Nguyen Xuan Binh 887799

Exercise 6A3

a) We know that only 1.5% of the string consists of protein coding regions

$\Rightarrow P_\Theta(\theta = 1) = 0.015$

$\Rightarrow P_\Theta(\theta = 0) = 0.985$

So any random location i has 1.5% chance in the protein coding regions

b) The prior distribution table

$f(\theta)$	$\theta = 1$	$\theta = 0$
A	0.17	0.25
C	0.29	0.25
G	0.33	0.25
T	0.21	0.25

The string has 16 As, 26 Cs, 38Gs and 20 Ts

Since $\theta = 0$

$\Rightarrow P(x_1 = A, \dots, x_{100} = A)$

$= 0.25^{16} \times 0.25^{26} \times 0.25^{38} \times 0.25^{20}$

$= 0.25^{100}$

$= 0.6223 \times 10^{-60}$

c) If $\theta = 1$

$\Rightarrow P(x_1 = A, \dots, x_{100} = A) = 0.17^{16} \times 0.29^{26} \times 0.33^{38} \times 0.21^{20}$

$\cancel{= 0.3443 \times 10^{-48}}$

$= 7.2019 . 10^{-59}$

d) Posterior distribution of Θ

$\theta = 1 \Rightarrow f_\Theta(1) . f(\text{sequence} \mid 1) = 0.015 \times 7.2019 . 10^{-59}$

$= 1.0803 \times 10^{-60}$

$\theta = 0 \Rightarrow f_\Theta(2) . f(\text{sequence} \mid 0) = 0.985 \times 0.6223 . 10^{-60}$

$= 6.1297 \times 10^{-61}$

$\Rightarrow c = 1.0803 \times 10^{-60} + 6.1297 \times 10^{-61} = 1.6933 . 10^{-60}$

$\Rightarrow$ Normalized distribution of Θ

θ	1	0
$f_\Theta(\theta) f(\text{sequence} \mid \theta)$	0.638	0.362

Exercise 6A4:

a) If $\theta = 0.5$, then it should rapidly changes its face on its flip. However, there are a long sequence of continuous 0 and 1, thus Θ cannot be 0.5

b) θ is probability of flipping the coin. 1 indicates flipping while 0 indicates no flip

$$\Rightarrow f_{X|\Theta}(x, \theta) = \begin{cases} 0, & \text{no flip} \\ 1, & \text{flipping} \end{cases}$$

There are 50 rounds and 9 flips $\Rightarrow$ $50 - 9 = 41$ no flips

$\Rightarrow$ $P(9 \text{ flip}, 41 \text{ no flips}) = \theta^9(1-\theta)^{41}$ Since prior is uniform $\Rightarrow f_\Theta(\theta) = 1$

$\Rightarrow$ Posterior distribution: $1 \times \theta^9(1-\theta)^{41} = \theta^9(1-\theta)^{41}$

PDF of Beta distribution: $f(\theta \mid a, b) = c \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1}$

Since it's not necessary to normalize, $\Rightarrow \alpha - 1 = 9 \Rightarrow \alpha = 10$

$\beta - 1 = 41 \Rightarrow \beta = 42$

$\Rightarrow \theta^9(1-\theta)^{41} = \text{Beta}(10, 42)$

$\Rightarrow$ Posterior mean: $\frac{\alpha}{\alpha+\beta} = \frac{10}{10+42} = \frac{5}{26} \approx 0.1923$

We have: $\frac{d}{d\theta}(\theta^9(1-\theta)^{41}) = 9\theta^8(1-\theta)^{41} - 41\theta^9(1-\theta)^{40} = 0$

$= -\theta^8(\theta-1)^{40}(50\theta - 9) = 0$

$\Rightarrow \theta = 0, 1, \frac{9}{50} \Rightarrow$ Posterior mode is $\frac{9}{50} \approx 0.18$

c) I use matlab and the interval it calculated is $[0.0982, 0.3087]$

d) Professor Abel is probably wrong because we know the mean and mode of θ to be 0.1923 and 0.18, because only at $\theta = [0.18, 0.2]$ would it be possible for very long sequence of non-flips $\Rightarrow \theta = 50\%$ is highly unlikely to be the true θ

e) We observe that the 10^{th} shake only has the same face as the initial position if it has to flip even times. For example if it flips 3 times after 10 shakes, the face would be different. There are 10 shakes so there's a maximum of 10 flips

$\Rightarrow$ In order to have the same face as the initial face, $\Rightarrow$ the flip is 0, 2, 4, 6, 8, 10

We have $\theta = 0.2$

Flip = 0 $\Rightarrow P(F=0) = C_0^{10}\, 0.2^0 \times 0.8^{10} = 0.1074$

Flip = 2 $\Rightarrow P(F=2) = C_2^{10}\, 0.2^2 \times 0.8^8 = 0.302$

Flip = 4 $\Rightarrow P(F=4) = C_4^{10}\, 0.2^4 \times 0.8^6 = 0.088$

Flip = 6 $\Rightarrow P(F=6) = C_6^{10}\, 0.2^6 \times 0.8^4 = 0.0055$

Flip = 8 $\Rightarrow P(F=8) = C_8^{10}\, 0.2^8 \times 0.8^2 = 0.0000737$

Flip = 10 $\Rightarrow P(F=10) = C_{10}^{10}\, 0.2^{10}\, 0.8^0 = 0.2^{10}$

$\Rightarrow$ Probability of having same face after 10 shakes $= \sum P(F = 0 \cdots 10)$

$= 0.50297$