BioMath - Probability - September 3, 2015

There is a list of formulas on the last sheet of the problem set. Refer to them if you need to.

1. **Mean and variance.** Let $\mathbf{x} = [1\ 2\ 3\ 4\ 5\ 10\ 15\ 20]$; and $\mathbf{p} = [0.1\ 0.1\ 0.1\ 0.1\ 0.2\ 0.2\ 0.1\ 0.1]$; represent outcomes and probabilities, respectively.

   • Calculate the mean μ, the variance σ^2, and the standard deviation σ of this distribution.

   • Now change $\mathbf{x}$ to $\mathbf{x}_{new} = \mathbf{x} + 15$. What are the mean, variance, and standard deviation now? How do they compare to the original values?

   • Starting from the original $\mathbf{x}$, now change $\mathbf{x}$ to to $\mathbf{x}_{new} = \mathbf{x} * 10$. What are the mean, variance, and standard deviation now? How do they compare to the original values?

   • Generally speaking, what happens to the mean, variance, and standard deviation if you add a constant α to the vector $\mathbf{x}$? What happens to the mean, variance, and standard deviation if you multiply $\mathbf{x}$ times some constant β? So if x is a distribution with mean 5 and variance 3, what is the mean, variance, and standard deviation of $y = 3 * x + 2$?

2. **The Poisson distribution**. Remember that for a Poisson distribution with mean μ, the probability of that the outcome is $x = k$, where k is an integer, is $p(x = k) = \frac{\mu^k}{k!}e^{-\mu}$.

   • (Extra optional: this part of the problem is another example of sampling from a distribution.) Write a Matlab function that, given a random number that is uniformly distributed between 0 and 1, and given μ, returns a sample x from the Poisson distribution with mean μ.

   • Show that the variance of a Poisson distribution is equal to μ.

   • Plot the Poisson distribution, i.e., $P(k)$ as a function of k, for $\mu = 0.1$, $\mu = 1$, $\mu = 10$, and $\mu = 100$. How does the shape change? Does it become less or more symmetrical?

   • The "coefficient of variation (CV)" is often defined as standard deviation over the mean (sometimes it's defined as variance over mean squared)– it literally tells you how much variation there is relative to the mean. How does the CV Poisson distribution change as μ grows?

   • Suppose x is drawn from a Poisson distribution with mean 2 and y is drawn from a Poisson distribution with mean 1.5. Use Matlab to empirically find the distribution of $z = x + y$. Compare the distribution of z to a Poisson distribution with mean 3.5. Similar? Different? Speculate on whether the sum of two Poisson distributions is also a Poisson distribution, yes or no. In the second probability lecture, we'll go over the tools needed to prove the answer.

3. **Conditional distributions**

You're doing a development experiment in which on each attempt of your experiment, you get either 1, or, two, or three dividing cells (we'll call the number of dividing cells y), and there are 1, or, two, or three genes of interest expressed (we'll call this number x). The possible outcomes occur with the following joint probability:

y=1	3/18	1/18	3/18
y=2	1/18	2/18	1/18
y=3	3/18	1/18	3/18
	x=1	x=2	x=3

• Compute the marginal distributions $p(x)$ and $p(y)$.

• What is the mean of x? What is its variance? What is its standard deviation?

• Compute the covariance between x and y for this distribution: $\mathrm{cov}(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle$. (Don't forget to subtract those means!) Based on your result, are x and y correlated or not? Based on that, would you conclude that the value of x affects the value of y, yes or no?

• What is $p(x = 1|y = 1)$? What is $p(x = 1|y = 2)$? Are x and y independent? How does this compare to what you found looking at their covariance?

4. **Joint probability distributions, independence and bayes' rule– synaptic protein staining**

   You're busy as a beaver, and in the same week, you've done yet another totally different experiment. This time, you're looking over neurons in a slice. You are interested in the relationship between presynaptic terminals and number of dendritic branches. You've stained for **synapsin**, a marker of presynaptic terminals. For each neuron, you count the number of dendritic branches, and you also count the number of synapsin puncta that are close to that neuron's dendrites. You do the experiment twice, once with a slice from a wild-type mouse, and the nenxt time with a slice from an MHC-knockout mouse. Load `Pjoint.mat`. You'll find two joint probability distributions, representing the outcome of the two experiments. You can plot them with:

```
>> figure(1);
>> imagesc(nbranches, npuncta, Pjoint1); axis xy;
>> xlabel('nbranches'); ylabel('npuncta'); colormap jet;
colorbar

>> figure(2);
>> imagesc(nbranches, npuncta, Pjoint2); axis xy;
>> xlabel('nbranches'); ylabel('npuncta'); colormap jet;
colorbar
```

$Pjoint1(m, n)$ is the fraction of neurons for which you counted n dendritic branches and m synpasin puncta.

• For each of the two distributions (two matrices) that you were given, calculate the marginal distributions of number of branches and of number of puncta. Plot them and compare the result from Pjoint1 to the result from Pjoint2 to each other.

• Take Pjoint1, calculate P(number of branches = $b|$ number of puncta= n). Plot this as a function of b. Do it for all of the possible values of n, putting all the plots on top of each other. Are the number of branches independent of the number of puncta? Repeat this for Pjoint2. How about now, are the number of branches independent of the number of puncta?

• Show that if x and y are independent, i.e., $P(x, y) = P(x)P(y)$, then $P(x|y = y_0) = P(x)$ for any y_0. Does this match what you saw in the previous bullet?

5. **Maximum Entropy**

Consider a weighted coin that gives heads with probability p and tails with probability $1-p$. We want to know what the entropy of this distribution is, a s a function of p.

[Note that $\log(0) = -\infty$, but, very helpfully the limit of $plog(p)$ as $p \to 0$ is 0. So we define $0 * log(0) = 0$.]

• plot the entropy H as a function of p. For what value(s) of p is H the biggest? For what value(s) of p is H the smallest? Does this make sense in terms of H quantifying the degree of disorder or unpredictability of the coin?

For the next bullet point, we're going to use *Lagrange multipliers*, which help to find extreme of functions under constraints. We won't prove how they work here. For now, just take it under faith that if you have a function $f(p)$, and you want to find an extremum over p (i.e., a maximum or minimum) subject to the constraint $c(p) = 0$, then the extremum will satisfy

$$\frac{\partial f}{\partial p} + \lambda \frac{\partial c}{\partial p} = 0 \tag{1}$$

where λ is what is known as a Lagrange multiplier, and is some constant that will be chosen such that $c(p) = 0$.

Sounds a bit loopy, I know. Let's use it and step through it and we'll make it work. Consider a probability distribution defined by a vector $\mathbf{p} = [p_1\ p_2\ ...\ p_i\ ...\ p_N]$. What we want to do is ask: "of all possible distributions (i.e., all possible vectors $\mathbf{p}$ subject to $\sum_i p_i = 1$), which one has the maximum entropy H?"

• Write down equation (1) for the case where f is the entropy of that distribution, the constraint is $\sum_i p_i - 1 = 0$, and the parameter p that we're trying to find the extremum over is p_k, the k^{th} entry in the vector $\mathbf{p}$.

• Show that $p_k = 1/N$ will satisfy the constraint and the extremum equation that you just wrote. This shows that for a discrete distribution defined over a finite set of N outcomes, the maximum entropy distribution is the uniform distribution. In other words, for a finite set of N outcomes, the most disordered or least predictable distribution is the uniform distribution.

Formulae

• Poisson distribution

$$F(k|\mu) = \frac{\mu^k}{k!} e^{-\mu} \tag{2}$$

k is the number of events
μ is the expected (average) number of events

• Exponential distribution

$$D(t|r) = re^{-rt} \tag{3}$$

r is the expected rate of events

• Gaussian distribution

$$D(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \tag{4}$$

μ is the mean
σ is the standard deviation

• Bayes rule

$$P(X, Y) = P(X|Y)P(Y) = P(Y|X)P(X) \tag{5}$$

X and Y are two random variables