1 Problem 1

Poisson distribution: $\frac{\lambda^n e^{-\lambda}}{n!}$, where λ is the expectation (and the variance), n is the number of occurrence.

Gaussian distribution: $e^{\frac{-(n-\lambda)^2/2\sigma^2}{\sqrt{2\pi\sigma^2}}}$, where σ is the standard deviation.

It's easier to work in log space, so let's take the log of Poisson distribution:

$$\log(\mathrm{PDF}) = -\lambda + \mathrm{n}\log(\lambda) - \log(\mathrm{n}!)$$

Now use Stirling's approximation: $n! \sim \sqrt{2\pi n}(n/e)^n$:

$$\log(\mathrm{PDF}) = -\lambda + \mathrm{n}\log(\lambda) - \mathrm{n}\log(\mathrm{n}) + \mathrm{n} - \frac{1}{2}\log(2\pi\mathrm{n})$$

Now let $n = \lambda + \delta$:

$$\log(\mathrm{PDF}) = -\lambda + (\lambda+\delta)\log(\lambda) - (\lambda+\delta)\log(\mathrm{n}) + (\lambda+\delta) - \frac{1}{2}\log(2\pi\mathrm{n})$$

$$\log(\mathrm{PDF}) = -(\lambda+\delta)\log\left(1+\frac{\delta}{\lambda}\right) + \delta - \frac{1}{2}\log(2\pi\mathrm{n})$$

Expand the log to second order: $\log(1+\epsilon) \sim \epsilon - \frac{\epsilon^2}{2}$:

$$\log(\mathrm{PDF}) = -(\lambda+\delta)(\frac{\delta}{\lambda} - \frac{\delta^2}{2\lambda^2}) + \delta - \frac{1}{2}\log(2\pi\mathrm{n})$$

$$\log(\mathrm{PDF}) = -\delta - \frac{\delta^2}{\lambda} + \frac{\delta^2}{2\lambda} + \frac{\delta^3}{2\lambda^2} + \delta - \frac{1}{2}\log(2\pi\mathrm{n})$$

We can ignore the δ^3 term because at the limit of large λ it won't matter much. So,

$$\log(\mathrm{PDF}) = -\frac{\delta^2}{2\lambda} - \frac{1}{2}\log(2\pi\mathrm{n})$$

Again at the limit of large λ, $n \approx \lambda$,

$$\log(\mathrm{PDF}) = -\frac{\delta^2}{2\lambda} - \frac{1}{2}\log(2\pi\lambda)$$

This is the log form of the Gaussian distibution:

$$\frac{e^{\delta^2/2\lambda}}{\sqrt{2\pi\lambda}}$$

, where $\delta = n - \lambda$, λ is both the variance and the mean.

2 Problem 2

See code at the end. 3σ agreement can be achieved at a few datapoints, while 5σ needs over 500. Because Poisson distribution is asymmetric, the exact number you get will be different depending on which side you approach this criteria or how you define "within a factor of two".

3 Problem 3

3.1

The mean of this distribution is $\frac{\Sigma(x_i)}{n}$, and we need to find the variance of it. We know that the variance of ax is $a^2Var(x)$, so the variance of $\frac{x_i}{n}$ is $\frac{\sigma^2}{n^2}$. We also know that the variance of the sum is the sum of the variance, so the variance of this mean should be:

$$\frac{\Sigma\sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma}{n}$$

So the standard error is $\frac{\sigma}{\sqrt{n}}$.

3.2

In class we learned that the expectation of least square estimate is essentially a weighted mean in the form of $\frac{\Sigma w_i d_i}{\Sigma w_i}$, where $w_i = \frac{1}{\sigma_i^2}$. So if we misweight half the data, then we are looking at the variance of $\frac{2x_1+x_2}{3}$. The variance is:

$$\frac{4\sigma^2}{9} + \frac{\sigma^2}{9} = \frac{5\sigma^2}{9}$$

If we didn't the mistake, we would be evaluating the variance of $\frac{x_1+x_2}{2}$, which is:

$$\frac{\sigma^2}{4} + \frac{\sigma^2}{4} = \frac{\sigma^2}{2}$$

The ratio is $(\frac{5}{9})/\frac{1}{2} = \frac{10}{9}$, meaning that our mistake gave us a $\frac{1}{9}$ bigger variance (5% increase in error bar), which is not as bad as I would have imagined..

3.3

If we throw away the datapoint, the difference in variance would go from $\frac{n}{100}$ to $\frac{n}{99}$, enlarging the error bar by 0.5%, which is not a big deal. However, if we decide to include the datapoint, we are looking at the variance of $\frac{100x_1+\Sigma_2^101x_i}{200}$, which is:

$$\frac{10^4\sigma^2+100\sigma^2}{4\times10^4}=\sigma^2/4$$

It's equivalent to the variance of only 4 good datapoints, which means that we have essentially thrown away 95% of our data!

4 Problem 4

See code example for how to calculate this. Basic problem is that if noise anti-correlates with the signal then we get a low answer, but also a low estimate of the noise. So, when we do the weighted average of a bunch of measurements, the total average is biased low. There are a bunch of ways of addressing this, depending on what your data look like. One would be to subtract off the best-fit signal, then re-estimate the noise. Repeat this a few times until the mean quits changing. If you think the noises ought to be similar, you could treat all your measurements the same, which we have seen gives an unbiased result. You might also try using the median instead of the standard deviation for the noise estimate. That will be quite a bit more robust to these problems as well.

5 Problem 5

In class we have seen that in the rotated space, $\chi^2=\tilde{\delta}^T\tilde{N}^{-1}\tilde{\delta}$, where $\tilde{\delta}=\tilde{d}-\tilde{A}(m)$. And this can be related back to the original space by $\tilde{\delta}=V\delta$, $\tilde{N}=VNV^T$, where $\delta=d-A(m)$, V and V^T are orthogonal matrices. So in this rotated space, the covariance is $<\tilde{\delta}\tilde{\delta}^T>$, where $\tilde{\delta}=V\delta$. Hence,

$$cov=<V\delta(V\delta)^T>=<V\delta\delta^TV^T>=VNV^T=\tilde{N}$$