MARCH 31

mean = 35.4 h

SD = 10.6 h

$$\mu_T = \ln\left(\frac{\text{mean}^2}{\sqrt{\text{mean}^2 + SD^2}}\right) = \ln\left(\frac{35.4^2}{\sqrt{35.4^2 + 10.6^2}}\right)$$

$$\sigma_T = \sqrt{\ln\left(1 + \frac{SD^2}{\text{mean}^2}\right)} = \sqrt{\ln\left(1 + \frac{10.6^2}{35.4^2}\right)}$$

$$L = \frac{1}{T\,\sigma_T\sqrt{2\pi}} \exp\left(-\frac{(\ln T - \mu_T)^2}{2\sigma_T^2}\right) \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{Y_i - A e^{-\frac{\ln(2)\cdot t_i}{T}}}{\sigma}\right)^2\right]$$

"L" is the posterior (unnormalized)

$$\ln L = -\ln\left(T\sigma_T\sqrt{2\pi}\right) - \frac{(\ln T - \mu_T)^2}{2\sigma_T^2} - N\ln\left(\sigma\sqrt{2\pi}\right)$$

$$-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left[y_i - Ae^{-\frac{\ln(2)}{T}t_i}\right]^2$$

$$-\ln L = \ln\left(T\sigma_T\sqrt{2\pi}\right) + \frac{(\ln T - \mu_T)^2}{2\sigma_T^2} + N\ln\left(\sigma\sqrt{2\pi}\right)$$

$$+\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left[y_i - Ae^{-\frac{\ln(2)}{T}t_i}\right]^2$$

$$= C$$

$$\frac{\partial C}{\partial A} = \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[y_i - Ae^{-\frac{\ln(2)}{T}t_i}\right]\left(-e^{-\frac{\ln(2)}{T}t_i}\right)$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[Ae^{-\frac{\ln(2)}{T}t_i} - y_i\right]e^{-\frac{\ln(2)}{T}t_i}$$

$$\frac{\partial^2 C}{\partial A^2} = \frac{1}{\sigma^2}\sum_{i=1}^{N} e^{-\frac{2\ln(2)}{T}t_i} = \frac{1}{\sigma^2}\sum_{i=1}^{N} \xi_i(T, t_i)$$

$$\frac{\partial^2 C}{\partial T\partial A} = \frac{\partial}{\partial T}\left\{\frac{1}{\sigma^2}\sum_{i=1}^{N}\left[Ae^{-\frac{2\ln(2)}{T}t_i} - y_i e^{-\frac{\ln(2)}{T}t_i}\right]\right\}$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[Ae^{-\frac{2\ln(2)\cdot t_i}{T}}\cdot\frac{2\ln(2)\cdot t_i}{T^2} - y_i t_i \cdot \frac{\ln(2)}{T^2}\cdot e^{-\frac{\ln(2)}{T}t_i}\right]$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N} h_i(A, T, t_i, y_i)$$

$$\frac{\partial C}{\partial T} = \frac{\sigma_T\sqrt{2\pi}}{T\sigma_T\sqrt{2\pi}} + \frac{2(\ln T - \mu_T)}{2\sigma_T^2}\cdot\frac{1}{T} + \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[y_i - Ae^{-\frac{\ln(2)}{T}t_i}\right] \cdot -Ae^{-\frac{\ln(2)\cdot t_i}{T}} \cdot \frac{\ln(2)t_i}{T^2}$$

$$\frac{\partial C}{\partial T} = \frac{1}{T} + \frac{(\ln T - \mu_T)}{T\sigma_T^2} + \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[Ae^{-\frac{\ln(2)}{T}t_i} - y_i\right] \cdot Ae^{-\frac{\ln(2)\cdot t_i}{T}} \cdot \frac{\ln(2)\cdot t_i}{T^2}$$

$$= \frac{1}{T} + \frac{(\ln T - \mu_T)}{T\sigma_T^2} + \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[A^2\cdot\frac{\ln(2)t_i}{T^2}e^{-\frac{2\ln(2)\cdot t_i}{T}} - Ay_i e^{-\frac{\ln(2)\cdot t_i}{T}}\cdot\frac{\ln(2)t_i}{T^2}\right]$$

$$\frac{\partial^2 C}{\partial T^2} = -\frac{1}{T^2} + \frac{\mu_T + 1 - \ln(T)}{T^2\sigma_T^2} + \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[\frac{2A^2(\ln(2)t_i)e^{-2(\ln 2)t_i/T}\cdot(\ln(2)t_i - T)}{T^4} - \frac{Ay_i(2\ln 2\cdot t_i)e^{-\ln(2)\cdot t_i/T}\cdot(\ln(2)\cdot t_i - 2T)}{T^4}\right]$$

$$= g(T) + \frac{1}{\sigma^2}\sum_{i=1}^{N} f_i(t_i, A, T, y_i)$$

$$\frac{\partial^2 C}{\partial A \partial T} = \frac{\partial^2 C}{\partial T \partial A}$$

Let A and T be the "true parameters":

$$\therefore \; I_{AA}(A,T) = \mathbb{E}\left(\frac{\partial^2 C}{\partial A^2}\right)$$

$$I_{TA}(A,T) = \mathbb{E}\left(\frac{\partial^2 C}{\partial T \partial A}\right) = I_{AT}(A,T)$$

$$I_{TT}(A,T) = \mathbb{E}\left(\frac{\partial^2 C}{\partial T^2}\right)$$

Linear operators, so we can separate terms corresponding to different time points t_i:

$$I = \begin{pmatrix} I_{AA} & I_{AT} \\ I_{TA} & I_{TT} \end{pmatrix} = \left(\begin{pmatrix} \mathbb{E}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N} \xi_i(T,t_i)\right] & I_{AT} \\ \mathbb{E}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N} h_i(A,T,t_i,y_i)\right] & \mathbb{E}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N} f_i(t_i,A,T,y_i)\right] \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}[g(T)] \end{pmatrix} \right)$$

$$= \sum_{i=1}^{N} \begin{pmatrix} \frac{1}{\sigma^2}\mathbb{E}[\xi_i(T,t_i)] & \frac{1}{\sigma^2}\mathbb{E}[h_i(A,T,t_i,y_i)] \\ \frac{1}{\sigma^2}\mathbb{E}[h_i(A,T,t_i,y_i)] & \frac{1}{\sigma^2}\mathbb{E}[f_i(t_i,A,T,y_i)] \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}[g(T)] \end{pmatrix}$$

$$\Rightarrow I = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}[g(T)] \end{pmatrix} + \sum_{i=1}^{N} \mathcal{I}_i(A, T, t_i)$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & g(T) \end{pmatrix} + \sum_{i=1}^{N} \mathcal{I}_i(A, T, t_i)$$

Notes: for a single (t_i, y_i),

$\frac{1}{\sigma^2}\mathbb{E}[\xi_i(T, t_i)] = \frac{1}{\sigma^2}\xi_i(T, t_i)$ (no dependence on y_i)

$$\Rightarrow I = \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} 0 & 0 \\ 0 & g(T) \end{pmatrix} + \sum_{i=1}^{N} \mathcal{I}_i(A, T, t_i)$$

$$\Rightarrow I = \sum_{i=1}^{N}\left[\mathcal{I}_i(A, T, t_i) + \frac{1}{N}\begin{pmatrix} 0 & 0 \\ 0 & g(T) \end{pmatrix}\right]$$

$$\Rightarrow I = \sum_{i=1}^{N} \mathcal{I}_i(A, T, t_i)$$

$\mathcal{I}_i$ got re-defined here.

Quantify "importance"/"value" of a data point by:

$$w_i = \det\left[\mathcal{I}_i(A, T, t_i)\right] = \lambda_{i1} \cdot \lambda_{i2}$$

where λ_{i1} and λ_{i2} are the eigenvalues of $\mathcal{I}_i$

Plan:

1. Select 2 points to estimate $\hat{A}$ and $\hat{T}$
2. Compute $\det\left[\mathcal{I}_i(\hat{A}, \hat{T}, t_i)\right]$ for all $i \in \{1, 2 \ldots N\}$
3. Find the 2 points with highest such "weights".
4. Repeat for <u>all</u> combinations of 2 points.
5. If there are "K" combinations, this leads to a $(K \times 2)$ matrix of most important indices, whose first column corresponds to indices that were ranked highest. Second column for indices ranked second highest.
6. Flatten this into a 1-D array and find the indices that occurred first and second most frequently.
7. These two indices correspond to our proposed two most valuable points.

How "expected $\frac{d^2C}{dTdA}$, $\frac{d^2C}{dT^2}$" were calculated:

Suppose we have two points (t_1, y_1), (t_2, y_2)

We showed:

$$\frac{d^2C}{dTdA} = f_1(A,T,t_1,y_1) + f_2(A,T,t_2,y_2)$$

$$\therefore E\left[\frac{\partial^2 C}{\partial T \partial A}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[f_1 + f_2]\, P(y_1|A,T)\, P(y_2|A,T)\, dy_1\, dy_2$$

$$= \int dy_1 \int dy_2\, f_1(y_1)\, P(y_1)\, P(y_2)$$

$$+ \int dy_1 \int dy_2\, f_2(y_2)\, P(y_1)\, P(y_2)$$

$$= \int dy_1\, f_1(y_1)\, P(y_1) \left[\int dy_2\, P(y_2)\right]$$

$$+ \int dy_2\, f_2(y_2)\, P(y_2) \left[\int dy_1\, P(y_1)\right]$$

$$= \int_{-\infty}^{\infty} f_1(A,T,t_1,y_1)\, P(y_1|A,T)\, dy_1$$

$$+ \int_{-\infty}^{\infty} f_2(A,T,t_2,y_2)\, P(y_2|A,T)\, dy_2$$