## APPM 4370/5370, Computational Neuroscience: Homework 3

Due Wed. Oct. 6, 2021
at 11am (in class)

THIS HOMEWORK IS ON FOUR PAGES. PAGE 1-2: THE FIRST THREE MATH EXERCISES.
PAGE 3: THE NO COLLABORATION PROBLEM. PAGE 4: PYTHON EXERCISES.

***Math exercises.*** Relevant Gerstner et al sections: 7.1-7.5, 8.1, 8.4, 10.1, 10.2; & probability primer

1. **Poisson process in discrete & continuous time (and renewal).** Here you will work through through computing statistics of a Poisson neuron. Mostly, this requires a bit of calculus and knowledge from the basic probability primer we covered in class.

   (a) In a small time interval $\Delta t$, the probability a neuron fires is $\nu \Delta t$. What is the probability it does not fire in an interval $\Delta t$?

   (b) What is the probability the neuron goes $N$ time steps $\Delta t$ without firing?

   (c) Assume the neuron has just fired at $t = 0$. What is the probability $p(N, \Delta t)$ it fires next exactly at time $N\Delta t$? Hint: You must compute the probability it does not fire within some time $(N-1)\Delta t$ times the probability it fires in an interval $\Delta t$.

   (d) Take the limit of your answer from (b) as $\Delta t \to 0$ while keeping $t = N\Delta t$ fixed (so $N \to \infty$). The result is the survivor function $S_0(t; \nu)$ or the probability the neuron survives to time $t$ without firing.

   (e) The spike time distribution (the probability density function for the time of the next spike) is $P_0(t;\nu) = -\frac{dS_0}{dt}$, but it can also be computed by taking the limit $\lim_{\Delta t \to 0} \frac{p(N,\Delta t)}{\Delta t}$ with $t = N\Delta t$ fixed. Show both approaches yield that $P_0(t;\nu) = \nu \mathrm{e}^{-\nu t}$.

   (f) You are given spike time data from one of two neurons. Neuron 1 has Poisson spike rate $\nu_1 = 4$ Hz. Neuron 2 has $\nu_2 = 2$ Hz. Over one time interval of length $t = 2$s, you notice there are no spikes. What is the probability this would occur in neuron 1? What about neuron 2? Based only on this observation, what neuron do you think this data came from?

2. **Inferring model parameters from data.** This problem builds on the end of the previous one. We will use knowledge of probability distributions associated with Poisson neuron models to develop a method to fit parameters to data.

   (a) The probability of observing $k$ spikes from a Poisson neuron with rate $\nu$ in a time interval $t$ is

   $$P(k;\nu,t) = \frac{(\nu t)^k}{k!}\mathrm{e}^{-\nu t}.$$

   With this in mind, is it more likely a Poisson neuron with rate $\nu_1 = 3$Hz has exactly 10 spikes in 3s or that a Poisson neuron with rate $\nu_2 = 4$Hz has exactly 10 spikes in 3s? If you observe a neuron emit 10 spikes in 3s, do you think it is more likely to be neuron 1 or 2? Note, this second question is different from the first.

   (b) Here, we make clear the connection between the first and second question using Bayes rule. If a Poisson neuron is observed, and we see $k$ spikes in $t$ seconds, the probability this came from a neuron with rate $\nu$ is (by Bayes rule):

   $$P(\nu|k,t) = \frac{(\nu t)^k}{k!}\mathrm{e}^{-\nu t}\frac{p_0(\nu)}{p(k,t)},$$

   where $p_0(\nu)$ is our prior belief about the likelihood of a Poisson neuron having rate $\nu$ (before we make any observations) and $p(k,t)$ is the probability of observing $k$ spikes over time interval $t$ across all neuron types. Assume an 'improper' flat prior ($p_0(\nu) = 1$). In this case, if we observe a neuron and see 4 spikes in 3s is it more likely this came from a $\nu_1 = 1$Hz neuron or a $\nu_2 = 2$Hz?

(c) Now, determine the Poisson rate $\nu$ with maximum likelihood from the data suggested in part (b). That is, compute the $\nu$ that maximizes $P(\nu|k,t)$ from part (b).

3. **Spike time distribution of noisy integrate-and-fire model.** In this problem, you will estimate the distribution of inter spike interval of the noisy, perfect integrate-and-fire (PIF) neuron:

$$\frac{du}{dt} = 1 + \xi(t), \qquad u(0) = v, \quad \text{and if } u(t) = \theta, \text{ end simulation} \tag{1}$$

where unit input is constant and $\xi(t)$ is called a *white noise* process (which we assume has unit amplitude). In this case, the probability $p(u,t)$ of finding the neuron at a given voltage $u$ at time $t$ is given by the *partial differential equation*

$$\frac{\partial p(u,t)}{\partial t} = -\frac{\partial p(u,t)}{\partial u} + \frac{\partial^2 p(u,t)}{\partial u^2},$$

with initial condition $p(u,0) = \delta(u-v)$ and boundary condition $p(\theta,t) = 0$. A sequence of calculations allows us to derive a second order differential equation for the mean time $T_1(v)$ for a spike to occur:

$$T_1''(v) + T_1'(v) = -1, \quad T_1(\theta) = 0, \quad \lim_{v\to-\infty} T_1''(v) = 0.$$

(a) Solve the above boundary value problem and show $T_1(v) = \theta - v$. Hint: Start by integrating the differential equation once. Then, solve using an integrating factor and both boundary conditions to specify the free constants. Why does $T_1(v)$ decrease as $v$ is moved closer to $\theta$?

(b) The 'second moment' for the distribution of spike times is related to the variance, and is determined by the following boundary value problem

$$T_2''(v) + T_2'(v) = -2T_1(v), \quad T_2(\theta) = 0, \quad \lim_{v\to-\infty} T_2'''(v) = 0.$$

Plug in $T_1(v)$ from part (a) and solve the above for $T_2(v)$, using a similar approach to before.

(c) Lastly, in the specific case where $v = 0$, compute $\mu(\theta) = T_1(v;\theta)$ and $\sigma^2(\theta) = T_2(v;\theta) - T_1(v;\theta)^2$ as functions of $\theta$. These are the mean and variance of a normal distribution approximating the distribution of interspike intervals. Plug $\mu(\theta)$ and $\sigma^2(\theta)$ into the normal distribution

$$f(t;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(t-\mu)^2/(2\sigma^2)}$$

and make a sketch of $f(t;\mu(\theta),\sigma^2(\theta))$ as a function of $t$ indicating how $\theta$ shapes the function. Give an explanation as to why.

4. ***NO COLLABORATION PROBLEM!*** **The McKean model.** As shown in class, spike solutions of piecewise linear versions of the Fitzhugh-Nagumo model can be approximately determined by hand using fast-slow methods. We'll go through a simpler version of the example from class here. Consider the McKean model:

$$\frac{du}{dt} = -u + H(u - 1/2) - w,$$
$$\frac{dw}{dt} = \epsilon(2u - w),$$

where

$$H(u - 1/2) = \begin{cases} 0, & u < 1/2, \\ 1, & u > 1/2, \end{cases}$$

is the Heaviside function and we've fixed all the parameters except the rate $\epsilon$ of the gating variable.

(a) There is only one fixed point. Find it and use linear stability analysis to show it is stable. Hint: $\frac{d}{du}H(u - 1/2) = 0$ everywhere except at $u = 1/2$.

(b) Assume $u(0) = 3/4$ and $w(0) = 0$. If $0 < \epsilon \ll 1$ (really small), explain why we expect $u \to 1$ quite quickly. Thereafter, we expect $u \to 1 - w$ quickly as $w$ slowly changes. Write the slow equation for $w$ and solve it to find the time $T$ at which $w(T) = 1/2$ (at which point the spike ends). What happens to $T$ as $\epsilon$ is decreased? Why does this make sense?

***python exercises.*** See relevant python and jupyter notebooks in github repo for reference.

5. **Spike time distribution of noisy integrate-and fire model.** Consider the noisy perfect integrate-and-fire model discussed in Exercise 3. Use the jupyter notebook on noisy integrate-and-fire models to help you work the following exercise which will numerically check your calculations from that previous problem.

   (a) Plot the mean spike time $\mu(\theta)$ as a function of $\theta$ for $\theta \in (0, 2]$ from Exercise 3a. Then, adapt code to numerically estimate the mean time $\mu(\theta)$ between spikes of the noisy perfect integrate and fire model, Eq. (1), given initially $u = 0$, by generating $N = 1000$ spikes. You can use 'np.mean().' Do this specifically for $\theta = 0.1, 0.5, 1, 2$ and plot each estimate as a point on top of your previous function plot. Do the results from your simulations match theory from Exercise 3a?

   (b) Now do the same for the variance $\sigma^2(\theta)$. Plot $\sigma^2$ as a function of $\theta$, as derived in Exercise 3b. Then, compute the variance numerically by finding a vector of times between spikes $k_{1:N}$, and computing 'np.var().' Do this in the specific cases $\theta = 0.1, 0.5, 1, 2$, and plot the estimates of variance as points on top of your previous function plot. Do they match theory from Exercise 3b?

   (c) Finally, check your estimate of the distribution of spike times. You can again borrow code on the leaky noisy integrate-and-fire model to do this. In the specific case where $\theta = 1$, plot the distribution $f(t; \mu(1), \sigma(1))$ from Exercise 3c. Then, run simulations, accumulate spike times in a vector, and use 'np.histogram', rescaling, and 'plt.bar' to generate a discrete estimate of the spike time distribution. How close it deviates from the normal estimate from Exercise 3c? Explain the origin of the deviation.

6. **Fitting Poisson spiking model to itself (maximum likelihood).** In this exercise, you will generate inter-spike times from an exponential distribution, and use this to define a Bayesian posterior over the possible rates of a Poisson process. You will be aided by python's built-in exponential random variable sampler.

   Write python code to perform the following steps:

   - Generate 1 sample $x_1 \sim \exp[1]$ from an exponential random variable with rate $\lambda = 1$ using the function 'np.random.exponential()'.
   - Now compute the likelihood function $L(\lambda|x_1)$ over possible arrival rates $\lambda$ given this observation, assuming an improper flat prior $p_0(\lambda)$, so by Bayes' rule:

   $$p(\lambda|x_1) = \frac{p(x_1|\lambda)p_0(\lambda)}{p(x_1)} = \frac{\lambda e^{-\lambda x_1}}{p(x_1)} = \frac{L(\lambda|x_1)}{p(x_1)} \quad \to \quad L(\lambda|x_1) = \lambda e^{-\lambda x_1}.$$

     Plot $L(\lambda|x_1)$ as a function of $\lambda$. Also, find the maximum of this function (the maximum likelihood) and plot it as a dot on the function. You may use calculus to find the maximum if you wish.
   - Next, compute 5 samples $x_{1:5}$, each from $\exp[1]$, and plot the associated likelihood function $L(\lambda|x_{1:5})$, given by

   $$p(\lambda|x_{1:5}) = \frac{p(x_{1:5}|\lambda)p_0(\lambda)}{p(x_{1:5})} = \frac{\lambda^5 e^{-\lambda \sum_{j=1}^{5} x_j}}{p(x_{1:5})} = \frac{L(\lambda|x_{1:5})}{p(x_{1:5})} \quad \to \quad L(\lambda|x_{1:5}) = \lambda^5 e^{-\lambda \sum_{j=1}^{5} x_j}.$$

     along with the maximum likelihood point. Is it closer or further from the true $\lambda = 1$ than when you used one sample?
   - Next compute 10 samples $x_{1:10}$ and plot $L(\lambda|x_{1:10}) = \lambda^{10} e^{-\lambda \sum_{j=1}^{10} x_j}$ along with the maximum likelihood point. Again, is this close or further from the true $\lambda$ than when you used five samples?
   - Repeat the above with 100 samples and $L(\lambda|x_{1:100}) = \lambda^{100} e^{-\lambda \sum_{j=1}^{100} x_j}$. How does the likelihood function look now? What do you think would happen if you increased the number of samples even more?