Open Science Framework DOI:10.31219/osf.io/***

Notes on inferring connectivity (Bente's problem) [draft]

P.G.L. Porta Mana <piero.mana@ntnu.no>

17 October 2019; updated 17 October 2019

Notes on Bente's problem about inferring connectivity. This is just a memo, not meant to be detailed or precise.

1 Synopsis of the problem

We're considering neurons in a specific region A of the mouse brain, and neurons in a specific region B which connects to those in region A.

The question: if we examine a new neuron in region A in a new mouse, how many neurons in region B will project to it? We want to calculate the probabilities of the possible numbers. Call this number c. A slightly more detailed question is this: in the long run, how often will we observe neurons in region A with 0 or 1 or 2 or 3 or... input neurons from region B?

It seems that the analysis will also yield probabilities for other another quantity: how many neurons are there in region B, in the next mouse we observe?

The observations and data we use: in several mice a small number of neurons in region A – *starter neurons* – are infected with a virus. This virus spreads, with time, to neurons – *input neurons* – that projects to the starter neurons. The virus can't spread further than that. Some of the infected neurons will be in region B. The virus also leads to colouring of starter and input neurons. Each mouse is killed after several days (different for each mouse) and the starter and input neurons are counted. So our data are the number of starter and input neurons for a number of mice.

We have this problem and data not just for one input region, but several. These joint data can be used to get more accurate inferences for each input region.

Complication: if the mouse is killed too early, the virus hasn't had the time to spread to all input neurons. If the mouse is killed too late, the starter cells decay and are no longer visible. So we also need to guess, for each observed mouse, whether there were additional starter neurons, no

longer visible, and what proportion of input neurons have been reached by the virus. We have, for each mouse, the number of days between virus injection and killing.

2 Strategies for solving the problem

It may be useful to give a first answer without the time-lapse complication, and add that complication afterward.

The core of the problem is trying to infer how many input neurons project to a starter neuron in the case in which we have several starter neurons. An input neuron can project to several starter neurons, so we can't just divide the number of input neurons by the number of starter neurons.