Question

A conversation involving a person A was observed at 5-second intervals. If A was speaking a 1 was recorded; if A was silent a 0 was recorded. The following results were obtained:

1111111110001111111111111111111111111111
1111111111111111111111111111111110000000
0000000000011111111111111111111111111111
1111100111111111111111110011111111111111
11111111110100000001111111

Describe the assumption which must be made in order to model this process as a two-state Markov chain.
Estimate the transition probability matrix.
If the person is observed talking estimate the probability that he will be taking.

(i) at the next observation of the conservation,

(ii) when the conservation is observed 500 seconds later.

Estimate the mean recurrence time for state 1 from the data and explain how this relates to the proportion of time spent in state 1.

Answer

The frequency matrix is as follows:

		To 0	To 1	Total
From	0	25	6	31
	1	6	161	167
				198

So an estimated transition matrix will be

$$P = \begin{pmatrix} \frac{25}{31} & \frac{6}{31} \\ \frac{6}{167} & \frac{161}{167} \end{pmatrix}$$

The assumptions are that the probabilities of transition depend only on the starting state, and not on past history.

(i) If $\mathbf{p}_0 = (0, 1)$ then at the next observation the distribution is

$$\mathbf{p}_0 P = \left(\frac{6}{167}, \frac{161}{167}\right)$$

so the probability that he is talking is $\dfrac{161}{167}$

(ii) This is a finite Markov chain so it has equilibrium distribution given by the stationary distribution. So $\pi P = \pi$

i.e. $\pi_0 \dfrac{25}{31} + \pi_1 \dfrac{6}{167} = \pi_0$ and $\pi_0 + \pi_1 = 1$

These give $\pi_0 = \frac{31}{198} \approx 0.157 \quad \pi_1 = \frac{167}{198} \approx 0.843$

OR use the standard formula for a 2×2 Markov chain to obtain:

$$P^n \to \begin{pmatrix} \frac{31}{198} & \frac{167}{198} \\ \frac{31}{198} & \frac{167}{198} \end{pmatrix}$$

From the data, recurrence times for state 1 are as follows

Time till return	1	2	3	4	8	17	Total
Frequency	161	1	2	1	1	1	167

The sample mean recurrence time is $\dfrac{198}{167} \approx 1.186$ The expected proportion of the time spent in state 1 should be π_1 for a long sequence, and should satisfy $\pi_1 = \frac{1}{\mu}$ where μ_1 is the mean recurrence time.

This is borne out by this example.