Tutorial problems for Lectures 1, 2, and 3. Due Wednesday September 19th.

Problem 1: Show that a Poisson distribution converges to a Gaussian in the limit of large λ. Hints - use Stirling's approximation plus use the first two terms in a log expansion.

Problem 2: The gold standard for a believable result is usually 5σ. Let's define the Gaussian approximation as "good enough" if it agrees with the Poisson to within a factor of 2. How large does n need to be for the Gaussian to be good enough at 5σ? How about at 3σ?

Problem 3: Let's say we have n Gaussian-distributed data points with identical standard deviations σ, and identical but unknown mean. What is the error on the maximum-likelihood estimate of the mean? Now let's say we got the errors on half the data wrong by a factor of $\sqrt(2)$ (so the variance is off by a factor of 2). What is the true error on the new non-optimal mean, and how does it compare to the maximum-likelihood you could have gotten had you gotten the noises right? How about if you underweight 1% of the data by a factor of $\sim$100? And if you overweight 1% of the data by a factor of 100? What type of mistake in weighting your data should you be most concerned about?

Problem 4: In linear least-squares, our estimate for fit parameters m is unbiased if $< m >= m_{true}$. If our model is correct, $< d >= \mathrm{A}m$, then show that the least-squares solution is unbiased. Show that this result does not depend on our noise matrix N actually being the noise in the data.

Problem 5: The preceding statement comes with an important caveat, namely that our noise estimate is not correlated with any residual signal in the data. Write a computer program that generates random Gaussian noise (numpy.random.randn may come in handy here), and adds a template (possibly a Gaussian as would be typical for a source seen by a telescope with a finite-resolution beam, but the details aren't important) to it. Estimate the noise by assuming it's constant and equal to the scatter in the observed data, which has the template added to it. Show that the least-squares estimate for any individual chunk is unbiased, but that the least-squares estimate for many data chunks analyzed jointly is biased low. Basically, your program should fit an amplitude and error to each individual chunk, then use that to get an overall amplitude/error. How might you go about mitigating this bias? Note that this is an extremely common situation when you say observed the same field/source several times and want to make your "best" estimate of what you have seen.

Problem 6 - bonus: In class it was asserted that adding orthogonal matrices into the expression for χ^2 let us work with correlated data. In particular, show that

$$\chi^2 = \delta^T V^T V N^{-1} V^T V \delta$$

, where $\delta_i = d_i - < d_i >$, is equivalent to

$$\chi^2 = \tilde{\delta}^T \tilde{N}^{-1} \tilde{\delta}$$

and that

$$\tilde{N}_{ij} =< \tilde{\delta}_i \tilde{\delta}_j >$$