

# TIME SERIES (20236) Assignment 2

Group 15: Alessandro Caggia, Edoardo Calleri di Sala, Emanuele Cinti, Stefano De Bianchi

## 1 Basic notions on Markov chains

**Question 1:** A discrete-time stochastic process  $(Y_t)_{t \geq 0}$  is said to be a Markov chain if the conditional distribution of  $Y_t$  depends only on the most recent past observation  $Y_{t-1}$ . This means that  $Y_t$  is independent of all previous states  $(Y_0, \dots, Y_{t-2})$ , given  $Y_{t-1}$ . Thus, for a Markov chain, the probability distribution of  $Y_t$  given  $(Y_0 = y_0, \dots, Y_{t-1} = y_{t-1})$  is

$$p(y_t | y_0, y_1, \dots, y_{t-1}) = p(y_t | y_{t-1}),$$

for any  $t \geq 1$ .

For example, let  $Y_t$  represent an individual's vote intention at time  $t$ , with values: 1 = Republican, 2 = Democrat, 3 = Don't know. Suppose that the transition matrix is:

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0 & 0.4 & 0.6 \end{bmatrix}$$

If the person is undecided at time  $t = 0$ , i.e.,  $Y_0 = 3$ , then:  $P(Y_1 = 2 | Y_0 = 3) = 0.4$ . Since each  $Y_t$  depends only on  $Y_{t-1}$  and not on earlier values, this process satisfies the Markov property. Therefore,  $(Y_t)_{t \geq 0}$  is a homogeneous Markov chain.

**Question 2:** The *initial distribution*  $p_0(\cdot)$  is the probability distribution of  $Y_0$ , i.e., where the process starts. The *transition matrix*  $P$  of a discrete-time Markov chain with finite state space  $\mathcal{Y} = \{1, 2, \dots, N\}$  is the  $N \times N$  matrix whose  $(i, j)$ -th entry is the one-step transition probability

$$p_{i,j} = P(Y_t = j | Y_{t-1} = i),$$

i.e., the probability of moving from state  $i$  to state  $j$  in one time step.

If  $P$  does not depend on  $t$ , the Markov chain is said to be *homogeneous*, and the same transition matrix  $P$  applies at every time step. Meaning that the probabilities of transition from a state  $i$  to a state  $j$  are constant over time. In case of non-homogeneous Markov chains the matrix would change over time, meaning that the initial distribution of  $Y_0$  is going to be possibly different from the one at time  $t$ ,  $Y_t$ .

**Question 3:** DAG is a method for showing the dependence structure of a Markov chain. In a Markov chain, the dependence structure is:

$$Y_0 \rightarrow Y_1 \rightarrow Y_2 \rightarrow \dots \rightarrow Y_{t-3} \rightarrow Y_{t-2} \rightarrow Y_{t-1} \rightarrow Y_t \rightarrow \dots$$

This graph allows us to see how each variable  $Y_s$  only depends directly on  $Y_{s-1}$ . According to the Markov property, for any  $t$ , we have:

$$P(Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_0) = P(Y_t | Y_{t-1})$$

Thus, given  $Y_{t-2}$ , the past variables  $(Y_1, \dots, Y_{t-3})$  influence  $Y_t$  only through  $Y_{t-2} \rightarrow Y_{t-1} \rightarrow Y_t$ . So by the D-separation rule in DAGs, once we condition on  $Y_{t-2}$ , all earlier variables become irrelevant for predicting  $Y_t$ .

$$\Rightarrow Y_t \perp (Y_1, \dots, Y_{t-3}) | Y_{t-2}$$

**Question 4:**

(a) We consider three possible paths:

1)  $Y_0 = 1 \rightarrow Y_1 = 2 \rightarrow Y_2 = 2$ , which has probability:

$$P(Y_1 = 2 | Y_0 = 1) \cdot P(Y_2 = 2 | Y_1 = 2) = 0.4 \times 0.7 = 0.28.$$

2)  $Y_0 = 1 \rightarrow Y_1 = 1 \rightarrow Y_2 = 2$ , which has probability:

$$P(Y_1 = 1 | Y_0 = 1) \cdot P(Y_2 = 2 | Y_1 = 1) = 0.6 \times 0.4 = 0.24.$$

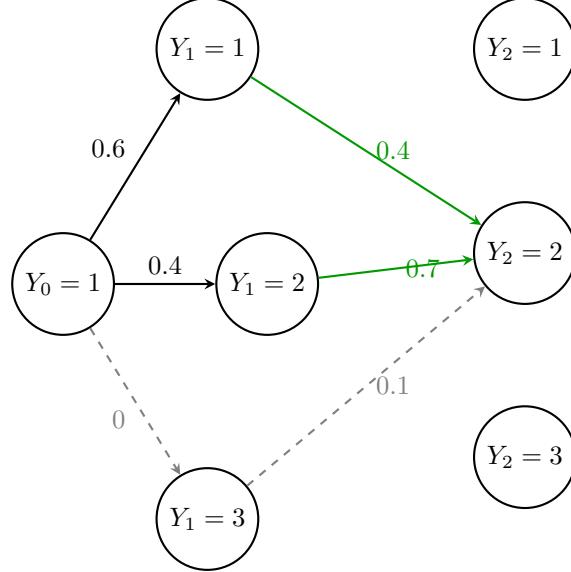
3)  $Y_0 = 1 \rightarrow Y_1 = 3 \rightarrow Y_2 = 2$ , which has probability:

$$P(Y_1 = 3 | Y_0 = 1) \cdot P(Y_2 = 2 | Y_1 = 3) = 0 \times P(Y_2 = 2 | Y_1 = 3) = 0.$$

Thus, the total probability that  $Y_2 = 2$  given  $Y_0 = 1$  is:

$$P(Y_2 = 2 | Y_0 = 1) = 0.28 + 0.24 + 0 = 0.52.$$

(b) Knowing that  $Y_0 = 1$  and  $Y_2 = 2$ , the probability that  $Y_1 = 1$  is given by the ratio:  $\frac{0.24}{0.52} = 0.4615$ .



### Question 5:

(a) The state space of the Markov chain is

$$\mathcal{Y} = \{1, 2, 3\}.$$

(b) To compute the marginal distribution of  $Y_2$  given  $Y_0 = 1$ , we use the transition matrix

$$P = \begin{bmatrix} 0.6 & 0.4 & 0.0 \\ 0.3 & 0.6 & 0.1 \\ 0.0 & 0.2 & 0.8 \end{bmatrix}.$$

We want the distribution of  $Y_2$  given  $Y_0 = 1$ , which is the first row of  $P^2$ . Compute:

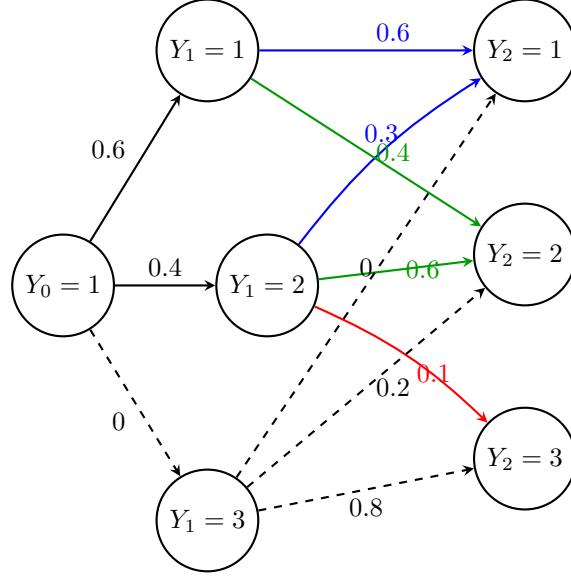
$$P_{1,1}^2 = 0.6 \cdot 0.6 + 0.4 \cdot 0.3 + 0 \cdot 0 = 0.36 + 0.12 = 0.48$$

$$P_{1,2}^2 = 0.6 \cdot 0.4 + 0.4 \cdot 0.6 + 0 \cdot 0.2 = 0.24 + 0.24 = 0.48$$

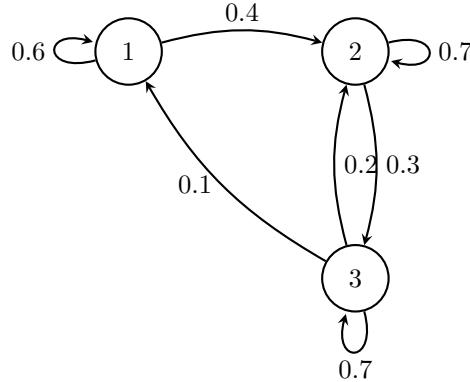
$$P_{1,3}^2 = 0.6 \cdot 0 + 0.4 \cdot 0.1 + 0 \cdot 0.8 = 0.04$$

Therefore, the marginal distribution of  $Y_2$  given  $Y_0 = 1$  is:

$$P(Y_2 = 1 | Y_0 = 1) = 0.48, \quad P(Y_2 = 2 | Y_0 = 1) = 0.48, \quad P(Y_2 = 3 | Y_0 = 1) = 0.04.$$



**Question 6:** A matrix is irreducible if there are no closed sets. We must check if there is the possibility to get stuck in a given combination of states with no possibility to go back to the other states. A quick analysis of the table allows us to see that there are no closed sets: from any state you can always reach any other state. We have tested multiple paths, and we compacted our analysis in the chart below



**Question 7: Limit Behavior of  $P(Y_n = k | Y_0 = j)$  and  $P(Y_n = k)$**

Let  $(Y_t)_{t \geq 0}$  be a Markov chain with finite state space  $\mathcal{Y} = \{1, \dots, N\}$ , transition matrix  $P$ , and initial state  $Y_0 = j$ . We want to study the asymptotic behavior of such Markov chain. In particular, we want to know the probability distribution of  $Y_n$  given the initial state  $Y_0 = j$ . In such a context we will have:

- The conditional distribution:

$$P(Y_n = k | Y_0 = j),$$

- The marginal distribution:

$$P(Y_n = k) = \sum_{j=1}^N P(Y_n = k | Y_0 = j) \cdot P(Y_0 = j).$$

As often we do not deal with stationary series, our  $Y_n$  will vary with  $n$ . However, we want to know whether asymptotically, i.e. for large  $n$ , the series tends towards a stage of equilibrium. We want to know whether the conditional distribution of  $Y_n$ , given the initial state  $j$ , converges to some limit distribution, and, if so, if such limiting distribution is reached for any initial state. In this case, the conditional and the marginal distribution would converge to the same limit distribution  $p^*(k)$ . Under the conditions that the Markov chain is *irreducible* and *aperiodic* and with a finite state space, we have:

1. The conditional distribution converges to a limit distribution  $p^*(k)$ , independent of  $j$ :

$$\lim_{n \rightarrow \infty} P(Y_n = k \mid Y_0 = j) = p^*(k), \quad \forall j, k,$$

where  $p^*(k) > 0$  and  $\sum_{k=1}^N p^*(k) = 1$ .

2. The limit distribution  $p^* = (p^*(1), \dots, p^*(N))$  satisfies the stationarity condition:

$$p^*(k) = \sum_{i \in \mathcal{Y}} p^*(i) \cdot p_{i,k}, \quad \text{for all } k \in \mathcal{Y},$$

meaning that the LR probability of staying in state  $k$  results from the LR probability of staying in  $i$  times the probability of jumping from  $i$  to  $k$ .

Then:

- The chain is called *ergodic* if the limit distribution  $p^*(k)$  exists and is independent of  $j$ .
- A distribution  $p(\cdot)$  satisfying  $p(k) = \sum_i p(i) \cdot p_{i,k}$  is called *stationary* or *invariant*.
- If  $Y_0 \sim p(\cdot)$ , then  $Y_n \sim p(\cdot)$  for all  $n \geq 1$ .

Below you can find a Monte Carlo simulation that we run. In short, we simulate 10,000 independent Markov chains over  $T = 500$  steps with 20 states and a transition matrix  $P$  featuring local, sticky transitions (e.g.,  $P[i, i] = 0.6$ ,  $P[i, i \pm 1] = P[i, i \pm 2] = 0.1$ ). All chains start at state 0; no wrap-around is allowed. We track  $P(Y_t = k)$  over time and compare it to the stationary distribution  $\pi$  (left eigenvector of  $P$  for eigenvalue 1). Initially, distributions are skewed; as  $t \rightarrow \infty$ , they converge to  $\pi$ , confirming ergodicity. The plot shows convergence and loss of memory of initial conditions.

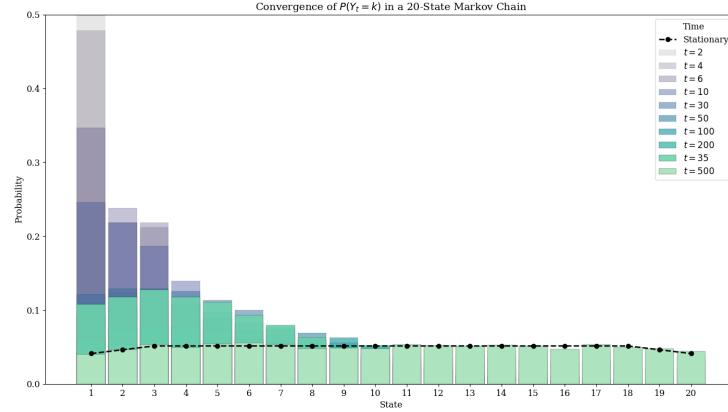


Figure 1: Monte Carlo simulation of 10,000 Markov Chains over 500 steps

## 2 Inference for Markov chains

### 2.1 Exercise 1

(a) **Maximum Likelihood Estimate (MLE) of  $p_{3,1} = P(Y_t = 1 | Y_{t-1} = 3)$ :**

We sum all transitions from state 3 to 1 across all time periods:

$$\begin{bmatrix} 105 & 10 & 15 \\ 40 & 55 & 25 \\ 25 & 55 & 70 \end{bmatrix}$$

Numerator ( $3 \rightarrow 1$ ) :  $10 + 10 + 0 + 5 = 25$ , Denominator (all from 3) :  $55 + 40 + 25 + 30 = 150$

$$\hat{p}_{3,1} = \frac{25}{150} = \frac{1}{6} \approx 0.1667$$

(b) **95% Asymptotic Confidence Interval for  $p_{3,1}$ :**

Our transition probabilities follow a multinomial distribution, hence we use the usual formula for the variance of the multinomial distribution:

$$SE = \sqrt{\frac{\hat{p}_{3,1}(1 - \hat{p}_{3,1})}{n}} = \sqrt{\frac{(1/6)(5/6)}{150}} \approx 0.0304$$

Then, as  $n$  is large enough (150) we can use the CLT and approximate the sampling distribution of  $\hat{p}_{3,1}$  with a normal distribution and compute CI.

$$\begin{aligned} & \left[ 0.1667 - 1.96\sqrt{\frac{0.1667(1 - 0.1667)}{150}}, \quad 0.1667 + 1.96\sqrt{\frac{0.1667(1 - 0.1667)}{150}} \right] \\ &= [0.1077; 0.2263] \end{aligned}$$

(c) **MLE of  $p_{3,1}(t)$  without the homogeneity assumption:**

We compute separate estimates for each time period.

**At  $t = 1$**  (transitions from  $t = 0$  to  $t = 1$ ):

$$\hat{p}_{3,1}(1) = \frac{10}{10 + 15 + 30} = \frac{10}{55} \approx 0.1818$$

**At  $t = 2$**  (transitions from  $t = 1$  to  $t = 2$ ):

$$\hat{p}_{3,1}(2) = \frac{10}{10 + 10 + 20} = \frac{10}{40} = 0.25$$

### 2.2 Exercise 2

(a) **Joint probability expression**

For individual 1, the observed path is:

$$(Y_{1,0}, Y_{1,1}, Y_{1,2}, Y_{1,3}, Y_{1,4}) = (1, 1, 3, 3, 2)$$

By the Markov property and homogeneity:

$$\begin{aligned} P(Y_{1,1} = 1, Y_{1,2} = 3, Y_{1,3} = 3, Y_{1,4} = 2 | Y_{1,0} = 1; \mathbf{P}) &= p_{1,1} \cdot p_{1,3} \cdot p_{3,3} \cdot p_{3,2} \\ &= 0.5 \cdot 0.33 \cdot 0.125 \cdot 0.75 \approx 0.015624 \end{aligned}$$

(b)

$$\begin{aligned}
L(p_{i,j}; Y_1, Y_2, Y_3, Y_{i,0}) &= \prod_{k=1}^{100} P(Y_k; p_{i,j}) \\
&= \prod_{k=1}^{100} P(Y_{k,0}, Y_{k,1}, \dots, Y_{k,T-1}) \\
&= \prod_{i=1}^{100} \prod_{j=1}^{100} p_{i,j}^{n_{i,j}}
\end{aligned}$$

Maximizing the likelihood, we obtain:

$$\hat{p}_{i,j} = \frac{n_{i,j}}{n_{i+}}$$

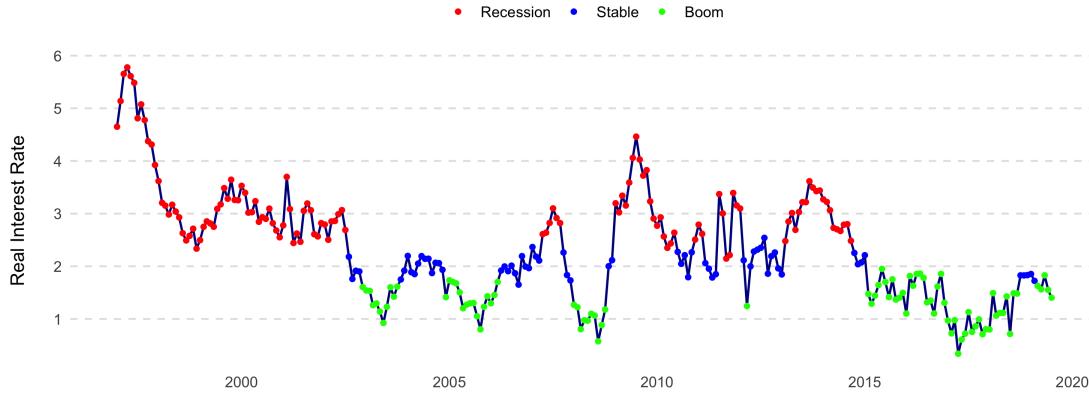
(c) The estimated probability that an individual who is in favour in May will turn negative in June is:  $\hat{p}_{j,t} = \frac{20}{120} = 0.767$

$$\left[ 0.767 - 1.65\sqrt{\frac{0.767(1-0.767)}{120}}, 0.767 + 1.65\sqrt{\frac{0.767(1-0.767)}{120}} \right] = [0.711, 0.823]$$

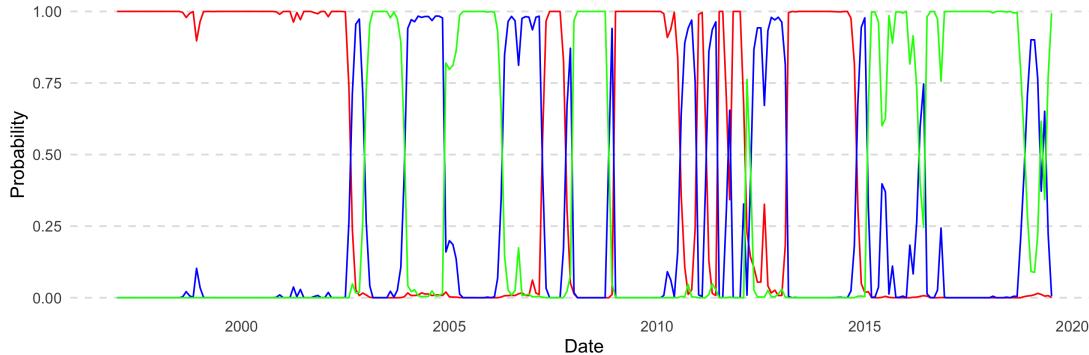
### 3 Hidden Markov Models with R

An HMM model may be reasonable when dealing with this time series as one may expect the bond market - and the stock market more in general - to be characterized by different "states" as for example phases of expansions, contraction and stability. Below we plot in panel (a) the time series and the identified states and in panel, in panel (b) the dynamic evolution of the HMM state probabilities.

Panel (a): Real Interest Rate and HMM States (Homogeneous)



Panel (b): HMM derived state probabilities



Below we report the MLEs of the unknown parameters of the model (and their standard errors).

State	Mean (Intercept)	Std. Dev. ( $\sigma$ )
State 1 (Boom)	3.169	0.737
State 2 (Stable)	2.019	0.207
State 3 (Recession)	1.290	0.351

Table 1: MLEs of state-dependent means and standard deviations from the homogeneous HMM.

The homogeneous HMM identifies three regimes of the real interest rate:

- State 1 (Boom): Highest mean (3.169) and relatively high volatility (0.737), suggesting high and volatile interest rates.
- State 2 (Stable): Intermediate mean (2.019) with low volatility (0.207).
- State 3 (Recession): Lowest mean (1.290) and moderate variability (0.351), capturing contractionary phases characterized by lower rates.

The decoded plot, displayed in panel (a) above, shows the regime sequence assigned to each point. Plus, below we plot a comparison between the state-dependent mean assigned to each time point, based on the decoded state.

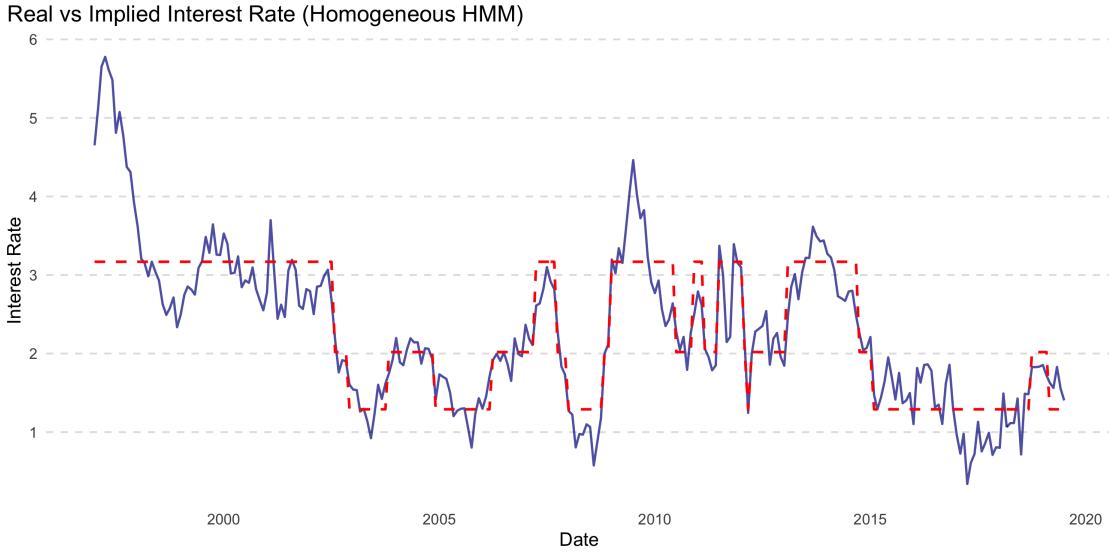


Figure 2: A comparison between State dependent means and raw data

Then, we allow the transition matrix to depend on covariates. In particular, we introduce inflation and gdp. Below we plot the newly decoded signals and the state-dependent means. In such a case, we can see how such model expansion does not significantly alter our results.

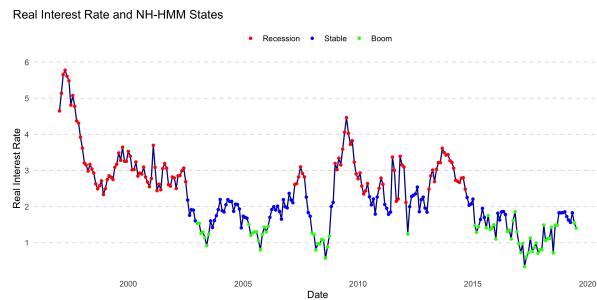


Figure 3: NH-HMM states

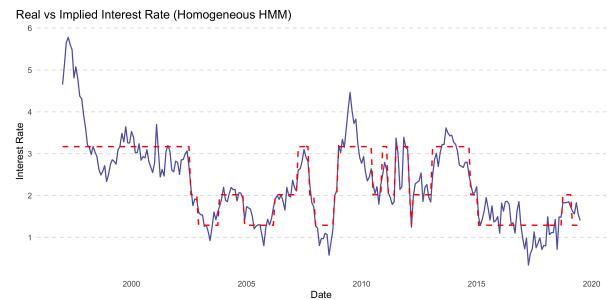


Figure 4: NH-HMM states implied mean