On Estimating Ion Channel Densities In Model Neurons From Simulated Patch Clamp Data

Sean G. Carver, Ph.D. carver@american.edu American University, Washington, DC 478.10 (B6)

Problem Statement

How well can ion channel densities be inferred from patch clamp data? Likelihood-based methods are statistically consistent, meaning the correct answer will become increasingly certain as the sample size increases. But how large does the sample size need to be? We solve this problem for an oversimplified "toy model" aiming to develop software applicable to more complicated problems.

Toy Model

The toy model consists of an ensemble of n independent and identical channels, each residing in one of two states: open or closed. The patch clamp unambiguously determines k_i, the number of open channels, at each of the r times that current is sampled. The sample interval is long, so we consider the samples independent. The probability of seeing each channel in the open state is p. Thus the measurements have Binomial distribution:

$$k_i \sim B(n,p) \text{ for } i = 1, 2, \ldots, r$$

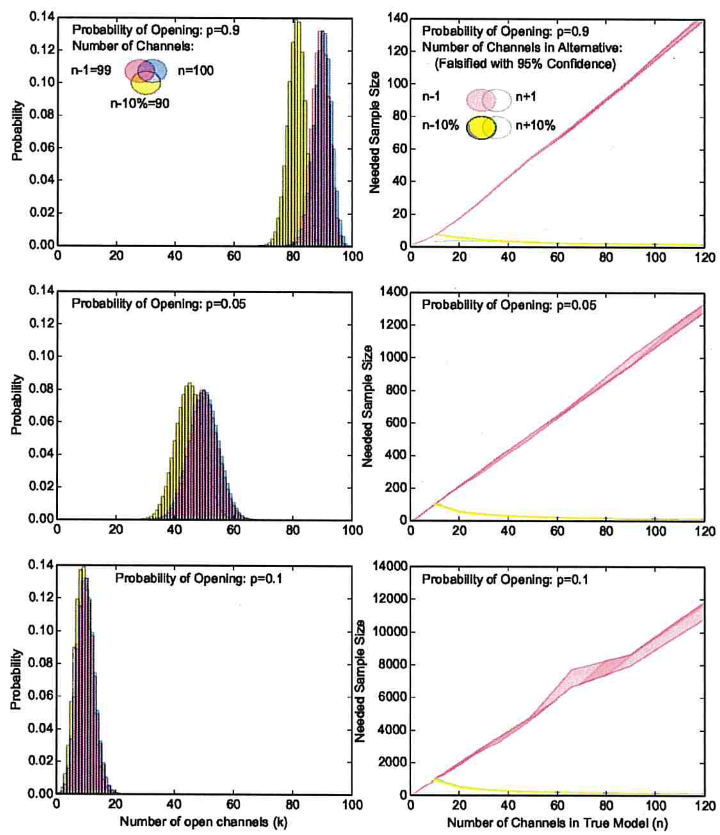
The problem is to estimate n observing $\{k_i\}$ and knowing p.

Relation to Previous Work

Our toy problem has been studied classically. The statistics literature includes a relevant paper by Student and another by Fisher, and more recent papers by other researchers including neuroscientists working in the context we consider here. These papers contribute alternative estimators for n, a hard problem when p is small, even if known. Maximum likelihood performs particularly poorly, showing wide variation of the estimate, for small perturbations of data. Our work considers the related problem of quantifying the difficultly of model selection between two simple alternatives. How many samples are needed to make the selection with 95% confidence?

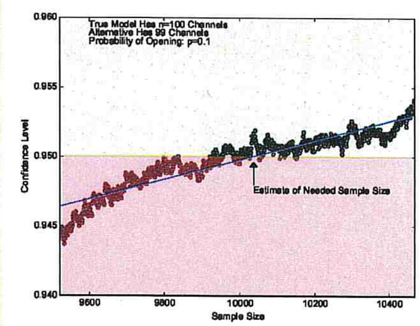
Results

The left column shows probability histograms for k with $n = 100$ (blue), $n = 99$ (red) and $n = 90$ (yellow). Overlap is shown as mixed color. The right column shows ± 1 standard error (using 4 calculations), as a function of n, estimating the number of samples needed to confidently falsify an alternative with $n-1$ (red), $n+1$ (grey), $n - 10\%$ (yellow) and $n + 10\%$ (green). Note the differing vertical scale.



Sample Size for 95% Confidence

Monte Carlo simulations estimate the confidence level of the likelihood-ratio test over a wide range of sample sizes. A cumulative sum of log-likelihood-ratios makes this computation efficient. We increase sample size until the confidence threshold is exceeded, repeating the calculation to attain precision. Bootstrapping and/or caching likelihoods can greatly reduce computation times. Sample size vs. estimated confidence remains noisy so we estimate the needed sample size as shown:



Conclusions and Future Work

Our results suggest that the amount of data needed to falsify an alternative differing from the true model by an absolute number of channels (e.g. 1) grows approximately linearly as n increases. The slope of this line increases with decreasing p. On the other hand, if the difference between the models is relative (e.g. 10%) the amount of data needed for 95% confidence reaches, surprisingly quickly, a floor of just one sample.

We aim to develop software tools to tackle related problems for more complicated models. How much more data would Hodgkin and Huxley have needed if they had used a current clamp instead of a voltage clamp? What is better—improving the measurement noise characteristics of your recordings, or their temporal resolution? With tools in place, students and researchers will be able to easily formulate and answer novel questions within the vast space of possible problems.