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

Guessing long-run mutual information [draft]

P.G.L. Porta Mana
<pgl@portamana.org>

C. Battistin
<>

S. Gonzalo Cogno
<@>

***; updated 11 July 2020

***

1 Introduction

We have two kinds of stimuli and, say, ten kinds of neural responses to these stimuli, under specific environmental and behavioural conditions, and for a specific kind of subjects. The responses could be, for example, ten different firing rates of a specific neuron, or ten different total-population activities of a brain region.

Imagine that a researcher told us the results of 10^{60} stimulus-response measurements, all performed for the specified conditions and subjects, and in which the two kinds of stimulus occurred equally often. From the long-run[1] joint frequencies of stimulus-response pairs we could calculate the long-run mutual information between stimulus and response.

But we don't have access to such a wealth of observations, and never will. Suppose we only have actual knowledge of a sample of few, say 20, measurements for each kind of stimulus, from such long-run sequence. We can calculate the joint frequencies of stimulus-response pairs in this sample and the resulting sample mutual information.

Our question is: how do the mutual information of the sample and of the long-run sequence differ?

Let's explore some possibilities.

[1] "But this *long run* is a misleading guide to current affairs. *In the long run* we are all dead." (**keynes1923_r2013**)

Case A. Suppose the long-run frequencies of responses conditional on each kind of stimulus are as given in the top panel of fig. 1[2]. In the long run all responses occur equally often for either stimulus. The long-run mutual information is 0 bit.

Now we consider a sample of 20 response measurements for each stimulus from such long-run distribution. From this sample we calculate the joint frequencies and their mutual information. The exact long-run sequence is unknown, so such a sample could yield different results. Considering all long-run sequences (having the same long-run distribution) as equally possible, we plot some possible results for the sample mutual information in the bottom panel. The long-run mutual information is indicated by the red line. The sampled mutual information is most surely always higher than the long-run one, with an average of 0.21 bit.

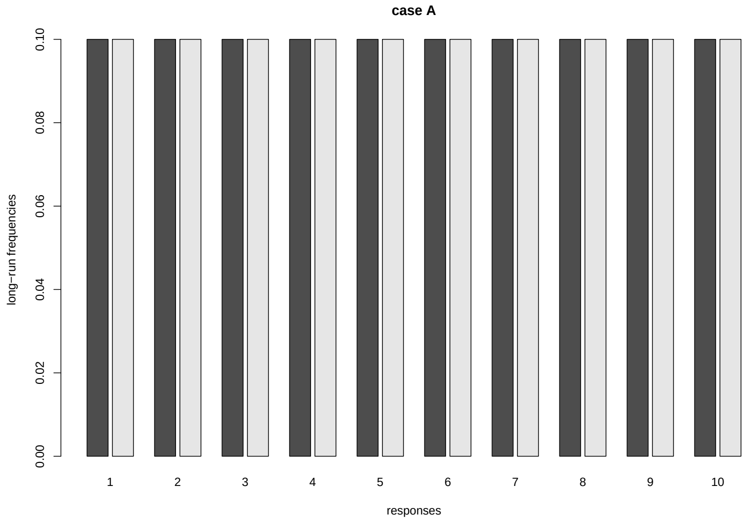
Case B. Suppose the long-run frequencies of responses conditional on each kind of stimulus are as given in the top panel of fig. 2. The responses do not occur equally often in the long run, and the frequencies of some of them are different for the two kinds of stimulus. The long-run mutual information is 0.43 bit.

We again consider a sample of 20 response measurements for each stimulus. The corresponding plot for the sampled mutual information is shown in the bottom panel. The sampled mutual information is most surely always lower that the long-run one, with an average of 0.22 bit.

Case C. Suppose the long-run frequencies of responses conditional on each kind of stimulus are as given in the top panel of fig. 3. Also in this case the responses do not occur equally often in the long run, and their frequencies are different for the two kinds of stimulus. The long-run mutual information is 0.22 bit.

We again consider a sample of 20 response measurements for each stimulus. The corresponding plot for the sampled mutual information is shown in the bottom panel. The sampled mutual information is on average 0.22 bit, equal to the long-run one.

[2] **panzerietal2007**.



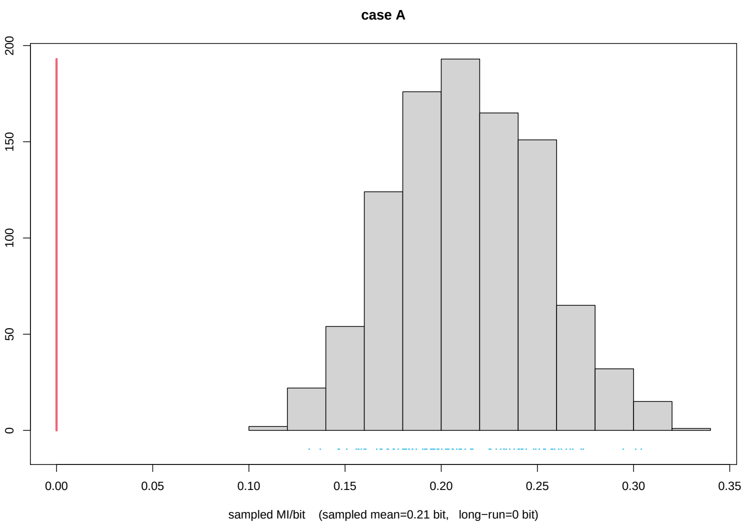


Figure 1 Case A

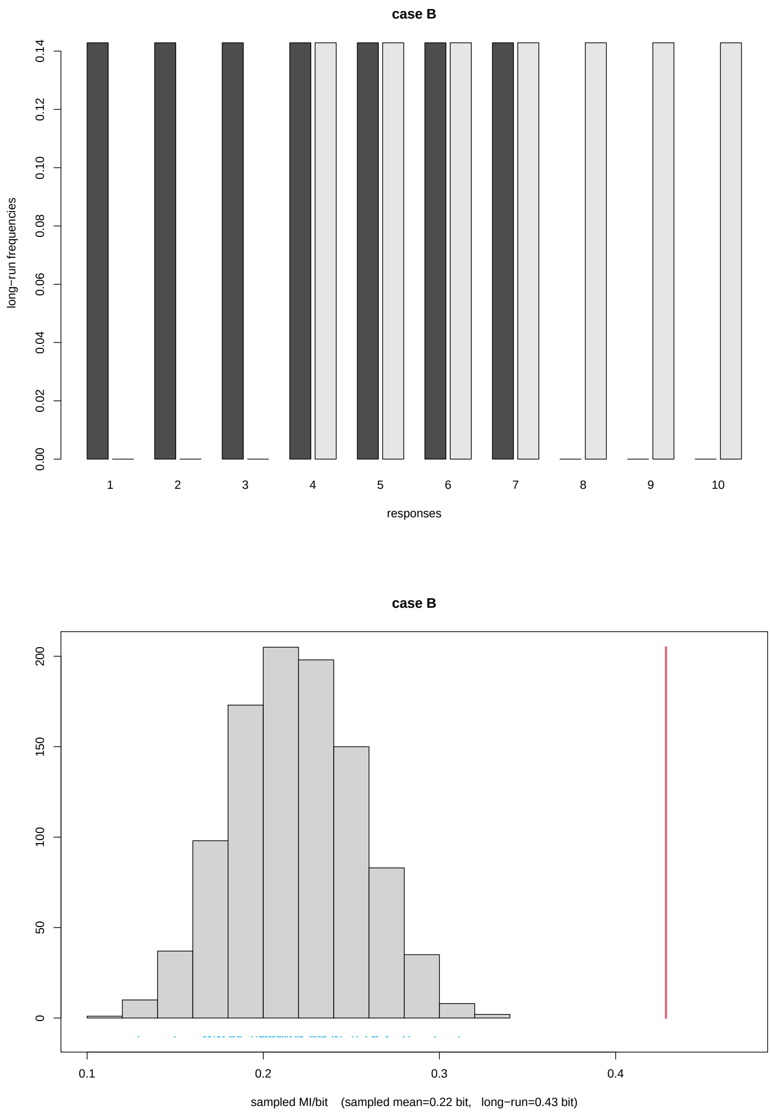


Figure 2   Case B

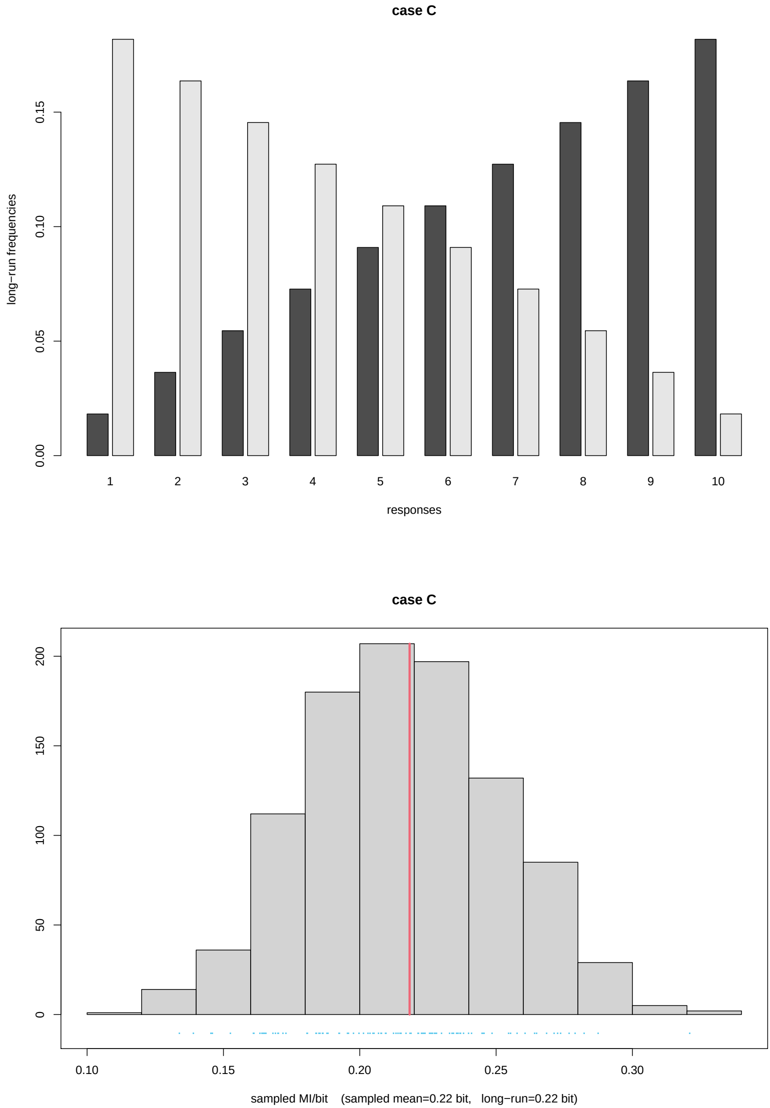


Figure 3 Case C