Sampling

fish5101fishsci Introduction to fish population dynamics

Gunnar Stefansson

December 19, 2016

Design issues

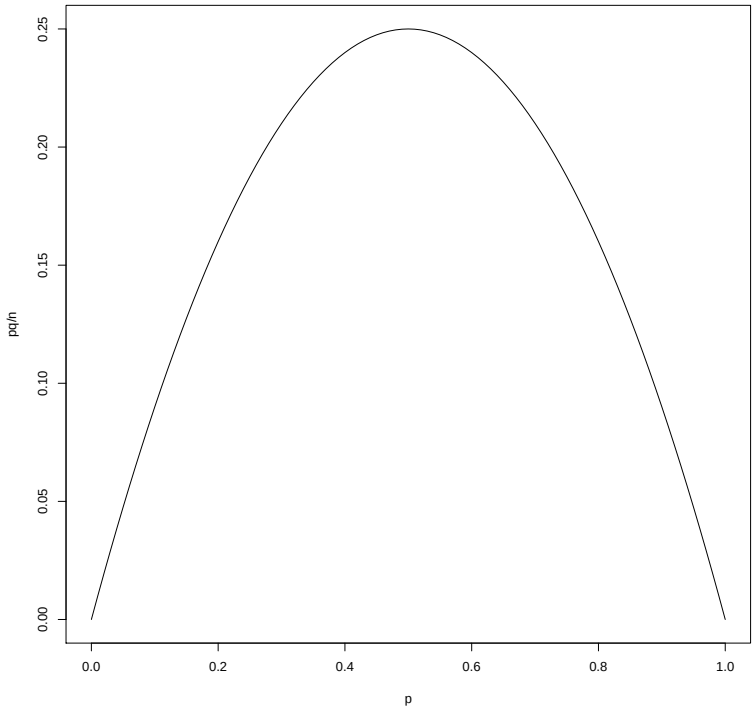


Figure : The variance vs mean relationship for an estimated proportion

Length distributions are basically count data, as are the number of fish at a given age if simple random sampling is used for age readings. In the case of simple random sampling from an entire population of fish, the number of fish in a given age or length cell should follow a binomial distribution.
This implies that the probability of obtaining exactly x fish in a given cell is given by binomial probabilities:

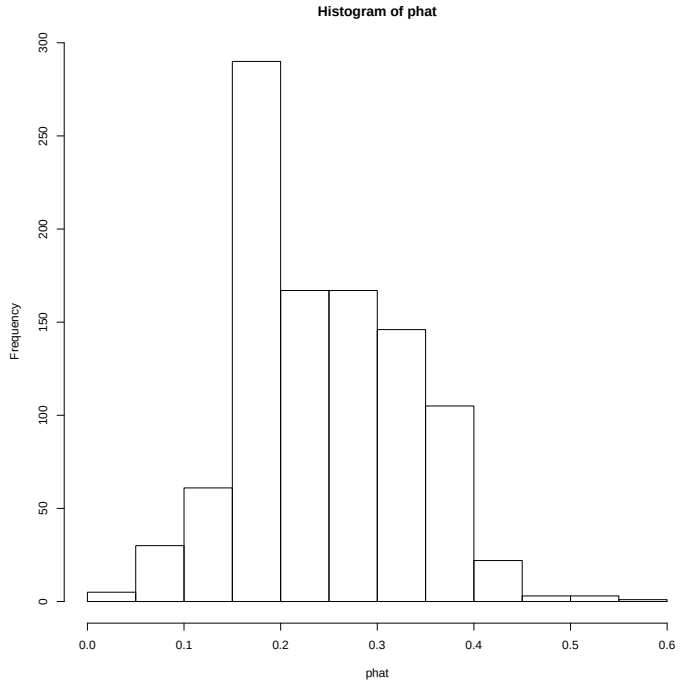
Simulating sampling schemes

Want to estimate a proportion p

Sample independently individual fish (in group=1; or not=0): get binomial distribution for x=number of positives

Estimated proportion: $\hat{p} = x/n$

Known expected value and variance: $E[\hat{p}] = p$, $V[\hat{p}] = np(1-p)$



Example: In R it is very easy to simulate the effects of different sampling schemes and there are several ways to do this. Consider the red/green marble experiment.
R code for the red/green marble experiment

```
n<-25
numred<-10
numgreen<-30
green<-rep("G",numgreen)
```

Correlation issues

When correlation issues occur the model is usually wrong

Primary issue: Intra-haul correlation, cf Pennington and Volstad

Fish within a station s are more similar than across stations.

A model of the process needs to include a (random) station effect: $y_{sj} = \mu + \alpha_s + \epsilon_{sj}$

Here, y_{sj} could be the length of fish j at station s.

The resulting correlation between fish at the same station is the **intra-haul** correlation