PAGE NO.:
DATE: / /

Q1. Two coins y_1 & y_2

$P(H|y_1) = 0.6$ ...(given)

$\therefore P(T|y_1) = 0.4$

Similarly $P(H|y_2) = 0.1$ ...(given)

$\therefore P(T|y_2) = 0.9$

Also probability of choosing one of the 2 coins at random $P(y_1) = P(y_2) = 0.5$

(a) To find $P(2H\,1T)$

$P(2H\,1T) = P(2H\,1T|y_1) \cdot P(y_1) + P(2H\,1T|y_2) \cdot P(y_2)$

Now $P(2H1T|y_1) = P(HHT|y_1) + P(THH|y_1) + P(HTH|y_1)$

$= (0.6)^2 \times 0.4 + (0.6)^2 \times 0.4 + (0.6)^2 \times 0.4$

$= 0.432$

& $P(2H\,1T|y_2) = P(HHT|y_2) + P(THH|y_2) + P(HTH|y_2)$

$= \cancel{(0.6)^2 \times 0.4 + (0.6)^2 \times 0.4 + (0.6)^2 \times 0.4}$

$= (0.1)^2 \times 0.9 + (0.1)^2 \times 0.9 + (0.1)^2 \times 0.9$

$= 0.027$

$\therefore P(2H\,1T) = P(2H\,1T|y_1)P(y_1) + P(2H\,1T|y_2)P(y_2)$

$= 0.432 \times 0.5 + 0.027 \times 0.5$

$= 0.2295$

(b) To find $P(y_1/2H\,1T)$

By bayes theoram,

$$P(y_1|2H\,1T) = \frac{P(2H\,1T|y_1) \cdot P(y_1)}{P(2H\,1T)}$$

PAGE NO.:
DATE: / /

$\therefore P(y_1|2H1T) = \frac{0.432 \times 0.5}{0.2295}$ ...... (using values from ⓐ)

$= \underline{0.941}$

ⓒ $P(H|2H1T) = ?$

Now $P(H|2H1T) = P(y_1|2H1T) \cdot P(y_1 = H) +$

$P(y_2|2H1T) \cdot P(y_2 = H)$ — ①

From ⓑ we have $P(y_1|2H1T) = 0.941$ — ②

$\therefore P(y_2|2H1T) = 1 - P(y_1|2H1T)$

$= 1 - 0.941$

$= 0.058$ — ③

$\therefore$ from ①, ② & ③ $P(H|2H1T) = 0.941 \times 0.6 + 0.058 \times 0.1$

$= 0.564 + 0.0058$

$= 0.57$

Q2. Given p.d.f $f(x) = \frac{\theta^x e^{-\theta}}{x!}$

$\therefore$ Likelihood $= \prod_{i=1}^{n} f(x)$

$= \prod_{i=1}^{n} \frac{\theta^x e^{(-\theta)}}{x!}$

$= \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}{x_1 x_2 x_3 \cdots x_n}$

PAGE NO.:
DATE: / /

Now taking log likelihood.

$$L(x) = \sum_{i=1}^{n} x_i \log\theta - n\theta - \sum_{i=1}^{n} \log(x_i!)$$

Here $n = 20$ ....(given)

$$\therefore L(x) = \sum_{i=1}^{20} x_i \log\theta - 20\theta - \sum_{i=1}^{20} \log(x_i!)$$

Now to find maximum likelihood for θ

$$\frac{d}{d\theta} L(x) = 0$$

$$\therefore \sum_{i=1}^{20} x_i \times \frac{1}{\theta} - 20 = 0$$

$$\therefore \theta \text{ or } \hat{\theta} = \frac{1}{20} \sum_{i=1}^{20} x_i$$

$$= \bar{x} \text{ (can be said as mean value)}$$

$$\therefore \boxed{\hat{\theta} = \bar{x} = 6.9} \text{ ...(given) } \bar{x} = 6.9$$

(b) To prove: Posterior is gamma if prior is gamma.

→ Given: $P(\theta / a_0, b_0)$ → gamma prior

$P(y|\theta)$ → poison likelihood

To prove: posterior $P(\theta | y, a_0, b_0) \sim \Gamma(a_n, b_n)$

→ We know that posterior $\propto$ likelihood × prior

$$\therefore P(\theta | y, a_0, b_0) \propto P(y|\theta) \cdot P(\theta | a_0, b_0)$$

$$\therefore p(\theta / y, a_0, b_0) \propto \frac{\theta^{\Sigma y_i} e^{-\theta n}}{y!} \times \frac{b_0^{a_0}}{\Gamma(b_0)} \theta^{a_0-1} e^{-\theta b_0}$$

We may neglect terms without the θ.

$$p(\theta / y, a_0, b_0) \propto \theta^{\Sigma y_i + a_0 - 1} e^{-\theta(b_0+n)}$$

Thus posterior is a gamma distribution by comparing above expression with gamma p.d.f expression.

Thus $\underline{\underline{a_n = \Sigma y_i + a_0}}$ & $\underline{\underline{b_n = b_0 + n}}$

(c) Given average = 5 visits per minute
and ~~Gamma~~ Prior chosen is Gamma $(2.5, 0.5)$
and sample mean = 6.9.

$a_0 = 2.5$ & $b_0 = 0.5$

we know that $E[\theta] = \frac{a_0}{b_0} = \frac{2.5}{0.5} = 5$

$$var[\theta] = \frac{a_0}{b_0^2} = \frac{2.5}{0.25} = 10$$

Both variance & expected value match with the given average minutes per minute
Thus given prior is a good assumption.

(d) Given that prior is gamma$(2.5, 0.5)$

$$a_n = \Sigma x_i + a_0$$

$$a_n = \sum x_i + a_0$$

$$= n\bar{x} + a_0$$

$$= 20 \times 6.9 + 2.5$$

$$= 140.5$$

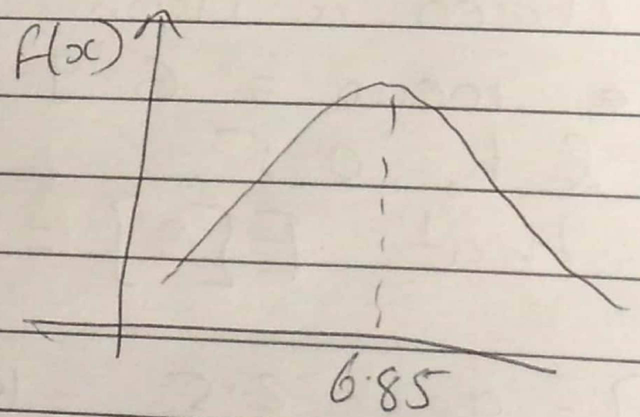
$$b_n = n + b_0$$

$$= 20 + 0.5$$

$$= 20.5$$

$$E(x) = \frac{a_n}{b_n} = 6.85$$

$$Var(x) = \frac{a_n}{b_n^2} = 0.334$$



```
markov1 <- function(rr = 1){

  y0 <- c(1, 0, 0, 0)

  trans_mat <- matrix(0, nrow = 4, ncol = 4)

  trans_mat[1, ] <-   c(0.721, 0.202, 0.067, 0.010)
  trans_mat[2, 2:4] <- c(0.581, 0.407, 0.012)
  trans_mat[3, 3:4] <- c(0.750, 0.250)
  trans_mat[4, 4] <- 1

  trans_mat[1, 2:4] <- trans_mat[1, 2:4] * rr
  trans_mat[1, 1] <- 1 - sum(trans_mat[1, 2:4])

  trans_mat[2, 3:4] <- trans_mat[2, 3:4] * rr
  trans_mat[2, 2] <- 1 - sum(trans_mat[2, 3:4])

  trans_mat[3, 4] <- trans_mat[3, 4] * rr
  trans_mat[3, 3] <- 1 - trans_mat[3, 4]

  y_store <- matrix(0, nrow = 20, ncol = 4)
  y_store[1, ] <-  y0 %*% trans_mat

  for(i in 2:20){
    y_store[i, ] <-  y_store[i-1, ] %*% trans_mat
  }
  sum(y_store[, 1:3])
}
```

Above function is the markov function to multiply the state matrix with the transition matrix and finally calculate the sum of resultant matrix.

3a. The best estimate for the relative risk r is ˆr = 0.509. Using this value, estimate the gain in life years for patients taking combination therapy versus monotherapy.

- ➔ For this we use the above defined function and calculate markov1(0.509)-markov1(1) which gives the result as 5.885. Thus the gain the life years using combination therapy is 5.88 as compared to monotherapy.

3b. Because it was estimated from a trial containing a limited number of patients, there is considerable uncertainty regarding the estimate of the relative risk r. Briggs et al propose to model this uncertainty by modelling r as a log normal distribution, log r ∼ N(−0.675, 0.16972). Using samples generated from this distribution, compute the expected number of life years gained by patients taking combination therapy versus monotherapy.

- ➔ For this we execute the below piece of code with the above function

```
rnorm <- rlnorm(1000,meanlog=-0.675,sdlog=0.1697*0.1697)
flag=0
rnorm_gain1=c()
for(i in rnorm){
  rnorm_gain1[flag]=markov1(i)-markov1(1)
```

```
  flag=flag+1
}
mean(rnorm_gain1)
```

So after calculating the mean of rnorm_gain, we get the output as **5.883 as the gain in life (years) when r is modelled as a log normal distribution.**

3c. Using the samples generated in b), compute lower and upper bounds yα and yβ such that P(yα < y < yβ) = 0.95, where y denotes the amount of life years gained by patients taking combination therapy versus monotherapy.

➔ For this we use the quantile function

```
quantile(rnorm_gain1,c(0.025,0.975))
```

output:

```
> quantile(rnorm_gain1,c(0.025,0.975))
    2.5%    97.5%
5.434911 6.308285
```

Thus the lower and upper bounds are 5.4 and 6.31 respectively