Probability Problem Set
42-699B/86-595 Neural Data Analysis
Due Tuesday 9/18/12 by noon.

1. (10 pts.) Prove that the variance of the Poisson distribution is equal to its mean.

2. (10 pts.) You are recording from two neurons. X_1 emits spikes with Poisson distributed counts at rate λ_1, and X_2 emits spikes with Poisson distributed counts at rate λ_2. Unfortunately, they have waveforms that are almost impossible to distinguish, and so when recording you lump them into one neuron, Z (the number of spikes we attribute to Z will be X_1+X_2). What is the distribution of Z's spike counts, assuming X_1 and X_2 are independent?
(Hint – remember the binomial theorem.)

3. (10 pts.) You are recording from a neuron that, in response to a particular stimulus, emits spike counts in a given interval according to a Poisson distribution with rate parameter λ. The spike waveform happens to be close to the threshold you've set for detection of a spike, so that each spike is detected with only 90% probability, independent of every other spike. What is the distribution of spike counts that you will record? Don't just state it, prove it.
(Hint – remember the Taylor series expansion of e^x.)

4. (10 pts.) Say you are measuring the intracellular membrane voltage of a neuron. Unbeknownst to you, there is a problem with your volt meter that causes the meter reading, V, to deviate from the true value, υ, such that the values V can be considered Gaussian random variables with mean $\mu=(\upsilon+b)$ and variance σ^2. (b here is called a bias – the measurement overestimates the true value on average by the bias). What is the mean squared error of your measurement?
(Note the mean squared error is defined as $MSE \equiv E[(V-\upsilon)^2]$.)

5. (10 pts.) Say you wish to design an experiment to measure reaction time: the subject must move as quickly as possible after he receives a cue. You don't want the subject to be able to anticipate the cue at all - so you decide to pull the cue time from a random distribution. What distribution should you use?

   5a. (5 pts.) Suppose you choose to draw the cue from a continuous distribution with pdf f(t) and cdf F(t). Prove that $\lambda(t)$, the probability of getting the cue in the next infinitesimal increment of time, given that you haven't received the cue up to time t, is equal to:

$$\lambda(t) = \frac{f(t)}{1 - F(t)}$$

   This is known as the *hazard function*.

5b. (5 pts.) Compute the hazard function for the exponential distribution and the Uniform(0,1) distribution. Which makes a better distribution for your cue times, and why?

6. (45 pts.) Matlab simulation: The threshold of action potential generation can affect correlation measurements.

What can the firing of one neuron tell us about how other neurons in the brain are responding? One way to study this process is by measuring the correlation between the firing rates of the two neurons. Understanding whether neurons change their firing rates in coordinated or independent ways can tell us a lot about the fundamental mechanisms the brain uses to store and represent information.

We learned in class that correlation measures *linear* dependence. However, we know that action potential generation is a non-linear process: membrane voltage must exceed a threshold value to generate a spike. You will simulate this thresholding operation and investigate its effect on firing rate correlations.

Assume that two neurons have correlated inputs that cause correlated intracellular membrane potentials. To model this process, you will model the intracellular voltage of each neuron as a random draw from a bivariate Gaussian. That is, assume that the voltage in the two neurons, $V=[v_1,v_2]^T$ is a random vector, V~Gaussian($\boldsymbol{\mu}$,$\boldsymbol{\Sigma}$):

$$\boldsymbol{\mu} = \begin{bmatrix} \mu \\ \mu \end{bmatrix}, \ \boldsymbol{\Sigma} = \begin{bmatrix} 1 & \rho_v \\ \rho_v & 1 \end{bmatrix}$$

Here, μ is the mean voltage in each neuron (we are assuming they have the same mean), and ρ_v is the correlation in voltage between the two neurons. We assume the variance for each neuron is 1.

Now model the thresholding of the neuron in the following way: assume that each neuron *i* turns its intracellular membrane potential into firing rate f_i through the equation

$$f_i = \begin{cases} v_i, & vi \geq 0 \\ 0, & otherwise \end{cases}$$

6a. (7 pts.) Set μ=2, ρ_v = 0.2, and draw 10000 random samples from V. (Look up the command mvnrand to see how to do this.) Use these to generate f_1 and f_2. Plot f_2 as a function of f_1. What is the correlation ρ_f you measure between f_1 and f_2?

6b. (7 pts.) Repeat 5a, for μ = -0.5.

6c. (7 pts.) Compute ρ_f for μ=[-2:0.25:2]. Plot ρ_f as a function of μ.

6d. (7 pts.) Repeat 5c, for $\rho_v = 0.9$.

6e. (7 pts.) Repeat 5c, for $\rho_v = -0.9$.

6f. (10 pts.) Write a paragraph describing how thresholding might affect correlation measurements in neurons.

7. (5 pts.) About how much time did you spend on each question of this problem set? Which problem taught you the most, and which taught you the least?