Marcus Merryfield

PHYS 641 Assignment 1

1.) Poisson dist^{n}: $P(X=x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$ $\quad x = 0,1,2,\ldots$

In the limit of large λ, show that Poisson dist^{n} $\to$ Normal dist^{n}

Sterling's Approximation: $x! \approx \sqrt{2\pi x}\left(\dfrac{x}{e}\right)^x$

good approx. if x is large. Since λ large, P is only appreciable if x also is large. So this approximation is quite good

$$P(X=x) \approx \frac{\lambda^x e^{-\lambda} e^x}{\sqrt{2\pi x}\, x^x}$$

in order to expand this, we will approximate. $x = \lambda + y$, where $|y| \ll \lambda$. Once again, this is since P is only appreciable when x is near λ.

$$P(X=x) \approx \frac{\cancel{\lambda^{\lambda+y}}\, \cancel{e^{-\lambda}}\, e^{\cancel{\lambda}+y}}{\sqrt{2\pi(\lambda+y)}\, (\cancel{\lambda}(1+\frac{y}{\lambda}))^{\lambda+y}}$$

approximate $\sqrt{\lambda+y} \approx \sqrt{\lambda}$ introducing negligible error as $(1+\frac{y}{\lambda}) \approx$

$$\ln P = y - \tfrac{1}{2}\ln(2\pi\lambda) - (\lambda+y)\ln(1+\tfrac{y}{\lambda})$$

Since $|\frac{y}{\lambda}| \ll 1$, expand $\ln(1+\frac{y}{\lambda}) \approx \cancel{\ln(1)}^{0} + \dfrac{y}{\lambda} - \dfrac{y^2}{2\lambda^2} + \ldots$

$$\ln P \approx y - \frac{1}{2}\ln(2\pi\lambda) - (\lambda+y)\left[\frac{y}{\lambda} - \frac{y^2}{2\lambda^2}\right]$$

$$\ln P \approx \cancel{y} - \frac{1}{2}\ln(2\pi\lambda) - \cancel{y} + \frac{y^2}{2\lambda} - \frac{y^2}{\lambda} + \frac{y^3}{2\lambda^2}$$

get rid of terms $\mathcal{O}\left(\frac{y^2}{\lambda^2}\right)$ as $\frac{y}{\lambda} \ll 1$:

$$\ln P \approx -\frac{1}{2}\ln(2\pi\lambda) - \frac{y^2}{2\lambda}$$

$$P \approx \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{y^2}{2\lambda}\right)$$ recall: $x = \lambda + y$ $\Rightarrow y = x - \lambda$

$$\Rightarrow P \approx \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{(x-\lambda)^2}{2\lambda}\right)$$

Given $\sigma^2 = \lambda$ for Poisson distributions, this form matches the Gaussian distribution.

2.) See Email w/ attached Github Repo!

3.) n Gaussian distributed data pts., all with true stdev $= \sigma$. All have true (unknown) mean λ. What is ML error on $\hat{\lambda}$?

model: $\langle d \rangle = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \lambda$ $\quad \hat{\lambda} = \frac{\sum w_i d_i}{\sum w_i} \quad w_i = \frac{1}{\sigma_i^2}$

all errors are the same: $\sigma_i = \sigma$: $\quad \hat{\lambda} = \frac{\sum^n \frac{d_i}{\sigma^2}}{\sum^n \frac{1}{\sigma^2}} = \frac{\frac{d_1}{\sigma^2}}{\sum^n \frac{1}{\sigma^2}} + \frac{\frac{d_2}{\sigma^2}}{\sum^n \frac{1}{\sigma^2}} + \dots$

$$\Rightarrow \mathrm{Var}(\hat{\lambda}) = \frac{\sum^n \frac{\mathrm{Var}(d_i)}{\sigma^4}}{\left(\sum^n \frac{1}{\sigma^2}\right)^2} = \frac{n \frac{\cancel{\sigma^2}}{\cancel{\sigma^4}\,\sigma^2}}{\left(n\frac{1}{\sigma^2}\right)^2} = \frac{\sigma^2}{n}$$

$$\text{error on } \hat{\lambda} = \Delta\lambda = \sqrt{\mathrm{Var}(\hat{\lambda})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

Now, if we get half of the errs on our data wrong by a factor $\sqrt{2}$, what is the true error on the new non-optimal mean?

$$\hat{\lambda} = \frac{\sum^{n/2} \frac{d_i}{\sigma^2} + \sum^{n/2} \frac{d_i}{(\sqrt{2}\,\sigma)^2}}{\sum^{n/2} \frac{1}{\sigma^2} + \sum^{n/2} \frac{1}{(\sqrt{2}\sigma)^2}}$$

$$\Rightarrow \mathrm{Var}(\hat{\lambda}) = \frac{\sum^{n/2} \frac{\mathrm{Var}(d_i)}{\sigma^4} + \sum^{n/2} \frac{\mathrm{var}(d_i)}{4\sigma^4}}{\left(\sum^{n/2} \frac{1}{\sigma^2} + \sum^{n/2} \frac{1}{2\sigma^2}\right)^2}$$

$$= \frac{\frac{n\cancel{\sigma^2}}{2\sigma^{4\,2}} + \frac{n\cancel{\sigma^2}}{8\sigma^{4\,2}}}{\left(\frac{n}{2\sigma^2} + \frac{n}{4\sigma^2}\right)^2} = \frac{\frac{5n}{8\sigma^2}}{\left(\frac{3n}{4\sigma^2}\right)^2} = \frac{10}{9}\frac{\sigma^2}{n}$$

$$\Rightarrow \Delta\lambda = \frac{\sqrt{10}}{3}\frac{\sigma}{\sqrt{n}}$$

factor of $\frac{\sqrt{10}}{3} \approx 1.054$ different than mean when we get errs correct

Now we underweight 1% of data by factor of 100: f/ 1%, $w_i = \frac{1}{100\sigma^2}$

$$\hat{\lambda} = \frac{\sum^{99n/100} \frac{d_i}{\sigma^2} + \sum^{n/100} \frac{d_i}{100\sigma^2}}{\sum^{99n/100} \frac{1}{\sigma^2} + \sum^{n/100} \frac{1}{100\sigma^2}}$$

$$\Rightarrow \mathrm{Var}(\hat{\lambda}) = \frac{\frac{99n}{100}\left[\frac{\sigma^2}{\sigma^4}\right] + \frac{n}{100}\left[\frac{\sigma^2}{100^2\sigma^4}\right]}{\left(\frac{99n}{100}\frac{1}{\sigma^2} + \frac{n}{100}\frac{1}{100\sigma^2}\right)^2} = \frac{0.990001\frac{n}{\sigma^2}}{\left(0.9803\frac{n}{\sigma^2}\right)^2} = 1.009898\frac{\sigma^2}{n}$$

$$\Rightarrow \Delta\lambda = 1.0049\frac{\sigma}{\sqrt{n}}$$

Close to true error: underweighting not a big deal

overweight 1%: f/ 1%, $w_i = \frac{100}{\sigma^2}$

$$\hat{\lambda} = \frac{\sum^{99n/100} \frac{d_i}{\sigma^2} + \sum^{n/100} \frac{d_i(100)}{\sigma^2}}{\sum^{99n/100} \frac{1}{\sigma^2} + \sum^{n/100} \frac{100}{\sigma^2}}$$

$$\Rightarrow \quad Var(\hat{\lambda}) = \frac{\frac{99n}{100}\left[\frac{\sigma^2}{\sigma^4}\right] + \frac{n}{100}\left[\frac{100^2\sigma^2}{\sigma^4}\right]}{\left[\frac{99n}{100}\frac{1}{\sigma^2} + \frac{n}{100}\frac{100}{\sigma^2}\right]^2}$$

$$= \frac{100.99\frac{n}{\sigma^2}}{\left(1.99\frac{n}{\sigma^2}\right)^2} = 25.50\,\frac{\sigma^2}{n}$$

$$\Rightarrow \quad \underline{\underline{\Delta\lambda = 5.05\frac{\sigma}{\sqrt{n}}}}$$

errors are a factor of 5 wrong! Overweighting errors very impactful.

4.) Assume model is correct: $\langle d\rangle = A m_{true}$

least squares solution: $\hat{m} = (A^T N^{-1} A)^{-1} A^T N^{-1} d$

$$\begin{aligned}\langle\hat{m}\rangle &= \langle (A^T N^{-1} A)^{-1} A^T N^{-1} d\rangle \\ &= (A^T N^{-1} A)^{-1} A^T N^{-1} \langle d\rangle \qquad \text{but } \langle d\rangle = A m_{true} \\ &= (A^T N^{-1} A)^{-1} A^T N^{-1} A\, m_{true} \\ &= I\, m_{true}\end{aligned}$$

$$\Rightarrow \quad \underline{\underline{\langle\hat{m}\rangle = m_{true}}}$$

5.) Github!