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>

B. Jacobsen
<bente.jacobsen@ntnu.no>

17 October 2019; updated 16 January 2020

Notes on Bente's problem about inferring connectivity.

1 Introduction

We are interested in quantifying the proportion with which several brain regions project to others. It's impossible, at least at present, to simply count individual projections. What's done experimentally, with very clever techniques, is analogous to survey sampling: we manage to see, more or less clearly, a small sample of projections. From this sample we want to infer the larger picture, for a single brain and on average for a 'typical' brain. We show here how to make such inferences using the probability-calculus[1], also known as Bayesian theory.

We accompany the method with an extensive discussion of its steps and assumptions, for the sake of researchers unfamiliar with it. But we believe that some of the points of view will be of interest also to experts in the method.

2 Synopsis of the problem

We're considering two sets of neurons, which we call *target* set and *input* set. The neurons in the target set are of several classes, for example parvalbumin or somatostatin interneurons; and are distributed across several regions, for example medial septal complex or thalamus). Same holds for the input set. In general we are interested in the number of projections of input neurons to target neurons, and in how such projections are distributed across regions, or classes, or both, of neurons. More specifically we focus here on the following kind of question: How many input neurons in a region R (independently of their class) project

[1] Jaynes 2003; Hailperin 1996; Sox et al. 2013.

This document is designed for screen reading and two-up printing on A4 or Letter paper

to target neurons of class C (in a specific location)? And we want to compare the answer for several combinations of input regions R and target classes C, thus obtaining the proportion for each combination.

But the question above is still imprecise. Are we asking about a specific mouse? or about the average across all mice having some specific characteristics? These two cases can have slightly different answers, and their difference is important for the application of the probability-calculus and for experimental verification. We'll focus on these two kinds of question:

Individual: If we were able to perfectly count, in the next mouse to be examined, the number of input neurons in region R projecting to target neurons of class C, – what number would we find?

Collective or average: If we were able to perfectly count, for each mouse (with some specific traits such as age or gender and so on) to be examined in the future or that could have been examined in the past, the number of input neurons in region R projecting to target neurons of class C, – what average number would we find?

It's intuitively clear that the answer to the second question comes from the combination of answers to the first, for many mice.

The observations and data we use: In a mouse, virus methodologies allow us to select several neurons in the target set and to collectively observe which neurons in the input set project to them. We can't distinguish, however, which input neuron projects to which target neuron. Such observations are made in several mice.

There's a complication: to make this kind of observation the mouse must be killed. If it's 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 neurons 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. As data we have the number of days between virus injection and killing, for each mouse. This complication for the moment is set aside to simplify the method, but will be taken into account later.

Now let's sketch a schematic picture of the strategy to answer the two questions above.

3 Formulation and strategies

3.0 Preliminary remarks

Our inference, based the probability-calculus, boils down to these steps:

1. Find a set of simple, factual statements, in terms of which we can formulate our questions and our data. These are called 'atomic statements'; atomic in the sense that we won't analyse them into more precise statements.
2. Assign a joint probability distribution for all the possible combinations of these atomic statements. This probability distribution is called a 'pre-data' or 'prior' distribution, because it depends on our background information and assumptions, denoted by I, but not yet on the experimental data.
3. Express our data and the possible answers to our questions in terms of the atomic statements.
4. Use the probability-calculus to determine, from the pre-data distribution of step 2., the probability for each answer of interest, conditional on the data observer and on our pre-data information:

$$\mathrm{p}(\text{answer} \mid \text{data}, I). \tag{1}$$

   The probabilities for all possible answers form the so-called 'post-data' or 'posterior' distribution.

Why do we need a pre-data distribution? can't we find the final probability from the data alone, without background information and assumptions? – No, we can't; it's a logical impossibility. Just like in the truth-calculus (formal logic) we can't prove any theorem if we don't specify some initial axioms or postulates, so in the probability-calculus we can't derive any probabilities without some initial 'probabilistic postulates'[2]. *All* statistical and probabilistic methods use assumptions. 'The main difference is that in a non-Bayesian analysis more is swept under the carpet. This makes non-Bayesian methods politically expedient. The Bayesian is forced to put his cards on the table instead of up his sleeve. He thus helps others to improve his analysis, and this is always possible in a problem concerning the real world.'[3].

[2] Hailperin 2011; Johnson 1924 p. 182. [3] Good 1969 § 2.3 p. 26.

[from here on: old text]

The first two steps are especially important: it's there that we decide on the level of simplification or realism of our formulation. Let's make a couple of remarks about this.

The possible formulations of our inference problem differ in their degree of simplification or realism, in the kinds of questions they allow us to answer, and in their mathematical complexity. Their degree of realism is, in the present case, tightly connected with the *hierarchic* depth at which we analyse our problem. What do we mean by 'hierarchy'? Here are two examples.

(i) For each pair of input-target kinds, we could study the similarities and dissimilarities (expressed as, say, means and variances) of the connections from the input kind to *one* target neuron. Then study how these similarities and dissimilarities are themselves similar and dissimilar across all target neurons of the same kind in *one* mouse. And finally study how these higher-order similarities and dissimilarities are themselves similar and dissimilar across all mice.

(ii) For each pair of input-target kinds, we could study the similarities and dissimilarities of the connections to all target neurons of the same kind in all mice at once, pooling across target neurons and mice.

The first example is more realistic because it acknowledges different levels of similarities and variations in our study. It also allows us to ask specific questions, such as 'how does the connectivity vary from mouse to mouse?'. It leads, however, to very complex maths (integrals over function spaces). The second example is drastically simplified and doesn't allow us to ask very specific questions. But it leads to much simpler mathematical expressions. Other approaches are also possible in between and outside of these two.

The number of hierarchies goes hand in hand with the number of *symmetries*. In the first example, our analysis treats in a similar – symmetric – way all input neurons to a specific target neuron, but not across target neurons; then it treats in a symmetric way the statistics of the connectivity of every target neuron in a mouse, but not across all mice; and so on. In the second example, our analysis treats all target neurons in all mice in a symmetric way. We'll shortly see how these symmetries have a mathematical counterpart and determine the formulae we use.

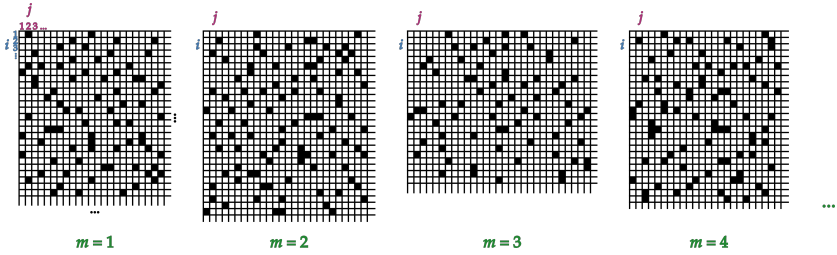

Figure 1 Imaginary layout of the connectivity from input to target region for all mice

3.1 Step 1: basic statements

We consider a very large number of mice: those we have observed and all those that we might observe in the future or we could have observed in the past. For our inferences to be valid, these mice must be judged 'similar' in some respects. Let's label each mouse with an index $m = 1, 2, \ldots$. For each mouse, we consider all neurons constituting its input region, labelling them $i = 1, 2, \ldots, K_m$, and all neurons constituting its target region, labelling them $j = 1, 2, \ldots, L_m$. The values of the three labels i, j, m may be arbitrary or carry some information. For example, m could reflect a temporal order, and i, j could each reflect a topographical order. The number K_m of neurons in the input region and the number L_m in the target region are different for each mouse m.

Now that we've labelled the mice and their neurons in the two regions, we can consider specific statements about their connections, for example

> *In mouse $m = 3$, the neuron $i = 5$ in the input region projects to the neuron $j = 4$ in the target region*

which can be either true or false. The situation described by these statements is visualized in fig. 1. We have a grid for each mouse m; each row i in the grid represents a neuron in the input region, each column j a neuron in the target region. If in mouse m neuron i projects to neuron j, then we colour the corresponding entry i, j in grid m in black, or we can associate the number 1 with it; if it doesn't project, then we use white or the number 0.

Mathematically we represent these statements and the grids of fig. 1 with the following quantities:

$$c_{ijm} = \begin{cases} 1 & \text{if neuron } i \text{ projects to neuron } j \text{ in mouse } m, \\ 0 & \text{if neuron } i \text{ doesn't project to neuron } j \text{ in mouse } m. \end{cases} \tag{2}$$

Let's call them *connections*.

For most of these statements we'll obviously never be able to ascertain whether they're true or false. But they serve us well as basic statements: we can precisely formulate our questions and data in terms of them.

The formulation just proposed is already based on some simplifications and the neglect of possible hierarchies: we could have distinguished the neurons in the two regions according to morphology or location, and the mice according to age, weight, and so on. But I take it as a good starting compromise between simplification and realism. We'll consider further simplifications in the next step.

3.2 Step 2: joint probability distribution for the connections

In order to assign the initial joint probability distribution $\mathrm{p}[(c_{ijm}) \mid I]$, for all values of the connections (c_{mij}), we explore the symmetries of our problem. Symmetries translate into equalities of some values of the joint distribution, and therefore restrict its possible values and mathematical form.

Let's give an example corresponding to the three-level hierarchy (i) of § 3.0. If we consider all input neurons (i) to a specific target neuron j in a specific mouse m as similar, then the joint probability should stay the same under exchanges of the indices i for which the connections have the same values. For example:

$$\begin{aligned} \mathrm{p}(c_{2\,jm} = 0, c_{5\,jm} = 1, c_{6\,jm} = 1, c_{9\,jm} = 0 \mid I) = \\ \mathrm{p}(\underbrace{c_{9\,jm} = 0, c_{5\,jm} = 1, c_{6\,jm} = 1, c_{2\,jm} = 0}_{2\leftrightarrow 9} \mid I) = \\ \mathrm{p}(\underbrace{c_{2\,jm} = 0, c_{6\,jm} = 1, c_{5\,jm} = 1, c_{9\,jm} = 0}_{5\leftrightarrow 6} \mid I) = \\ \mathrm{p}(\underbrace{c_{9\,jm} = 0, c_{6\,jm} = 1, c_{5\,jm} = 1, c_{2\,jm} = 0}_{2\leftrightarrow 9,\quad 5\leftrightarrow 6} \mid I). \end{aligned} \tag{3}$$

We can't exchange $i = 2$ and $i = 5$ because their connections have different values. Because of this symmetry in i the joint probability must have a specific form, by de Finetti's representation theorem:[4]

$$\mathrm{p}[(c_{ijm}) \mid I] = \int \prod_{jm} \Big[\mathrm{d}f_{jm}\; f_{jm}^{\sum_i c_{ijm}}\; (1 - f_{jm})^{\sum_i (1 - c_{ijm})} \Big]\; \mathrm{p}[(f_{jm}) \mid I]. \quad (4)$$

In this expression, f_{jm} is the total number of neurons projecting to target neuron j in mouse m, and what remains to be assigned is the probability distribution $\mathrm{p}[(f_{jm}) \mid I]$ over such numbers for all j and m. It may look like a complicated expression, but it has considerably reduced the possible assignments of joint probabilities.

The second level of the hierarchy of this example was to consider a similarity of connectivity statistics – which are the quantities f_{jo} – among all target neurons in each specific mouse. This means that the joint probability $\mathrm{p}[(f_{jm}) \mid I]$ in the formula above is symmetric in the index j for each fixed m. This leads, again by de Finetti's theorem, to a new expression similar to the one above, with new m-dependent statistics as parameters and a joint distribution for htem. The final level in the hierarchy says that this distribution is itself symmetric in the index m, thus leading to a final integral expression in terms of a single parameter only with an associated probability distribution.

The difficulty of the three-level formulation just sketched is mathematical: the integrals involved are over more complex spaces which are numerically more expensive to handle.

As a first, simpler formulation we consider instead the more unrealistic but powerful symmetry of example (ii): the joint distribution $\mathrm{p}[(c_{ijm}) \mid I]$ stays the same under exchanges of *any* indices (i, j, m) for which the connections have the same values. For example,

$$\begin{aligned} &\mathrm{p}(c_{237} = 0, c_{552} = 1, c_{914} = 0 \mid I) = \mathrm{p}(c_{937} = 0, c_{552} = 1, c_{214} = 0 \mid I) = \\ &\mathrm{p}(c_{217} = 0, c_{552} = 1, c_{934} = 0 \mid I) = \mathrm{p}(c_{234} = 0, c_{552} = 1, c_{917} = 0 \mid I) = \\ &\mathrm{p}(c_{917} = 0, c_{552} = 1, c_{234} = 0 \mid I) = \cdots \quad \text{and so on.} \end{aligned} \quad (5)$$

We can only exchange indices with the same connection values, but we have full freedom otherwise.

[4] De Finetti 1930; Hewitt et al. 1955; Heath et al. 1976; Diaconis 1977; Diaconis et al. 1980; Dawid 2013.

Again by de Finetti's theorem, this symmetry forces our joint probability distribution to have this form:

$$\begin{aligned} \mathrm{p}[(c_{ijm}) \mid I] &= \int \mathrm{d}q \; q^{\sum_i c_{ijm}} \, (1-q)^{\sum_i (1-c_{ijm})} \, \mathrm{p}(q \mid I) \\ &\equiv \int \mathrm{d}q \; q^{Rf} \, (1-q)^{R\,(1-f)} \, \mathrm{p}(q \mid I) \end{aligned} \tag{6}$$

The symmetries briefly discussed previously

Let's analyse two approximate symmetries first. Consider a specific mouse m.

For each neuron i in the input region, we can approximately judge as equally likely that it will connect to any of the neurons j in the target region. This means that our initial probability distribution should remain the same if we swap some values of the index j, with m and i fixed. For example, the probability that neuron #3 in the input region projects to neuron #2 in the target region and does not project to neuron #5 should be the same if we exchange 2 and 5 in that statement. And similarly for combinations of such statements:

$$\mathrm{p}(c_{m\,32}=1, c_{m\,35}=0 \mid I) = \mathrm{p}(c_{m\,32}=0, c_{m\,35}=1 \mid I). \tag{7}$$

This symmetry is an approximation: if the indices i and j bear some information about the neurons' locations, our probability should reflect the possibility that nearby neurons may connect more easily, for example.

We have an analogous approximate symmetry under exchanges of the input-neuron index i for fixed target neuron j.

These two approximate symmetries together are important because they greatly restrict the possible values of our initial probability distribution $\mathrm{p}[(c_{mij}) \mid I]$. The result they lead to[5] is mathematically difficult for me, however. To further simplify the problem I make an even rougher approximation: that the symmetries above hold not just for i and j separately, but for any pairs of them. This assumption is even more unrealistic than the previous two, because it gives the same probability to the situation where each input neuron is connected with only one, distinct, target neuron, as to the situation where only one input neuron is connected to all target neurons, and all other input neurons are unconnected.

[5] Hoover 1979; Aldous 1981; Diaconis et al. 1981.

With this approximation the joint probability distribution for a given mouse m assumes the following form, by de Finetti's representation theorem:[6]. Take a set of k_m input neurons and l_m target neurons, and consider all possible connections c_{mij} between them, with $i = 1, \dots, k_m$ and $j = 1, \dots, l_m$. Then

$$\begin{aligned} \mathrm{p}[(c_{mij}) \mid m, I] &= \int_0^1 \mathrm{d}\nu_m \; \mathrm{p}(\nu_m \mid I) \; \prod_{ij} \left[{\nu_m}^{c_{mij}} \, (1-\nu_m)^{1-c_{mij}} \right] \\ &\equiv \int_0^1 \mathrm{d}\nu_m \; \mathrm{p}(\nu_m \mid I) \; {\nu_m}^{k_m l_m f_m} \, (1-\nu_m)^{k_m l_m (1-f_m)} . \end{aligned} \tag{8}$$

Here $k_m l_m f_m = \sum_{ij} c_{mij}$ is the total number of connections between the k_m input neurons and the l_m target neurons, and $0 \leqslant \nu_m \leqslant 1$ is the population-averaged number of connections from the *whole* input region to the *whole* target region. The probability distribution $\mathrm{p}(\nu_m \mid I)$ expresses our beliefs about the total number of connections from input to target region in mouse m. The formula above is valid only if k_m and l_m are small (say, one tenth at most) compared to their region sizes.

Let's see how our question and data can be expressed in term of these statements and the quantities (c_{mij}).

We imagine to check every neuron i in the target region of every mouse m. For each such neuron we count how many neurons from the input region project to it. Let's call this number N_{mi}. From the point of view of fig. 1 we're considering every column from every grid, in turn, and counting how many black entries it has. If we were able to perform these observations, at the end we could build a histogram of the numbers (N_{mi}), showing the proportion of neurons in the target region that have 0 inputs from the input region, or 1 input, 2 inputs, n inputs, and so on. These values would be the bins of the histogram. The number of bins is large – as much as the largest number of neurons that could make the input region of a mouse. Denote by F_n the proportion (between 0 and 1) of neurons that receive n inputs. This proportion is calculated from all numbers (N_{mi}) by

$$F_n = \frac{\sum_{mi} \delta(N_{mi} = n)}{\sum_m L_m} \equiv \frac{\text{number of } N_{mi}\text{s equal to } n}{\text{total number of neurons in target regions}} . \tag{9}$$

The whole histogram is the set of proportions $\boldsymbol{F} \coloneqq (F_0, F_1, \dots)$.

[6] De Finetti 1930; Hewitt et al. 1955; Heath et al. 1976; Diaconis 1977; Diaconis et al. 1980; Dawid 2013.

This histogram can give one definite meaning to the question 'how many inputs from the input region does a neuron in the target region have on average?': we are asking 'what the mean of the histogram?', that is,

$$\text{average number of inputs} = \sum_n n\, F_n \equiv \frac{\sum_{mi} N_{mi}}{\sum_m L_m}. \tag{10}$$

But the histogram would tell us even more, for example the percentage of target neurons that receive between 100 and 200 inputs from the input region, obtained by summing $F_{100} + \cdots + F_{200}$.

It's good to keep in mind that our question could also be approached in slightly different ways, leading to different secondary questions. For example, we could construct a histogram for the number of inputs for each mouse individually, thus obtaining several $\boldsymbol{F}^{(m)} = (F^m_0, F^m_1, \dots)$, one for each m. Then we could ask: 'what is the average connectivity histogram for a mouse?', that is,

$$\frac{1}{\text{number of mice}} \sum_m \boldsymbol{F}^{(m)} \tag{11}$$

and also ask how much such histogram can vary from mouse to mouse.

These alternative questions are important in view of possible approximations in our assumptions and maths.

For example, if we assume that all target neurons from all mice can be somehow pooled together – which makes the maths easier – then we are forsaking the possibility of analysing the individual connectivity histograms.

Although we'll never be able to construct such a histogram, we can nevertheless guess, from our observations in a small set of mice, what kind of shape and numerical values it could have. So our goal is to determine the probability

$$\mathrm{p}(\text{histogram} \mid \text{data}, I) \tag{12}$$

for each possible histogram, given the data collected in our observations and any other information I we have, including our present knowledge of brain biology.

These probability above can be found using the rule for conditional probabilities:

$$\mathrm{p}(\text{histogram} \mid \text{data}, I) \propto \mathrm{p}(\text{data} \mid \text{histogram}, I) \times \mathrm{p}(\text{histogram} \mid I). \qquad (13)$$

To calculate these we need to determine our initial probability distribution $\mathrm{p}(\text{histogram} \mid I)$ for all possible histograms. In fact we need to determine the initial probability of a couple more things, for example the total number of neurons in the input region K_m, for each mouse m. All other probabilities are calculated from this initial probability distribution.

In the next section we discuss and analyse several alternative assumptions that we can use to build the distribution $\mathrm{p}(\text{histogram} \mid I) \equiv \mathrm{p}(\boldsymbol{F} \mid I)$. Some of these assumptions are approximate or unrealistic, yet unfortunately necessary for making the problem mathematically tractable. After discussing the alternatives we'll choose specific ones to arrive at formulae that we can use with the data. The discussion of the alternatives is useful because we can step back and try to change the less realistic assumptions if we see that our formulae don't work well.

4 Initial probability distribution

The assumptions we use to build our initial probability distribution $\mathrm{p}(\boldsymbol{F} \mid I)$ are about the symmetries of our problem.

The first symmetry is very intuitive and compelling. Take the distribution of numbers of inputs for a given mouse m: (N_{m1}, N_{m2}) ***

In order to formulate the initial probability distribution $\mathrm{p}[(c_{mij}) \mid I]$ we explore the symmetries of our problem.

Let's analyse two approximate symmetries first. Consider a specific mouse m.

the answers to our question Q2 and the statement of our experimental observations are built up from the specific statements (17). For example, the statement *In mouse # 3, neuron # 1 in the target region receives input from 2 neurons in the input region* is equivalent to the composite statement *In mouse # 3, neuron # 1 in the target region receives input from neurons # 1 and # 2, or # 1 and # 3, or # 1 and # 4, or … in the input region*, which we can express as

$$(C_{311} \text{ and } C_{312}) \text{ or } (C_{311} \text{ and } C_{313}) \text{ or } \cdots \\ \text{or } (C_{312} \text{ and } C_{313}) \text{ or } (C_{312} \text{ and } C_{314}) \text{ or } \cdots \tag{14}$$

Let's denote such kind of statement with

$$C_{mij} \coloneqq \textit{In mouse } m\textit{, neuron } i \textit{ in input region projects to neuron } j \textit{ in target region} \tag{15}$$

and let's introduce the quantity

$$c_{mij} = \begin{cases} 1 & \text{if } C_{mij} \text{ is true,} \\ 0 & \text{if } C_{mij} \text{ is false.} \end{cases} \tag{16}$$

We can rephrase it in two different ways; it's the second that probably interests us:

Q1. If we examine a new neuron in region A in a new mouse, how many neurons in region B will project to it? The possible answers to this question are: 0, 1, 2, and so on.

Q2. If we examine all neurons in region A in a very large sample of mice, and for each neuron we count how many neurons from region B project to it, then what are the frequencies with which we'll observe 0 connections, 1 connection, 2 connections, and so on? One possible answer could be this distribution:

$$0\text{: } 0.1\%, \quad 1\text{: } 0.3\%, \quad \dots, \quad 1\,000\text{: } 24\%, \quad \dots$$

another answer could be this distribution:

$$0\text{: } 0.02\%, \quad 1\text{: } 0.07\%, \quad \dots, \quad 1\,000\text{: } 31\%, \quad \dots$$

and so on, for all possible distributions of frequencies.

For the first question we must give a distribution of probability over all possible answers 0, 1, and so on. For the second the probability is distributed over all possible frequency distributions.

The answers to these two questions and their probabilities are connected; in particular, from the second we can obtain the first. The second is more informative, because it gives us also a range of variability.

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

We must consider the set of mice that have been observed and those that will be observed in the future. Let's label each mouse with an index m. For our inferences to be valid, these mice must be judged 'similar'.

Now that we've labelled the mice and their neurons in the two regions, we can make several concrete statements, such as

> *In mouse # 64, the neuron # 931 in the input region projects to the neuron # 1 488 in the target region.*

which can be either true or false. Let's denote such kind of statement with

$$C_{mij} \coloneqq \textit{In mouse } m\textit{, neuron } i \textit{ in input region projects to neuron } j \textit{ in target region} \tag{17}$$

and let's introduce the quantity

$$c_{mij} = \begin{cases} 1 & \text{if } C_{mij} \text{ is true,} \\ 0 & \text{if } C_{mij} \text{ is false.} \end{cases} \tag{18}$$

Obviously for most of these statements we'll never be able to ascertain whether they're true or false. But the answers to our question Q2 and the statement of our experimental observations are built up from the specific statements (17). For example, the statement *In mouse # 3, neuron # 1 in the target region receives input from 2 neurons in the input region* is equivalent to the composite statement *In mouse # 3, neuron # 1 in the target region receives input from neurons # 1 and # 2, or # 1 and # 3, or # 1 and # 4, or … in the input region*, which we can express as

$$(C_{311} \text{ and } C_{312}) \text{ or } (C_{311} \text{ and } C_{313}) \text{ or } \cdots \\ \text{or } (C_{312} \text{ and } C_{313}) \text{ or } (C_{312} \text{ and } C_{314}) \text{ or } \cdots \tag{19}$$

The probabilities of such combinations are completely determined once we give an initial joint probability distribution for all statements (17) and their negations, which is equivalent to the joint distribution of

$$\mathrm{p}[(c_{mij}) \mid I] \equiv \mathrm{p}(c_{111}, c_{112}, c_{113}, \dots \mid I) \tag{20}$$

conditional on the information I that we have before doing the experiments and the observations.

Now, we'll want the probability of a specific answer A to our question, conditional on the data D we observed: $\mathrm{p}(A \mid D, I)$. By the rule of conditional probability we obtain it as

$$\mathrm{p}(A \mid D \text{ and } I) = \frac{\mathrm{p}(A \text{ and } D \mid I)}{\mathrm{p}(D \mid I)}. \tag{21}$$

The probabilities in the numerator and denominator can be obtained from our main joint probability (20) using the basic rules of the probability-calculus.

So our strategy is as follows:

1. Quantitatively determine the initial joint probability $\mathrm{p}[(c_{mij}) \mid I]$.
2. Resolve the answers to question Q2 in terms of the basic statements C_{mij}.
3. Resolve the experimental observations in terms of the basic statements C_{mij}.
4. Calculate the probabilities in (21).

We now face these steps in turn. We shall discuss several possible alternative assumptions and conceptual or mathematical approximations that we can make in order to proceed.

5 Initial probability distribution

In order to formulate the initial probability distribution $\mathrm{p}[(c_{mij}) \mid I]$ we explore the symmetries of our problem.

Let's analyse two approximate symmetries first. Consider a specific mouse m.

For each neuron i in the input region, we can approximately judge as equally likely that it will connect to any of the neurons j in the target region. This means that our initial probability distribution should remain

the same if we swap some values of the index j, with m and i fixed. For example, the probability that neuron #3 in the input region projects to neuron #2 in the target region and does not project to neuron #5 should be the same if we exchange 2 and 5 in that statement. And similarly for combinations of such statements:

$$\mathrm{p}(c_{m\,3\,2}=1, c_{m\,3\,5}=0 \mid I) = \mathrm{p}(c_{m\,3\,2}=0, c_{m\,3\,5}=1 \mid I). \tag{22}$$

This symmetry is an approximation: if the indices i and j bear some information about the neurons' locations, our probability should reflect the possibility that nearby neurons may connect more easily, for example.

We have an analogous approximate symmetry under exchanges of the input-neuron index i for fixed target neuron j.

These two approximate symmetries together are important because they greatly restrict the possible values of our initial probability distribution $\mathrm{p}[(c_{mij}) \mid I]$. The result they lead to[7] is mathematically difficult for me, however. To further simplify the problem I make an even rougher approximation: that the symmetries above hold not just for i and j separately, but for any pairs of them. This assumption is even more unrealistic than the previous two, because it gives the same probability to the situation where each input neuron is connected with only one, distinct, target neuron, as to the situation where only one input neuron is connected to all target neurons, and all other input neurons are unconnected.

With this approximation the joint probability distribution for a given mouse m assumes the following form, by de Finetti's representation theorem:[8]. Take a set of k_m input neurons and l_m target neurons, and consider all possible connections c_{mij} between them, with $i = 1, \dots, k_m$ and $j = 1, \dots, l_m$. Then

$$\begin{aligned}\mathrm{p}[(c_{mij}) \mid m, I] &= \int_0^1 \mathrm{d}\nu_m \; \mathrm{p}(\nu_m \mid I) \; \prod_{ij} \left[{\nu_m}^{c_{mij}} \; (1-\nu_m)^{1-c_{mij}}\right] \\ &\equiv \int_0^1 \mathrm{d}\nu_m \; \mathrm{p}(\nu_m \mid I) \; {\nu_m}^{k_m l_m f_m} \; (1-\nu_m)^{k_m l_m (1-f_m)}.\end{aligned} \tag{23}$$

Here $k_m l_m f_m = \sum_{ij} c_{mij}$ is the total number of connections between the k_m input neurons and the l_m target neurons, and $0 \leqslant \nu_m \leqslant 1$ is the

[7] Hoover 1979; Aldous 1981; Diaconis et al. 1981.

[8] De Finetti 1930; Hewitt et al. 1955; Heath et al. 1976; Diaconis 1977; Diaconis et al. 1980; Dawid 2013.

population-averaged number of connections from the *whole* input region to the *whole* target region. The probability distribution $\mathrm{p}(\nu_m \mid I)$ expresses our beliefs about the total number of connections from input to target region in mouse m. The formula above is valid only if k_m and l_m are small (say, one tenth at most) compared to their region sizes.

The probability distribution for the connections present in different mice can be obtained by combining the expressions above for each m. To do this we need the initial joint probability distribution

$$\mathrm{p}[(\nu_m) \mid I] \equiv \mathrm{p}(\nu_1, \nu_2, \dots \mid I) \tag{24}$$

for the total number of connections in several mice.

Our question has another symmetry, more realistic than the previous two. We don't believe any of the mice to be examined to be special with respect to the others, nor we do believe that the particular order in which they are observed has any relevance. This means that our initial probability distribution above should be invariant with respect to permutations of the index m. Then we obtain the following expression, again by de Finetti's theorem:

$$\mathrm{p}[(\nu_m) \mid I] = \int \mathrm{d}\theta \; \mathrm{p}(\theta \mid I) \prod_m q(\nu_m \mid \theta), \tag{25}$$

where q is a distribution that depends on the parameter θ, and $\mathrm{p}(\theta \mid I)$ is our initial probability distribution for this parameter.

The parameter θ in principle represents every possible distribution – or histogram – of the total number of connections across the mice. For example, a particular θ can say that 7% of mice had 10^6 connections, 11% had 5×10^7 connections, and so on. The above formula sums over all such possible distributions. There is an infinite of such distributions, so the integral is over an infinite-dimensional space. When we calculate this integral with this full generality, we're using a so-called *non-parametric* method. But for mathematical simplicity we can also decide to restrict our attention to histograms of specific shape, such as a normal for example. In this case the integral is restricted to a finite-dimensional space and we're using a *parametric* method.

Combining the initial probability distributions (23) and (25) we finally obtain our initial probability distribution for a joint set of statements

(C_{mij}):

$$\mathrm{p}[(c_{mij}) \mid I] = \iint \mathrm{d}\theta \; \mathrm{d}\boldsymbol{\nu} \; \mathrm{p}(\theta \mid I) \; \prod_{mij} \Big[q(\nu_m \mid \theta) \; {\nu_m}^{c_{mij}} \; (1-\nu_m)^{1-c_{mij}} \Big], \tag{26}$$

where $\boldsymbol{\nu} \coloneqq (\nu_m)$.

6 Statements of the questions and of the data

Bibliography

('de *X*' is listed under D, 'van *X*' under V, and so on, regardless of national conventions.)

Aldous, D. J. (1981): *Representations for partially exchangeable arrays of random variables*. J. Multivariate Anal. **11**[4], 581–598.

Damien, P., Dellaportas, P., Polson, N. G., Stephens, D. A., eds. (2013): *Bayesian Theory and Applications*. (Oxford University Press, Oxford).

Dawid, A. P. (2013): *Exchangeability and its ramifications*. In: Damien, Dellaportas, Polson, Stephens (2013), ch. 2, 19–29.

de Finetti, B. (1930): *Funzione caratteristica di un fenomeno aleatorio*. Atti Accad. Lincei: Sc. Fis. Mat. Nat. **IV**[5], 86–133. http://www.brunodefinetti.it/Opere.htm.

Diaconis, P. (1977): *Finite forms of de Finetti's theorem on exchangeability*. Synthese **36**[2], 271–281. http://statweb.stanford.edu/~cgates/PERSI/year.html.

Diaconis, P., Freedman, D. (1980): *Finite exchangeable sequences*. Ann. Prob. **8**[4], 745–764.

— (1981): *On the statistics of vision: the Julesz conjecture*. J. Math. Psychol. **24**[2], 112–138. http://statweb.stanford.edu/~cgates/PERSI/year.html.

Good, I. J. (1969): *A subjective evaluation of Bode's law and an 'objective' test for approximate numerical rationality*. J. Am. Stat. Assoc. **64**[325], 23–49. Partly repr. in Good (1983) ch. 13.

— (1983): *Good Thinking: The Foundations of Probability and Its Applications*. (University of Minnesota Press, Minneapolis, USA).

Hailperin, T. (1996): *Sentential Probability Logic: Origins, Development, Current Status, and Technical Applications*. (Associated University Presses, London).

— (2011): *Logic with a Probability Semantics: Including Solutions to Some Philosophical Problems*. (Lehigh University Press, Plymouth, UK).

Heath, D., Sudderth, W. (1976): *De Finetti's theorem on exchangeable variables*. American Statistician **30**[4], 188–189.

Hewitt, E., Savage, L. J. (1955): *Symmetric measures on Cartesian products*. Trans. Am. Math. Soc. **80**[2], 470–501.

Hoover, D. N. (1979): *Relations on probability spaces and arrays of random variables*. Tech. rep. (Institute for Advanced Study, Princeton). http://www.stat.berkeley.edu/~aldous/Research/hoover.pdf.

Jaynes, E. T. (2003): *Probability Theory: The Logic of Science*. (Cambridge University Press, Cambridge). Ed. by G. Larry Bretthorst. First publ. 1994. https://archive.org/details/XQUHIUXHIQUHIQXUIHX2, http://www-biba.inrialpes.fr/Jaynes/prob.html.

Johnson, W. E. (1924): *Logic. Part III: The Logical Foundations of Science*. (Cambridge University Press, Cambridge). https://archive.org/details/logic03john.

Sox, H. C., Higgins, M. C., Owens, D. K. (2013): *Medical Decision Making*, 2nd ed. (Wiley, New York). First publ. 1988.