Math and Medicine: Homework Assignment 2
Answers

1. **Answer:** To find the tree, we follow the UPGMA algorithm.

	lc1	lc5	ptb	ptd	denptc
lc1	0	95	58	85.0	72.0
lc5	95	0	39	37.0	41.0
ptb	58	39	0	43.0	23.0
ptd	85	37	43	0.0	44.5
denptc	72	41	23	44.5	0.0

	lc1	lc5	ptd	ptbdenptc
lc1	0	95	85.00	65.00
lc5	95	0	37.00	40.00
ptd	85	37	0.00	43.75
ptbdenptc	65	40	43.75	0.00

	lc1	ptbdenptc	lc5ptd
lc1	0	65.000	90.000
ptbdenptc	65	0.000	41.875
lc5ptd	90	41.875	0.000

	lc1	ptbdenptclc5ptd
lc1	0	77.5
ptbdenptclc5ptd	77.5	0

Now the controls do not cluster together (lc1 is off on its own), making it less clear that the dentist and the patients form a separate group from the controls.

2. **Answer:**

a.

$$\begin{aligned} L(\Lambda) &= \Pr(10|\Lambda) = \frac{\Lambda^{10}e^{-\Lambda}}{10!} \\ S(\Lambda) &= \ln(L(\Lambda)) = 10\ln(\Lambda) - \Lambda - \ln(10!) \\ \frac{dS}{d\Lambda} &= \frac{10}{\Lambda} - 1 \end{aligned}$$

This has a maximum at $\Lambda = 10$ where it switches from increasing (positive derivative) to decreasing (negative derivative).

b. If D and P are closest, the likelihood function is

$$\begin{aligned} L_{DP}(\Lambda_1, \Lambda_2) &= \frac{\Lambda_1^{15}e^{-\Lambda_1}}{15!} + \frac{\Lambda_2^{38}e^{-\Lambda_2}}{38!} + \frac{\Lambda_2^{39}e^{-\Lambda_2}}{39!} \\ S_{DP}(\Lambda_1, \Lambda_2) &= 15\ln(\Lambda_1) - \Lambda_1 + 77\ln(\Lambda_2) - 2\Lambda_2 - H \end{aligned}$$

Here H is some huge constant with logs and factorials that is the same for all the models, and which disappears when we take the derivative. This will have a maximum where $\Lambda_1 = 15$ and $\Lambda_2 = 38.5$. We can find S_{DC} in the same way, with $\Lambda_1 = 38$ and $\Lambda_2 = 27$, and similarly, S_{PC} has $\Lambda_1 = 39$ and $\Lambda_2 = 26.5$. Then (this ignores H, which would subtract 237.50 from each of these values),

$$\begin{aligned} S_{DP}(15, 38.5) &= 229.72 \\ S_{DC}(38, 27) &= 224.20 \\ S_{DP}(39, 26.5) &= 224.57 \end{aligned}$$

The model with D and P closest is still the best.