Tutorial problems for lectures 1-4. Due Wednesday September 21

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 class, we showed that the maximum likelihood estimator is unbiased, in the sense that $< m >= m_{true}$. This claim comes with an important caveat, namely that our noise estimate is not correlated with any residual signal in the data, *i.e.* that $< \mathrm{N}^{-1}n >= 0$. It's painful to analytically estimate the impact of $\mathrm{N} - n$ coupling, but relatively straightforward on a computer.

Write a python program that generates random Gaussian noise (numpy.random.randn may come in handy here), and adds a signal. Please use 51 points from -5 to 5 (x=np.linspace(-5,5,51)), and let the template signal be a unit Gaussian ($\exp(-x^2/2)$). Estimate the noise by assuming it's constant within a chunk and equal to the scatter in the observed data, which has the signal in 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. Your program should fit an amplitude and error to each individual chunk, then use that to get an overall amplitude/error using the weighted average of the individual chunks. What is the bias in your estimate of the template amplitude? How might you go about mitigating this bias? NB - I see the bias very clearly when averaging over 10,000 chunks.

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 5: In class we will see that we can add rotation matrices into our expression for χ^2 to get a new expression for chi^2 that has a non-diagonal noise matrix. Show that in the new, rotated space, $\mathrm{N}_{ij} = \langle n_i n_j \rangle$. This is the key result that lets us use the exact same mathematical framework with correlated noise that we developed for uncorrelated noise.