

Assignment 3

Theoretical Neuroscience

TAs:

Roman Pogodin (roman.pogodin.17@ucl.ac.uk)
Franziska Broeker (franziska.broeker.15@ucl.ac.uk)
Ilyes Khemakhem (ilyes.khemakhem.17@ucl.ac.uk)

due: 8 March

1 Point Processes

Question 1. Doubly stochastic Poisson processes and spike patterns.

In the 1980s Abeles suggested that the integrative properties of neurons, coupled with the density of connections between them, would lead to self-supporting synchronous volleys of firing that could propagate between different constellations of neurons with extremely high temporal precision (a phenomenon called a “synfire chain”). This prompted an experimental search for the precisely timed spike patterns that might be a signature of such a phenomenon. A single neuron might participate in more than one synchronous volley of a synfire chain. Thus, in part because of technological limitations, many experiments looked for patterns in the spike train of a single cell. Here, we will look at one such hypothetical experiment.

Suppose the mean response rate of a neuron to a stimulus flashed shortly before time 0, is given by the function

$$\bar{\lambda}(t) = \Theta(t)\bar{\rho}e^{-t/T}$$

where $\Theta(t)$ is the Heaviside function (0 if $t < 0$ and 1 if $t \geq 0$) and $\bar{\rho}$ and T are constants. We begin by making the common assumption that the firing of the neuron is described by an inhomogeneous Poisson process with intensity $\bar{\lambda}(t)$.

1. On average, how many spikes will the cell emit in response to the stimulus (assume the experimental counting interval is $\gg T$).
2. Under the inhomogeneous Poisson model, what is the intensity with which we would observe spikes within small intervals around three specific times $t, t + \tau_1$ and $t + \tau_2$ all greater than 0.

Hint. We want the marginal probability of those 3 times – don’t assume anything about what the cell is doing at any other time.

3. Integrate your expression with respect to t to find $\sigma(\tau_1, \tau_2)$, the intensity of observing a pattern with intervals τ_1 and τ_2 at any point. Assume τ_1 and τ_2 are positive.
4. An experimenter reports that, looking at a neuron with $\bar{\rho} = 80\text{s}^{-1}$ and $T = 0.05\text{s}$ and binning spikes in 1 ms intervals, he observed the pattern (5, 50) (i.e., $\tau_1 = 5$ ms and $\tau_2 = 50$) 8 times in 1000 trials. Given your result above, is this surprising? Assume that he looked only for the (5, 50) ms pattern. [**Bonus.** Why should that matter to your answer?]

Looking more closely at his data, you note that the Fano Factor of the spike count is about 2. This leads you to consider a doubly stochastic Poisson process model instead, with an intensity

$$\lambda(t) = \Theta(t)\rho e^{-t/T}$$

which depends on a random variable $\rho \sim \text{Gamma}(\alpha, \beta)$.

5. **Bonus.** Use moment matching to estimate values of the parameters α and β . [That is, find an expression for the variance of a Poisson *counting* distribution with random mean parameter drawn from $\text{Gamma}(\alpha, \beta)$. Find values of α and β for which this expression matches the observed Fano factor.]
6. **Bonus.** Repeat the calculation for the expected number of (5,50) ms patterns. [**Hint.** You'll need the third moment of the Gamma distribution]. Is the experimental result surprising now?

Question 2. The expected autocorrelation function of a renewal process.

In class, we analysed the autocorrelation function of a point process in terms of its intensity function $\lambda(t, \dots)$. For a self-exciting point process, λ depends on the past history of spiking, and so computing the expected value of the correlation in this way can be quite difficult. Fortunately, for the special case of a renewal process (i.e. a point process with iid inter-event intervals), there is an alternative way to compute the autocorrelation function.

Consider a neuron whose firing can be described by a renewal process with inter-spike interval probability density function $p(\tau)$.

1. Given an event at time t , the probability that the next spike arrives in the interval $I_\tau = [t + \tau, t + \tau + d\tau)$ is $p(\tau)d\tau$. What is the probability that the *second* spike after the one at t arrives in I_τ instead? The third spike?
2. What is the probability that, given a spike at t , there is a spike in I_τ , regardless of the number of intervening spikes? **Hint.** You can solve the resulting equation using the Laplace transform by setting $p(\tau) = 0$ for $\tau < 0$.
3. Your answer to the previous question has given you the positive half of the autocorrelation function. What does the negative half look like? What happens at $\tau = 0$?

Question 3. Fisher information and refractory firing

As a reminder, Fisher information of a distribution $p(t|s)$ is defined as

$$J(s) = \left\langle -\frac{\partial^2 \log p(t|s)}{\partial s^2} \right\rangle_{p(t|s)} = \left\langle \left(\frac{\partial \log p(t|s)}{\partial s} \right)^2 \right\rangle_{p(t|s)}.$$

It measures how much information a response t carries about a stimulus s .

Consider a hypothetical cell, which responds to the presentation of a stimulus with a continuous feature s by firing at a homogeneous rate $f(s)$ in a (fixed) interval $[0, T]$. Assume that the firing rate is 0 outside this interval and that there was a spike at $t = 0$. We will be interested in the contributions made to the Fisher information by spike-timing, with and without a refractory period.

First, assume that the firing is Poisson.

1. What is the probability of observing spikes at times $\{t_1 \dots t_n\} \subset [0, T]$?
2. What is the Fisher information $J_{t, \text{Pois}}(s)$ associated with this probability density function, assuming that the relevant interval $[0, T]$ is known? How does it compare to the Fisher information $J_{n, \text{Pois}}$ associated with the distribution of spike counts $P(n|s)$?

Now consider refractory firing. Recall that one way to model a refractory period is to use a gamma-interval renewal process in place of a Poisson process. Thus, now assume that the cell's firing follows a gamma-interval process with the same mean rate $f(s)$ and with gamma order $\gamma \in \mathbb{N}$.

3. What is the probability of observing spikes at times $\{t_1 \dots t_n\} \subset [0, T]$ from this process?
4. What is the Fisher information $J_{t, \text{Gamma}}(s)$ associated with the new probability density function? You may assume that T is long enough to neglect contributions due to the first spike, and due to the silence after the last spike.

Finally, we wish to see how much of this information gain is available in the spike count.

5. Which signal (count or spike-timing) do you expect to carry more information for this process? Why?
6. Find an expression for the the distribution of spike counts $P(n|s)$ under the gamma-interval model.
7. Write down the expression for the corresponding Fisher information $J_{n,\text{Gamma}}$, and thus for $J_{n,\text{Gamma}} - J_{t,\text{Gamma}}$. You need not necessarily evaluate the expectation. Identify the term(s) responsible for the difference between $J_{n,\text{Gamma}}$ and $J_{t,\text{Gamma}}$.
8. **Bonus.** It is possible to repeat the derivations for $\gamma \in \mathbb{R}_+$ if you assume a spike at $t = T$. What can you say about the difference between $J_{n,\text{Gamma}}$ and $J_{t,\text{Gamma}}$ in this case?

2 Information Theory

Question 4. Warmup KL.

1. Differential entropy is an extension of entropy to continuous distributions. What is the differential entropy of the continuous distribution with density $u_a(x)$ that is uniform on the interval $[-a/2, a/2]$ and 0 elsewhere? When is this entropy equal to 0? Compare to the case of the entropy of a discrete distribution being 0. What happens when $a = 0.5$? What happens as $a \rightarrow \infty$? Do these results correspond to your intuitive notion of uncertainty in a distribution?
2. Is the KL divergence defined on continuous distributions “well-behaved”? Calculate $KL[u_1||u_2]$ as well as $KL[u_2||u_1]$ (u_a as defined above). Interpret these results in terms of the “coding penalty” discussed in class.
3. **Bonus.** Use second-order conditions (i.e. positive-semidefiniteness of the Hessian) to prove that the KL divergence $KL[p||q]$ is convex in the pair (p, q) ; i.e., if (p_1, q_1) and (p_2, q_2) are two pairs of distributions and $a \in [0, 1]$, then

$$KL[ap_1 + (1-a)p_2||aq_1 + (1-a)q_2] \leq aKL[p_1||q_1] + (1-a)KL[p_2||q_2].$$

Question 5. Communication through a probabilistic synapse

1. The Blahut-Arimoto algorithm.

In this part of the question, we derive an algorithm to find an input distribution that achieves the capacity of an arbitrary discrete channel.

- (a) Given a channel characterised by the conditional distribution $P(R|S)$, we wish to find a source distribution $P(S)$ that maximises the mutual information $I(R; S)$. Show that

$$I(R; S) \geq \sum_{s,r} P(s)P(r|s) \log \frac{Q(s|r)}{P(s)}$$

for any conditional distribution $Q(S|R)$. When is equality achieved?

- (b) Use this result to derive (in closed form) an iterative algorithm to find the optimal $P(S)$. This is called the Blahut-Arimoto algorithm. Prove that the algorithm converges (showing that it converges to a unique maximal value of $I(R; S)$ is more challenging). **Hint.** By analogy to EM, alternate maximisations of the bound on the right hand side with respect to Q and to $P(S)$.

2. Synaptic failure.

Many synapses in the brain appear to be unreliable; that is, they release neurotransmitter stochastically in response to incoming spikes. Here, we will build an extremely crude model of communication under these conditions.

Assume that the input to the synapse is represented by the number of spikes arriving in a 10 ms interval, while the output is the number of times a vesicle is released in the same period. Let the minimum inter-spike interval be 1 ms (taking into account both the length of the spike and the refractory period), and assume that at most 1 vesicle is released per spike. Thus, both input and output symbols on this channel are integers between 0 and 10 inclusive.

Let the probability of vesicle release be independent for each spike in the input symbol, and be given by α^n where α is a measure of synaptic depression and n is the number of spikes in the symbol. (We are neglecting order-dependent effects within each 10ms symbol, and any interactions between successive symbols. This is a terrible model of synaptic behaviour).

- (a) Generate (in MATLAB/Python/etc) the conditional distribution of output given input for this synapse. Take $\alpha = 0.9$. Use Blahut-Arimoto to derive the capacity-achieving input distribution and plot it.
- (b) Try to interpret your result intuitively. Might this have anything to do with the short “bursts” of action potentials found in many spike trains?
- (c) **Bonus.** Improve on the model of synaptic transmission. Consider 5 ms input and output symbols, each being a 5-bit binary number where a 1 indicates a spike or a vesicle release. The probability of transmission for each spike in the symbol is again α^n but n is now the number of vesicles released *so far* for this symbol. Construct a new conditional distribution table and repeat the optimisation. Do you get a qualitatively similar result?