HMM

Problem

Continuously cast normal dice or fake dice a total of T times. HMM is used to predict which dice were used from the observed data. Let $Y \in \{1, \ldots, 6\}$ be an observation variable (a roll of the dice) and

$$X = \begin{cases} 1, & \text{if fake dice is used} \\ 0, & \text{if normal dice is used} \end{cases}$$

be a hidden variable.

The hidden variable transition model is

$P(X_{t+1} \mid X_t)$	$X_{t+1} = 0$	$X_{t+1} = 1$
$X_t = 0$	0.9	0.1
$X_t = 1$	0.1	0.9

The observation process model is

$P(Y_t \mid X_t)$	$Y_t = 1$	$Y_t = 2$	$Y_t = 3$	$Y_t = 4$	$Y_t = 5$	$Y_t = 6$
$X_t = 0$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$X_t = 1$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{2}$

The observation data is
4,5,4,2,3,3,6,4,5,5,3,4,4,1,4,5,3,6,5,3,3,3,5,5,3,5,6,5,5,1,3,4,3,1,2,6,1,6,1,5,4,2,4,1,5,4,1,1,1,1
,5,6,6,6,6,1,6,2,6,2,6,1,6,6,6,6,6,3,2,6,6,6,1,6,6,2,6,6,5,6,6,5,6,6,6,6,4,3,6,6,5,2,5,4,5,6,5,4,4.

Algorithm

The posterior probability of hidden variables $\gamma(X_t)$ is calculated using the Forward-Backward algorithm. The Forward-Backward algorithm is

$$\begin{aligned} \alpha(X_t) &\equiv P(Y_1, \ldots, Y_t, X_t) \\ &= P(Y_t \mid X_t) \sum_{t_{t-1}} \alpha(X_{t-1}) P(X_t \mid X_{t-1}) \end{aligned}$$

$$\beta(X_t) \equiv P(Y_{t+1}, \ldots, Y_T | X_t)$$
$$= \sum_{X_{t+1}} \beta(X_{t+1}) P(Y_{t+1}|X_{t+1}) P(X_{t+1}|X_t)$$
$$\gamma(X_t) \equiv P(X_t|Y_1, \ldots, Y_T) \propto \alpha(X_t)\beta(X_t)$$

The initial values are

$$\alpha(X_0 = 0) = P(Y_0|X_0 = 0) = \frac{1}{6}$$
$$\alpha(X_0 = 1) = P(Y_0|X_0 = 1) = 0.1$$
$$\beta(X_T = 0) = P(Y_T|X_T = 0) = \alpha(X_T = 0)$$
$$\beta(X_T = 1) = P(Y_T|X_T = 1) = \alpha(X_T = 1)$$

The posterior probability of fake dice is

$$\frac{\gamma(X_t = 1)}{\gamma(X_t = 0) + \gamma(X_t = 1)}$$

Result