

SEFSC computation of variance estimates for custom data aggregations
from the Marine Recreational Information Program

Kyle Dettloff, Vivian M. Matter, and Matthew Nuttall

SEDAR68-DW-10

11 March 2020



This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

Please cite this document as:

Dettloff, Kyle, Vivian M. Matter, and Matthew Nuttall. 2020. SEFSC computation of variance estimates for custom data aggregations from the Marine Recreational Information Program. SEDAR68-DW-10. SEDAR, North Charleston, SC. 2 pp.

SEFSC computation of variance estimates for custom data aggregations from the Marine Recreational Information Program

NOAA Fisheries
Southeast Fisheries Science Center
Fisheries Statistics Division
75 Virginia Beach Drive
Miami FL 33149

Kyle Dettloff, Vivian M. Matter, and Matthew Nuttall

03-11-2020

Coefficient of variation (CV) estimates for Marine Recreational Information Program (MRIP) survey catch totals are provided for stock assessments by the Southeast Fisheries Science Center (SEFSC). General MRIP survey design information is detailed in Matter and Nuttall 2020 and associated references. CV estimates for custom data aggregations (e.g. by year and mode within a specified sub-region) for each species have historically been provided as sums of variances of individual cells computed at the year, wave, sub-region, state, mode, and area fished level. Variances of individual cells, however, are only additive if each cell is computed at the level of a stratum identifier. This is not the case for MRIP intercept data, which is collected from a survey design stratified by *year, wave, sub-region, state, and mode*, but not *area fished*. As such, variances must be computed directly from the raw data to obtain CVs appropriate for custom aggregations.

Variances of total catch estimates are now computed using standard survey methods for a complex multistage design using the Taylor series (linearization) method (Woodruff 1971). Details of statistical computations are outlined in the SAS PROC SURVEYMEANS user manual (SAS Institute 2019).

Traditionally, observations in strata with a single primary sampling unit (PSU) are excluded from variance calculations. This method, however, can lead to zero-variance estimates if the selected subset of strata

each contain only a single PSU. An adjustment to this approach that includes the contribution of “lonely” PSU strata to the overall variance estimate centers these observations at the sample grand mean rather than the stratum mean. This adjustment has been previously recommended to the MRIP Oregon Recreational Boat Survey to handle lonely PSU cases (Starcevich 2013) and is equivalent to setting `options(survey.lonely.psu = "adjust")` in the R *survey* package (Lumley 2019). It is known that this method tends to produce conservative variance estimates, but it is preferable to excluding these cases entirely to avoid anti-conservative or zero-variance estimates. It should be noted that the impact of this adjustment tends to be very minor at the year/subregion/mode level, as strata with only a single observation at this level of aggregation are relatively uncommon in the MRIP intercept data.

References

- Lumley, T. (2019). “survey: Analysis of Complex Survey Samples.” R package version 3.35-1.
- Matter, V.M. and M.A. Nuttall. (2020). “Marine Recreational Information Program Metadata for the Atlantic, Gulf of Mexico, and Caribbean Regions.” NMFS Southeast Fisheries Science Center, Miami, FL. SEDAR 68-DW-13. 14pp.
- SAS Institute Inc. (2019). SAS/STAT® 15.1 User’s Guide. The SURVEYMEANS Procedure: Statistical Computations. Cary, NC. Updated December 13, 2019.
https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_surveymeans_details06.htm&docsetVersion=15.1&locale=en
- Starcevich, L.A.H. (2013). “Design Based Analysis for the Oregon Recreational Boat Survey. Weighting Approach and Analysis Methods.” Oregon Department of Fish and Wildlife. 30pp.
- Woodruff, R.S. (1971). “A Simple Method for Approximating the Variance of a Complicated Estimate.” *Journal of the American Statistical Association* **66**(334): 411-414.