Mattias Lazda

## Question # 1:

Poisson distribution: $P(\lambda) = \dfrac{\lambda^{k} e^{-\lambda}}{k!}$

First, use Stirling's Approximation to remove the factorial:

$$k! \approx \sqrt{2\pi k}\left(\frac{k}{e}\right)^{k}$$

$$\Rightarrow P(\lambda) = \lambda^{k} e^{-\lambda} \cdot \frac{e^{k}}{k^{k}} \cdot \frac{1}{\sqrt{2\pi k}} = \frac{1}{\sqrt{2\pi k}}\left(\frac{\lambda}{k}\right)^{k} e^{-(\lambda-k)}$$

Taking the log of both sides:

$$\ln P = \ln\left(\frac{1}{\sqrt{2\pi k}}\left(\frac{\lambda}{k}\right)^{k} e^{-(\lambda-k)}\right)$$

$$= \ln\left(\frac{1}{\sqrt{2\pi k}}\right) + \ln\left(\left(\frac{\lambda}{k}\right)^{k}\right) + \ln\left(e^{-(\lambda-k)}\right)$$

$$= -\frac{1}{2}\ln\left(2\pi k\right) + k\ln\left(\frac{\lambda}{k}\right) - \lambda + k$$

Recalling that $\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(x-1)^{n}}{n} \approx x - 1 - \frac{(x-1)^2}{2}$,
setting $x = \frac{\lambda}{k}$,

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

$$= -\frac{1}{2}\ln(2\pi k) + \cancel{\lambda} - \cancel{k} - \frac{(\lambda - k)^2}{2k} - \cancel{\lambda} + \cancel{k}$$

$$= -\frac{1}{2}\ln(2\pi k) - \frac{(\lambda - k)^2}{2k}$$

However, in the limit of $\lambda >> 1 \rightarrow k \approx \lambda$, so the above becomes:

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

$$\Longrightarrow \boxed{P\Big|_{\lambda>>1} \approx \frac{1}{\sqrt{2\pi\lambda}}\, e^{-\frac{(k-\lambda)^2}{2\lambda}}}$$

<u>Question #2</u>: See attached Jupyter Notebook.

## Question # 3:

$$\mathcal{L} = \prod_{i=1}^{N} G(x_i; \mu, \sigma) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}}$$

$$-2\ln\mathcal{L} = -2\sum_{i=1}^{N} -\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2} - 2\ln\left((2\pi\sigma)^{-1/2}\right)$$

$$= \sum_{i=1}^{N} \frac{(x_i-\mu)^2}{\sigma^2} + \ln(2\pi\sigma^2)$$

$$\chi^2 \equiv -2\ln\mathcal{L}$$

$$\frac{d\chi^2}{d\mu} = \frac{d}{d\mu}\left(\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{\sigma^2}\right) = \cancel{+}\sum_{i=1}^{N}\frac{\cancel{2}(x_i-\mu)}{\cancel{\sigma^2}} = 0$$

$$\Rightarrow \sum_{i=1}^{N} x_i - \sum_{i=1}^{N}\mu = 0$$

$$\Rightarrow \mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

To get the error:

$$\mathrm{Var}[\mu] = \mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N} x_i\right] = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}[x_i]$$

$$= \frac{1}{N}\sigma^2$$

$$\Rightarrow \boxed{\sigma_\mu = \frac{\sigma}{\sqrt{N}}}$$

This is the error on the maximum llhd estimate of the mean.

Next, using the equations for the weighted mean derived in class:

$$\mu = \frac{\sum_i w_i x_i}{\sum_i w_i}, \qquad w_i = \frac{1}{\sigma_i^2}$$

The general expression for the variance of the the weighted mean is

$$\mathrm{Var}[\mu] = \mathrm{Var}\left[\frac{\sum_i w_i x_i}{\sum_i w_i}\right]$$

$$= \frac{\sum_i w_i^2 \,\mathrm{Var}[x_i]}{\left(\sum_i w_i\right)^2} = \frac{\sum_i \frac{1}{\sigma_i^4}\sigma_i^2}{\left(\sum \frac{1}{\sigma_i^2}\right)^2} = \frac{\sum_i \cancel{\frac{1}{\sigma_i^2}}}{\left(\sum_i \frac{1}{\sigma_i^2}\right)^{\cancel{2}}}$$

$$= \frac{1}{\sum_i \frac{1}{\sigma_i^2}}$$ ← This is what we'll work off of.

Case 1: 50% of points have incorrect variance (factor of 2):

$$\mathrm{Var}[\mu] = \frac{1}{\frac{N}{2}\cdot\frac{1}{\sigma^2} + \frac{N}{2}\cdot\frac{1}{2\sigma^2}} = \frac{1}{\frac{N}{\sigma^2}\left(\frac{1}{2}+\frac{1}{4}\right)} = \frac{4}{3N}\sigma^2$$

$$\Rightarrow \boxed{\sigma_\mu = \frac{2}{\sqrt{3N}}\sigma}$$

The above result differs from our original estimate by $\frac{2}{\sqrt{3}}$.

**Case #2:** Underweight 1% of data by factor of 100. eg. $Var = \frac{1}{100}\sigma^2$

$$Var[\mu] = \frac{1}{\frac{0.99N}{\sigma^2} + 0.01N \frac{1}{\frac{\sigma^2}{100}}}$$

1% are underweighted

99% are normal

$$= \frac{\sigma^2}{N} \cdot \left( \frac{1}{\underbrace{0.99 + 1}_{\approx 2}} \right) = \frac{\sigma^2}{2N} \Rightarrow \boxed{\sigma_\mu = \frac{\sigma}{\sqrt{2N}}}$$

**Case #3:** Overweight 1% of data by 100, e.g. $Var = 100\sigma^2$

$$Var[\mu] = \frac{1}{\frac{0.99N}{\sigma^2} + \frac{0.01N}{100\sigma^2}} = \frac{\sigma^2}{N}\left( \frac{1}{\underbrace{0.99 + 0.00001}_{\approx 1}} \right)$$

$$\approx \frac{\sigma^2}{N} \Rightarrow \boxed{\sigma_\mu = \frac{\sigma}{\sqrt{N}}}$$

The idea (I think) is that assuming we had gotten it right the first time, and $\sigma_\mu = \frac{\sigma}{\sqrt{N}}$ was the true error, then it would be worse if we had set $Var = \frac{1}{100}\sigma^2$ for 1% than $Var = 100\sigma^2$ for 1%, since in the former case, the new error gets shifted by a much larger fraction than in the latter case.

Question #4: See attached Jypyter Notebook.

Question # 5:

$$\chi^2 = (d - A(m))^T V V^T N^{-1} V V^T (d - A(m))$$

$$= ((d - A(m))^T V)(V^T N^{-1} V)(V^T (d - A(m)))$$

$$= (d^T V - A(m)^T V)(V^T N^{-1} V)(V^T d - V^T A(m))$$

Assume $V$ orthogonal $\Rightarrow V^T V = I$

Let $\tilde{N}^{-1} = V^T N^{-1} V \Rightarrow \tilde{N} = (V^T N^{-1} V)^{-1} = V^T N V$

$*$ We wish to show that $\tilde{N}_{ij} = \langle \tilde{n}_i \tilde{n}_j \rangle$. $*$

Let $n \equiv d - d_t \xrightarrow[\text{frame}]{\text{rotated}} \tilde{n}_i = \tilde{d}_i - \tilde{d}_{t_i}$

where $\tilde{d}_i = V_{ij}^T d_j$, $\tilde{d}_{t_i} = V_{ij}^T d_{t_j}$

$$\Rightarrow \tilde{n}_i = V_{ij}^T (d_j - d_{t_j})$$

$$\Rightarrow \langle \tilde{n}_i \tilde{n}_j \rangle = \langle V_{ij}^T (d_j - d_{t_j})(V_{ji}^T (d_i - d_{t_i}))^T \rangle$$

$$= \langle V_{ij}^T (d_j - d_{t_j})(d_i - d_{t_i})^T V_{ij} \rangle$$

Recall: $\langle abc \rangle = \langle a \rangle \langle b \rangle \langle c \rangle$ if independent. Since our choice of $V_{ij}$ is arbitrary, the above must hold true for any $V_{ij}$.

$$= \langle V_{ij}^T \rangle \langle (d_j - d_{t_j})(d_i - d_{t_i})^T \rangle \langle V_{ij} \rangle$$

$$= V_{ij}^T \langle n_i n_j \rangle V_{ij}$$

$$= V_{ij}^T N_{ij} V_{ij}$$

$$= \tilde{N}_{ij} \blacksquare$$