salmon\_permit\_theory

Ben Buzzee

February 12, 2019

All data is taken from the permit\_records and harvest\_records sheets in the original excel document. First we will take note of the form of the data we have:

We have the total number of permits, and the mailing status of each:

|  |  |
| --- | --- |
| mailing | count |
| 0 | 17006 |
| 1 | 3269 |
| 2 | 2041 |
| 9 | 7268 |
| total | 29584 |

And of all those that reported fishing, (mailing status 0, 1, or 2), we have the number of fish they caught at each location they fished at:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| permit | fishery | red | king | coho | pink | chum |
| 2 | KASILOF | 40 | 0 | 0 | 0 | 0 |
| 4 | KENAI | 45 | 0 | 0 | 0 | 0 |
| 8 | KENAI | 16 | 1 | 0 | 0 | 0 |
| 13 | KASILOF | 1 | 0 | 0 | 0 | 0 |
| 14 | KENAI | 3 | 0 | 0 | 0 | 0 |
| 15 | UNKNOWN | 33 | 0 | 0 | 0 | 0 |

Our goal is to estimate how many fish were harvested from each location by the 7268 permit holders that never responded.

First, we observe that the total number of permits issued can be split into three groups:

where:

* is the number of compliant households that responded without a reminder or after the first reminder (mailing = 0 or 1)
* is the number of households that responded after the second reminder (mailing = 2)
* is the number of households that never responded. (mailing = 9)

We will assume that is a population of “noncompliant” households and is a random sample from said population. Further, note that , and the harvests from those households are fixed and known. We only need to estimate the harvest of households that never responded.

Since the households making up are a random sample, both and share the same population mean harvest and proportion of permit holders that fished. So we will use observations from to estimate those parameters. From the households making up , we will use the estimates:

* as the proportion of that participated in fishing: = where is the number of households responding to the second reminder (mailing = 2) that reported fishing.
* and as the average harvest at location **among all of those that fished** : over households and fixed location

By the central limit theorem,

follows a distribution with

and is with where

So the estimated number of non-respondents that fished is: and the estimated total harvest at each location by non-respondents is . Our estimated variance of the harvest estimate is then:

by Goodman (1960)

Finally, our estimated total harvest for each location would be where both and are known. Since they are both known, the total variance would be the same as above.