u>among</u> state/mode/wave strata, data must be reweighted prior to pooling. Reweighting of data is necessary due to non-random sampling of the catch among strata. Distribution of sampling among state/mode/wave strata is not representative of the true size distributions due to the following factors: 1) heavier sampling of the boat-mode catch; 2) variation in sampling levels among states; 3) variation in sampling levels among waves; and 4) differential success in obtaining fish length measurements. Variability in the proportion of fish measured for length among states and particularly among modes of fishing is demonstrated through a comparison of the distribution of measured bluefish by state and mode of fishing in the mid-Atlantic subregion for Wave 4, 1992 of the intercept survey (Figure 16) with the distribution of total estimated catch (Figure 17). Higher proportions of fish lengths are obtained in the private/rental boat mode. Weighting of length frequency data by the total estimated landings will adjust for the effects of non-random sampling of the catch across strata.

Number of Fish Lengths Obtained
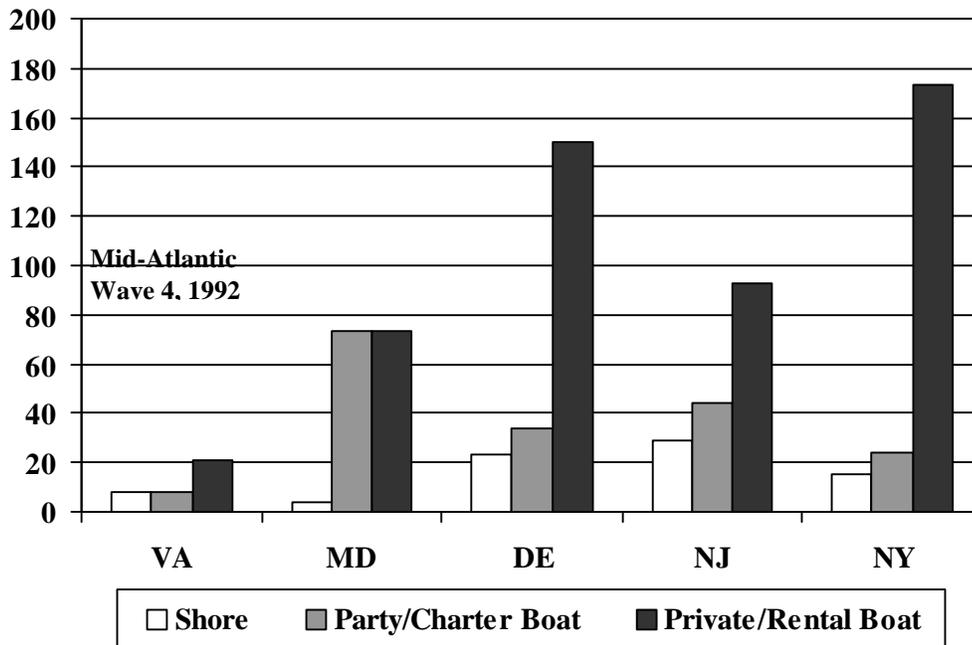
Figure 16.   Distribution of measured bluefish by state and mode of fishing in the mid-Atlantic subregion during Wave 4 of the 1992 MRFSS intercept survey.

Estimated Number of Fish Landed (thousands)
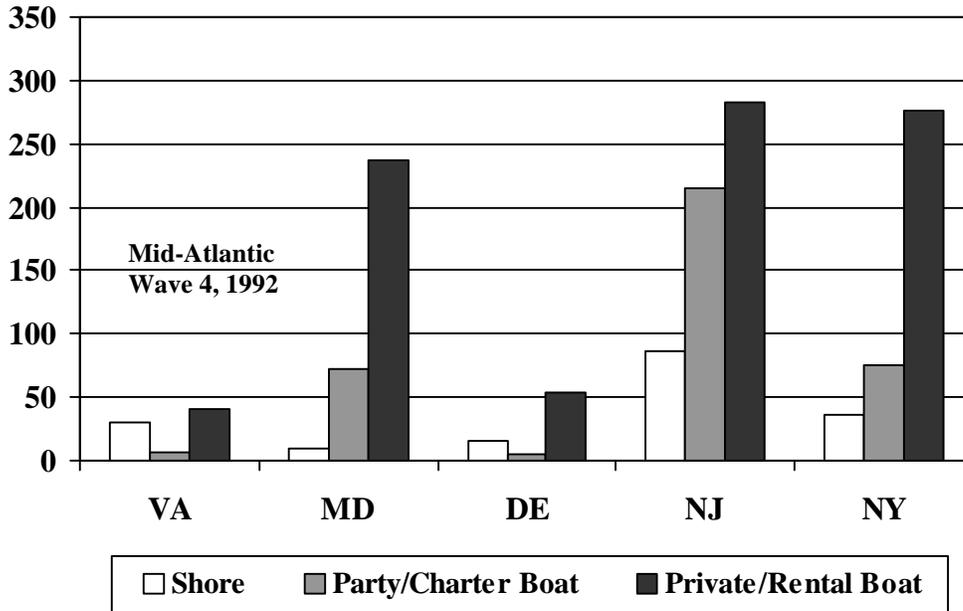
Figure 17.   Distribution of bluefish landings by state and mode of fishing in the mid-Atlantic
                subregion during Wave 4 of the 1992 MRFSS intercept survey.

        To estimate the number of fish landed in a given size class for pooled state/mode/wave strata,
the number of fish landed and measured in a given size class and state/mode/wave stratum is weighted
by the estimated total number of fish landed in the given state/mode/wave stratum prior to summing
across pooled strata as follows:

$$g_{S=X} = \sum_{ijk} \frac{l_{S=X\ ijk} * g_{ijk}}{\sum_{S} l_{ijk}}$$

where
i	= state
j	= mode
k	= wave
S	= size class
X	= range of sizes
$\gamma_{S=X}$	= estimated number of fish landed of size X over all pooled state/mode/wave strata

γ_{ijk} = estimated number of fish landed in a given state/mode/wave stratum

$l_{S=X\ ijk}$ = number of fish landed and measured of size X in a given state/mode/wave stratum

Σl_{ijk} = total number of fish landed and measured in a given state/mode/wave stratum

Computational Steps:

Pooling among states for length frequency analyses is demonstrated using hypothetical MRFSS data for bluefish landed in Wave 3 in the private/rental boat mode for the states of North Carolina and South Carolina. The number of bluefish landed and measured with a length of 300-400 mm ($l_{S=300\text{-}400\ ijk}$), the total number of bluefish landed and measured pooled over all size classes (Σl_{ijk}), and the estimated total catch of bluefish for each state/mode/wave strata (γ_{ijk}) are shown in Table 6.

Table 5.   Hypothetical MRFSS data for North Carolina and South Carolina for bluefish landed in Wave 3 for the private/rental boat mode, including data on the number of fish sampled in the state/mode/wave stratum having lengths between 300-400 mm, the total measured sample of bluefish in each state/mode/wave stratum, and the estimated total catch of bluefish in the state/mode/wave stratum.

State	Mode	Wave=3	Sample, 300-400 mm	Total Measured Sample	Catch Estimate
NC	PR	May/Jun	30	300	50,000
SC	PR	May/Jun	40	100	30,000

The <u>incorrect</u> procedure for estimating the total number of bluefish landed having lengths between 300-400 mm pools North Carolina and South Carolina data without weighting as follows:

$$g_{S=300\text{-}400} = \frac{(l_{S=(300\text{-}400)NC,PR,3} + l_{S=(300\text{-}400)SC,PR,3}) \times (g_{NC,PR,3} + g_{SC,PR,3})}{(l_{NC,PR,3} + l_{SC,PR,3})}$$

$$g_{S=300\text{-}400} = \frac{(30+40) \times (50{,}000 + 30{,}000)}{(300 + 100)} = 14{,}000$$

The <u>correct</u> procedure for estimating the total number of bluefish landed having lengths between 300-400 mm for pooled data over North Carolina and South Carolina is as follows:

$$g_{S=300\text{-}400} = \frac{(l_{S=(300\text{-}400)NC,PR,3} \times g_{NC,PR,3})}{l_{NC,PR,3}} + \frac{(l_{S=(300\text{-}400)SC,PR,3} \times g_{SC,PR,3})}{l_{SC,PR,3}}$$

$$g_{S=300\text{-}400} = [\frac{(30 \times 50{,}000)}{300}] + [\frac{(40 \times 30{,}000)}{100}] = 17{,}000 \ fish \ landed \ within \ length$$
$$interval \ between \ 300 \ and \ 400 \ mm$$

Relative catch frequency for 50 mm length classes for landed recreational catch of bluefish (Figure 18) and weakfish (Figure 19) in the mid-Atlantic subregion Wave 4 was calculated using the incorrect unweighted estimation procedure and the correct weighting procedure, demonstrating the differences between the relative frequencies for length classes obtained using the correct and incorrect estimation methods.

Relative Frequency

Mid-Atlantic Subregion
All Modes
Wave 4, 1992

□ Weighted ■ Unweighted   Length Class (mm)
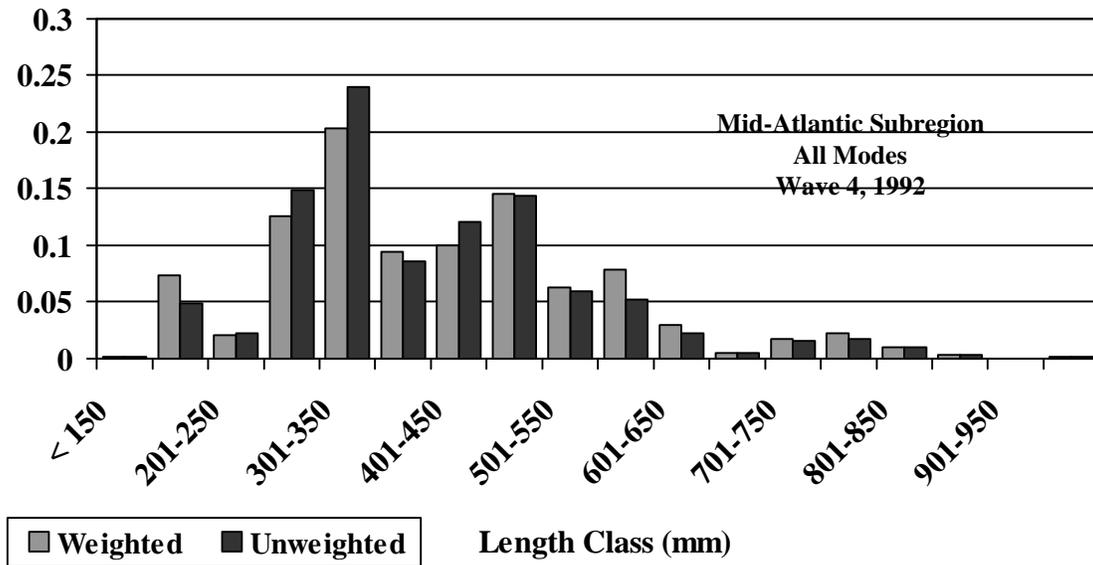
Figure 18.   Length-frequency distribution of bluefish in the mid-Atlantic subregion during Wave 4 of the 1992 MRFSS. Relative frequency was calculated using an unweighted and weighted length frequency analysis for pooling among states and modes of fishing.

Figure 19.   Length-frequency distribution of weakfish in the mid-Atlantic subregion during Wave 4 of the 1992 MRFSS. Relative frequency was calculated using an unweighted and weighted length frequency analysis for pooling among states and modes of fishing.

Relative Frequency

Mid-Atlantic Subregion
All Modes
Wave 4, 1992

□ Weighted ■ Unweighted   Length Class (mm)
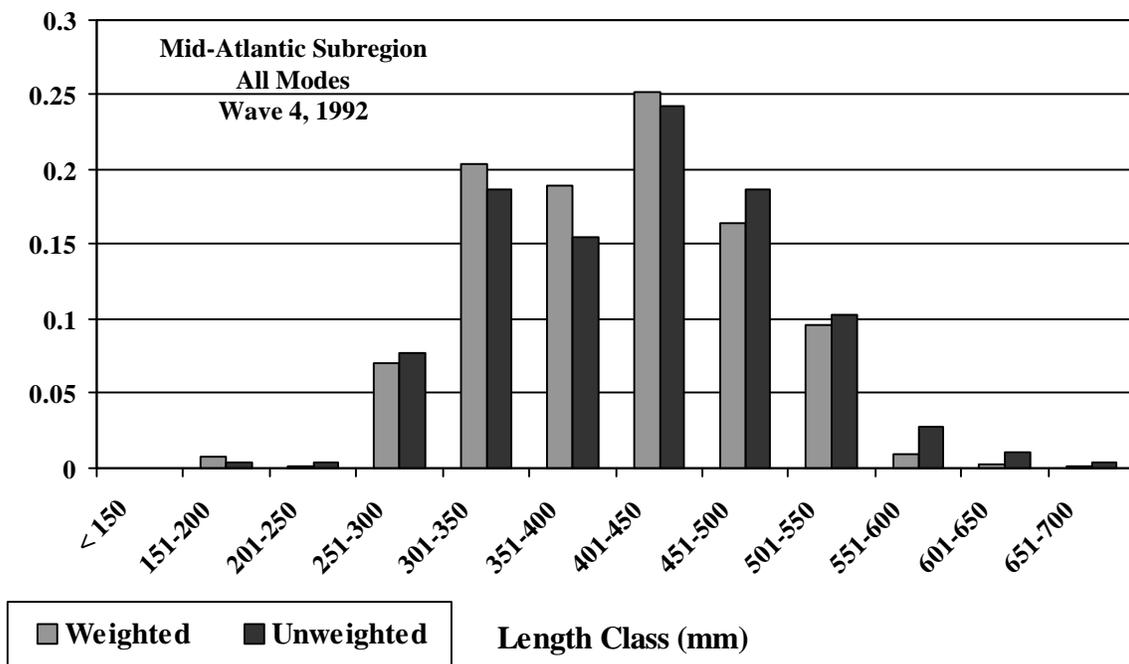
The specific steps in analysis of fish length measurements using the MRFSS SAS datasets are as follows:

1.   Delete records with missing length data.

    a.   From the Type 3 records, select records with the given SP_CODE.
    b.   From the Type 3 records, drop the records with missing length measurements (LNGTH=.)

2.   Merge the Type 3 data with the Type 1 records.

3.   Calculate length frequencies by state/mode/wave strata.

4.   Calculate the total number of records by state/mode/wave strata.

5.   Merge the length frequencies with the total number of records by state/mode/wave strata.

6.   From the catch estimate files, select the records for the given state/mode/wave stratum to be pooled and calculate estimated landings (ESTLAND = ESTCLAIM + ESTHARV).

7.   Merge the estimated landings with the relative length frequencies by state/mode/wave strata.

8.   Calculate estimated landings for each length class by state/mode/wave strata, by multiplying each relative length class frequency times the estimated landings (ESTLAND).

9.   Sum the estimated landings for each length class over state/mode/wave strata to be pooled.