Probabilistic Modelling and Reasoning, Tutorial Answer Sheet 3 (for Week 5)

School of Informatics, University of Edinburgh

Instructor: Amos Storkey
Prob. Sheet. ID: 11922.1.2
Thanks also to Matt Graham for some of the answers.

1.

$$
\begin{aligned}
P(n_h, n_t|\theta) &= \theta^{n_h}(1-\theta)^{n_t} && \text{(Binomial likelihood)} \\
\hat{\theta}_{\mathrm{ML}} &= \frac{n_h}{n_h+n_t} && \text{(Maximum likelihood solution, see last week's tutorial)} \\
p(\theta) &\propto \theta^{\alpha_h-1}(1-\theta)^{\alpha_t-1} && \text{(Beta prior)} \\
\mathrm{E}[\theta] &= \frac{\alpha_h}{\alpha_h+\alpha_t} && \text{(Prior mean, property of beta distribution)} \\
p(\theta|n_h, n_t) &\propto P(n_h, n_t|\theta)p(\theta) \propto \theta^{\alpha_h+n_h-1}(1-\theta)^{\alpha_t+n_t-1} && \text{(Posterior is also a beta distribution)} \\
\mathrm{E}[\theta|\mathcal{D}] &= \frac{\alpha_h+n_h}{\alpha_h+\alpha_t+n_h+n_t} && \text{(Posterior mean, property of beta distribution)} \\
&= \frac{\alpha_h}{\alpha+n} + \frac{n_h}{\alpha+n} && (\alpha = \alpha_h+\alpha_t,\ n = n_h+n_t) \\
&= \frac{\alpha}{\alpha+n}\frac{\alpha_h}{\alpha} + \frac{n}{\alpha+n}\frac{n_h}{n} \\
&= \lambda \mathrm{E}[\theta] + (1-\lambda)\hat{\theta}_{\mathrm{ML}} && (\lambda = \alpha/(\alpha+n))
\end{aligned}
$$

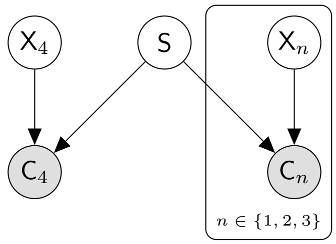
2.

$$
\begin{aligned}
\frac{d\log p}{d\theta} &= \frac{\alpha-1}{\theta} - \frac{\beta-1}{1-\theta} \\
\frac{\theta^*}{\alpha-1} &= \frac{1-\theta^*}{\beta-1} \qquad \text{at maximum} \\
\theta^* &= \frac{\alpha-1}{\alpha+\beta-2} \qquad \text{is the value of } \theta \text{ at maximum}
\end{aligned}
$$

For same positions of mean and mode

$$
\begin{aligned}
\frac{\alpha-1}{\alpha+\beta-2} &= \frac{\alpha}{\alpha+\beta} \\
\implies \alpha &= \beta
\end{aligned}
$$

3. The graphical model and conditional probabilities are given by.



$$\mathbb{P}\left[\mathsf{X}_n = x\right] = \begin{cases} \alpha & : x = 1 \\ 1-\alpha & : x = 2 \end{cases}$$

$$\mathbb{P}\left[\mathsf{S} = s\right] = \begin{cases} \beta & : x = 1 \\ 1-\beta & : x = 2 \end{cases}$$

$$\mathbb{P}\left[\mathsf{C}_n = c \,|\, \mathsf{X}_n = x, \mathsf{S} = s\right] = \begin{cases} 1-\lambda & : c = 1, x = 1, s = 1 \\ \lambda & : c = 2, x = 1, s = 1 \\ \delta & : c = 1, x = 2, s = 1 \\ 1-\delta & : c = 2, x = 2, s = 1 \\ \lambda & : c = 1, x = 1, s = 2 \\ 1-\lambda & : c = 2, x = 1, s = 2 \\ 1-\delta & : c = 1, x = 2, s = 2 \\ \delta & : c = 2, x = 2, s = 2 \end{cases}$$

Here $\alpha = 0.3, \beta = 0.25, \gamma = 0.1, \delta = 0.01$.

Given data $\boldsymbol{c}_{1:3} = \{1, 1, 2\}$ probability of next crowd-sourced image label

$$\begin{aligned}
&\mathbb{P}\left[\mathsf{C}_4 = 1 \,|\, \mathbf{C}_{1:3} = \boldsymbol{c}_{1:3}\right] \\
&= \sum_{x_4 \in \{1,2\}} \sum_{s \in \{1,2\}} \left\{\mathbb{P}\left[\mathsf{C}_4 = 1, \mathsf{X}_4 = x_4, \mathsf{S} = s \,|\, \mathbf{C}_{1:3} = \boldsymbol{c}_{1:3}\right]\right\} \\
&= \sum_{x_4 \in \{1,2\}} \sum_{s \in \{1,2\}} \left\{\mathbb{P}\left[\mathsf{C}_4 = 1 \,|\, \mathsf{X}_4 = x_4, \mathsf{S} = s, \cancel{\mathbf{C}_{1:3} = \boldsymbol{c}_{1:3}}\right] \mathbb{P}\left[\mathsf{X}_4 = x_4 \,|\, \cancel{\mathsf{S} = s}, \cancel{\mathbf{C}_{1:3} = \boldsymbol{c}_{1:3}}\right] \mathbb{P}\left[\mathsf{S} = s \,|\, \mathbf{C}_{1:3} = \boldsymbol{c}_{1:3}\right]\right\} \\
&= \sum_{s \in \{1,2\}} \left(\sum_{x_4 \in \{1,2\}} \left\{\mathbb{P}\left[\mathsf{C}_4 = 1 \,|\, \mathsf{X}_4 = x_4, \mathsf{S} = s\right] \mathbb{P}\left[\mathsf{X}_4 = x_4\right]\right\} \mathbb{P}\left[\mathsf{S} = s \,|\, \mathbf{C}_{1:3} = \boldsymbol{c}_{1:3}\right] \right)
\end{aligned}$$

Probability of supervisor identity given data

$$\begin{aligned}
\mathbb{P}\left[\mathsf{S} = s \,|\, \mathbf{C}_{1:3} = \{1,1,2\}\right] &\propto \mathbb{P}\left[\mathbf{C}_{1:3} = \{1,1,2\} \,|\, \mathsf{S} = s\right] \mathbb{P}\left[\mathsf{S} = s\right] \\
&= \prod_{n \in \{1,2,3\}} \left\{\mathbb{P}\left[\mathsf{C}_n = c_n \,|\, \mathsf{S} = s\right]\right\} \mathbb{P}\left[\mathsf{S} = s\right] \\
&= \prod_{n \in \{1,2,3\}} \left\{ \sum_{x_n \in \{1,2\}} \left(\mathbb{P}\left[\mathsf{C}_n = c_n \,|\, \mathsf{X}_n = x_n, \mathsf{S} = s\right] \mathbb{P}\left[\mathsf{X}_n = x_n\right]\right) \right\} \mathbb{P}\left[\mathsf{S} = s\right]
\end{aligned}$$

Substituting in values gives

$$
\begin{aligned}
\mathbb{P}\left[\mathsf{S}=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\{1,1,2\}\right] &\propto ((1-\lambda)\alpha+\delta(1-\alpha))^2\,(\lambda\alpha+\delta(1-\alpha))\,\beta \\
&= 0.277^2 \times 0.723 \times 0.75 = 0.0416 \\
\mathbb{P}\left[\mathsf{S}=2 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\{1,1,2\}\right] &\propto (\lambda\alpha+\delta(1-\alpha))^2\,((1-\lambda)\alpha+\delta(1-\alpha))\,(1-\beta) \\
&= 0.723^2 \times 0.277 \times 0.25 = 0.0362 \\
\therefore \mathbb{P}\left[\mathsf{S}=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\{1,1,2\}\right] &= \frac{0.0416}{0.0416+0.0362} = 0.535 = \epsilon \\
\text{and } \mathbb{P}\left[\mathsf{S}=2 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\{1,1,2\}\right] &= \frac{0.0362}{0.0416+0.0362} = 0.465 = 1-\epsilon
\end{aligned}
$$

Substituting this into above gives

$$
\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right] = \left[(1-\gamma)\alpha+\delta+\delta(1-\alpha)\right]\epsilon + \left[\gamma\alpha+(1-\delta)(1-\alpha)\right](1-\epsilon) = 0.485 = \zeta
$$

For final part, the probability of the true label being a car given the now observed crowd-sourced label $\mathsf{C}_4 = 1$

$$
\begin{aligned}
&\mathbb{P}\left[\mathsf{X}_4=1 \,|\, \mathsf{C}_4=1, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right] = \frac{\mathbb{P}\left[\mathsf{X}_4=1, \mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]}{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]} \\
&= \frac{\sum_{s\in\{1,2\}}\left\{\mathbb{P}\left[\mathsf{X}_4=1, \mathsf{C}_4=1, \mathsf{S}=s \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]\right\}}{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]} \\
&= \frac{\sum_{s\in\{1,2\}}\left\{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \mathsf{X}_4=1, \mathsf{S}=s, \cancel{\boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}}\right]\mathbb{P}\left[\mathsf{X}_4=1, \mathsf{S}=s \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]\right\}}{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]} \\
&= \frac{\sum_{s\in\{1,2\}}\left\{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \mathsf{X}_4=1, \mathsf{S}=s\right]\mathbb{P}\left[\mathsf{X}_4=1 \,|\, \cancel{\mathsf{S}=s}, \cancel{\boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}}\right]\mathbb{P}\left[\mathsf{S}=s \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]\right\}}{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]} \\
&= \frac{\mathbb{P}\left[\mathsf{X}_4=1\right]\sum_{s\in\{1,2\}}\left\{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \mathsf{X}_4=1, \mathsf{S}=s\right]\mathbb{P}\left[\mathsf{S}=s \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]\right\}}{\mathbb{P}\left[\mathsf{C}_4=1 \,|\, \boldsymbol{\mathsf{C}}_{1:3}=\boldsymbol{c}_{1:3}\right]} \\
&= \frac{\alpha\left\{(1-\gamma)\epsilon+\gamma(1-\epsilon)\right\}}{\zeta} = 0.327
\end{aligned}
$$