effective\_number\_strays\_note

Molly Payne

5/24/2022

The total effective number of strays column for a given stream-year in the new\_response\_var object (your new response variable as of May 2022) will not always equal the number of hatchery strays / proportion sampled for a stream-year because you set the proportions sampled that were greater than 1 equal to 1 instead. An example is Admiralty Creek 2009, which had 2 surveys:

survey\_data8 <- readRDS("survey\_data8.rds")  
survey\_data8[survey\_data8$StreamName == "Admiralty Creek" &  
 survey\_data8$Year == "2009", c(2,3,18:30)]

## StreamName SurveyDate DeadCount PreviouslySampledCount TotalCount  
## 1100 Admiralty Creek 2009-08-12 117 NA NA  
## 2100 Admiralty Creek 2009-08-17 14 NA NA  
## NumberofSpecimens NumberStrays PercentStrays numReadOK numMarked  
## 1100 96 39 NA NA NA  
## 2100 21 9 NA NA NA  
## numUnMarked SurveySamplers Proportion\_sampled Year Effective\_number\_strays  
## 1100 NA <NA> 0.8205128 2009 47.53125  
## 2100 NA <NA> 1.0000000 2009 9.00000

new\_response\_var <- readRDS("new\_response\_var.rds")  
new\_response\_var[new\_response\_var$StreamName == "Admiralty Creek" &  
 new\_response\_var$Year == "2009",]

## StreamName Year Total\_effective\_strays Number\_of\_surveys  
## 6 Admiralty Creek 2009 56.53125 2  
## Avg\_number\_strays  
## 6 28.26562

The proportion sampled in Admiralty Creek in 2009 would be the sum( Number of Specimens) / sum (Dead Count ) for all the surveys (n = 2 in this case) in Admiralty Creek in 2009. So, using the info in survey\_data8, this would be (96 + 21) / (117 + 14) = 117/131 = 0.89. Then, the total number of hatchery strays detected (39 + 9 = 48) should be divided by that proportion: 48 / 0.89 = 53.8

However, in new\_response\_var, the total\_effective\_strays in 2009 for Admiralty Creek is 56.5, not 53.8. This is because for the second Admiralty Creek 2009 survey, the NumberofSpecimens was greater than the DeadCount, resulting in a proportion sampled > 1. As per a decision made by Curry, Peter, and myself in April 2022, all proportions sampled > 1 were set equal to 1. So, the proportion sampled for the second Admiralty Creek 2009 survey was reduced to 1, resulting in the effective number of strays for the season being a little bit larger (bc proportion sampled is the denominator when calculating effective strays, thus a smaller denominator = large effective strays; 56.5 > 53.8)