FISH 621 Homework #3: Stratified Sampling and CPUE

Curry Cunningham 2022

Instructions

Please make sure to follow the instructions in the homework and, to the extent you can, feel free to **work with others to learn from each other**. If you get stuck on a problem and can't figure out how to proceed on your own, feel free to e-mail others in this class or your peers. For homework purposes, perhaps you can work together via e-mail, Zoom, Google Hangouts, or whatever works best for you. However, each of you will need to submit your own completed homework assignment for evaluation.

Please post to the Canvas Discussion Forum, and feel free to e-mail me if you get stuck and have exhausted your immediate options (google, classmates, friends). No need to bang your head against the wall, first ask your peers for help and if all else fails feel free to email me directly.

This homework assignment is due by 11:59 pm on Friday April 1, 2022.

Please submit all components of the homework assignment (i.e. word, R script, or Excel file) via Canvas, and name each file with the homework number and your first and last name (e.g. **Hwk3_FirstName_LastName.docx**, **Hwk3_FirstName_Last.xlsx**, **Hwk3_FirstName_LastName.R**). In the word document please describe your results and answer any questions posed in the homework document, along with any useful and/or necessary model output or figures. Please ensure that the R script you submit would allow anyone with the same input files to recreate your analysis, and include comments as necessary to guide review of your work.

In the event you are unable to submit your completed assignment via Canvas, please email it to me at: cjcunningham@alaska.edu. Late homework assignments will be penalized 2 points per day overdue. Please contact me ***ahead of time*** if there are circumstances requiring late submission.

Evaluation

This homework assignment will be graded out of ***40 possible points***.

Homework Contents

- **621_Homework 3.pdf** (this file)
- **Yukon Moose Survey.csv** Moose aerial survey data from the Yukon-Charley Rivers National Preserve, Alaska in 1997.
- **GOA BTS POP.csv** NOAA-NMFS Bottom Trawl Survey data for Pacific ocean perch in the Gulf of Alaska.

- **Stratum Descriptions.csv** Descriptions of NOAA-NMFS Bottom Trawl Survey strata, including the area of each necessary for generating design-based biomass indices.

Problem 1 – Stratified Random Sampling (8 points)

You are charged with estimating the abundance of wolves in a 450-km^2 region in interior Alaska. Based on your knowledge of the system you have identified a stratified random sample design as most appropriate, with three strata representing habitats of varying quality. Within each of these three strata you randomly sample 1-km x 1-km areas by air, counting the number of wolves within each sampling unit or area (y_{hi}). Stratum 1 is 100-km^2 , stratum 2 is 50-km^2, and stratum 3 is 300-km^2. Within each stratum you sample $n_h = 50$, 1-km^2 sampling units.

At the conclusion of your survey you calculate both the average abundance per 1-km^2 sampling unit ($\bar{y}_h$) and the sample variance s_h^2, for each strata.

The data for the three strata are as follows:

Stratum	1	2	3
Number of 1km^2 units sampled (n_h)	50	50	50
Average abundance per sampling unit ($\bar{y}_h$)	10	20	30
Sample variance for each strata (s_h^2)	2800	700	600

Please calculate the following quantities based upon your stratified random sample design:

1. An estimate of total wolf abundance within each stratum: $\hat{\tau}_h$.
2. The estimated wolf population total in the entire 450-km^2 region: $\hat{\tau}_{st}$.
3. An unbiased estimate of the variance of the wolf population total: $\widehat{var}(\hat{\tau}_{st})$.
4. The coefficient of variation for the estimate of the wolf population total.
5. An approximate 95% confidence interval for the population total.
6. An unbiased estimate of mean wolf abundance per 1-km^2 for the entire population: $\hat{\mu}_{st} = \bar{y}_{st}$
7. An unbiased estimate of the variance of the mean wolf abundance per 1-km^2: $\widehat{var}(\hat{\mu}_{st}) = \widehat{var}(\bar{y}_{st})$.
8. An approximate 95% confidence interval for the mean wolf abundance per 1-km^2.

Please summarize your results and inference in the word (.docx) document you submit for this assignment along with your R script.

Problem 2 - Moose Sampling (10 points)

For this problem we will utilize data from the 1997 aerial moose survey along the Yukon River corridor, Yukon-Charley Rivers National Preserve, Alaska: Project report, November, 1997 (Burch and Demma, 1997). Three strata were sampled by air counting the abundance of moose within sampling units in each stratum. The sampling units within each stratum were selected randomly, out of a total of possible $N_1 = 122$ units in stratum 1, $N_2 = 57$ possible sampling units in stratum 2, and $N_3 = 22$ possible sampling units in stratum 3.

Data from this moose aerial survey are contained within the ***Yukon Moose Survey.csv*** file associated with this homework assignment. Each row in the data frame are counts of moose in a given sampling unit. The columns in this data frame are the stratum number (`str`) and the number of moose counted (`moose`) in each surveyed sampling unit.

Begin by reading these data into R, and please calculate the following quantities:

1. The number of units sampled per stratum: n_h
2. The sample mean for each stratum: $\bar{y}_h$
3. An unbiased estimate of the mean number of moose per sampling unit, for the population (i.e. the population mean): $\hat{\mu}_{st} = \bar{y}_{st}$
4. The sample variance per strata: s_h^2
5. An unbiased estimate of the variance for the population mean abundance per sampling unit: $\widehat{var}(\hat{\mu}_{st}) = \widehat{var}(\bar{y}_{st})$
6. An approximate 95% confidence interval for the mean abundance per sampling unit
7. An estimate of total moose abundance within each stratum: $\hat{\tau}_h$
8. The estimated total moose abundance across strata: $\hat{\tau}_{st}$
9. An unbiased estimate of the variance of the total moose population: $\widehat{var}(\hat{\tau}_{st})$
10. An approximate 95% confidence interval for the moose population total

Please summarize your results and inference in the word (.docx) document you submit for this assignment along with your R script.

Problem 3 – Abundance Indices for GOA Pacific Ocean Perch (22 points)

Fishery-independent indices of abundance are important inputs for the assessment of stock status for many marine species. The NMFS Bottom Trawl Survey provides CPUE data from which we can construct indices of abundance for benthic species, in addition to important age and length composition samples.

Please find and load the ***GOA BTS POP.csv*** file which contains bottom trawl survey catch observations for Pacific ocean perch in the Gulf of Alaska, as well as the ***Stratum Descriptions.csv*** file which contains information on the survey strata, most importantly the area in km^2 of each strata.

Please address the following questions with figures and specific values or tables as appropriate:

1. Please begin by plotting the trend in total (i.e. the sum of) biomass cpue across years. Why would this be an inappropriate treatment of these data?
2. Next please plot the total number of stations sampled by year.
3. Please calculate the design-based (area-weighted) estimate of total Pacific ocean perch biomass across time for the Gulf of Alaska.
    a. For each year, calculate:
        i. The total biomass estimate
        ii. The variance of the biomass estimate
        iii. The standard deviation of the biomass estimate
        iv. The coefficient of variation for the biomass estimate
4. Finally, please calculate a model-based biomass index using a GLM of appropriate structure with `log(Weight.CPUE..kg.km2.+1)` as the response variable.
    a. Next, compare GLM index standardization models with and without stratum effects, plotting their respective trends and discussing the appropriateness of each.

Please summarize your results and inference in the word (.docx) document you submit for this assignment along with your R script.

Time Allocation

At the end of the word (.docx) document you submit for this assignment, please estimate the amount of time you spent in total on this assignment.