Lecture 12

Markov Sequences

Note: We will always assume a descrete or continuous density function

Defn: The sequence Y_n is Markov if

$$p(y_n \mid y_i \; i<n) = \underbrace{p(y_n \mid y_{n-1})}_{\text{transition density}}$$

- y_n only depends on the last sample

Example 1)

X_n are i.i.d. $N(0, \sigma^2)$

$$Y_n = \sum_{i=1}^{n} X_n \qquad n \geq 1$$

$$Y_n = Y_{n-1} + X_n$$

$$p(y_n \mid y_i \; i<n) = p(y_n \mid x_i \; i<n)$$

$$\sim N\left(\sum_{i=1}^{n-1} x_i, \sigma^2\right)$$

$$\sim N(y_{n-1}, \sigma^2)$$

$$= p(y_n \mid y_{n-1})$$

Distribution for Markov Chain

$$p(y) = p_n(y_n \mid y_i \; i<n)\, p(y_i \; i<n)$$

$$= p_n(y_n \mid y_{n-1})\, p(y_i \; i<n)$$

$$= p_n(y_n \mid y_{n-1})\, p_{n-1}(y_{n-1} \mid y_{n-2}) \cdots p_2(y_2 \mid y_1)\, p_1(y_1)$$

$$\boxed{p(y) = \left(\prod_{i=2}^{n} p_i(y_i \mid y_{i-1}) \right) p_1(y_1)}$$

Example) binary Y_i from before

$$p(y) = \left(\prod_{i=2}^{N+1} p_i(y_i \mid y_{i-1}) \right) p_1(y_1)$$

$$p(y_i \mid y_{i-1}) = \begin{cases} 1-p & y_i = y_{i-1} \\ p & y_i = 1 - y_{i-1} \end{cases}$$

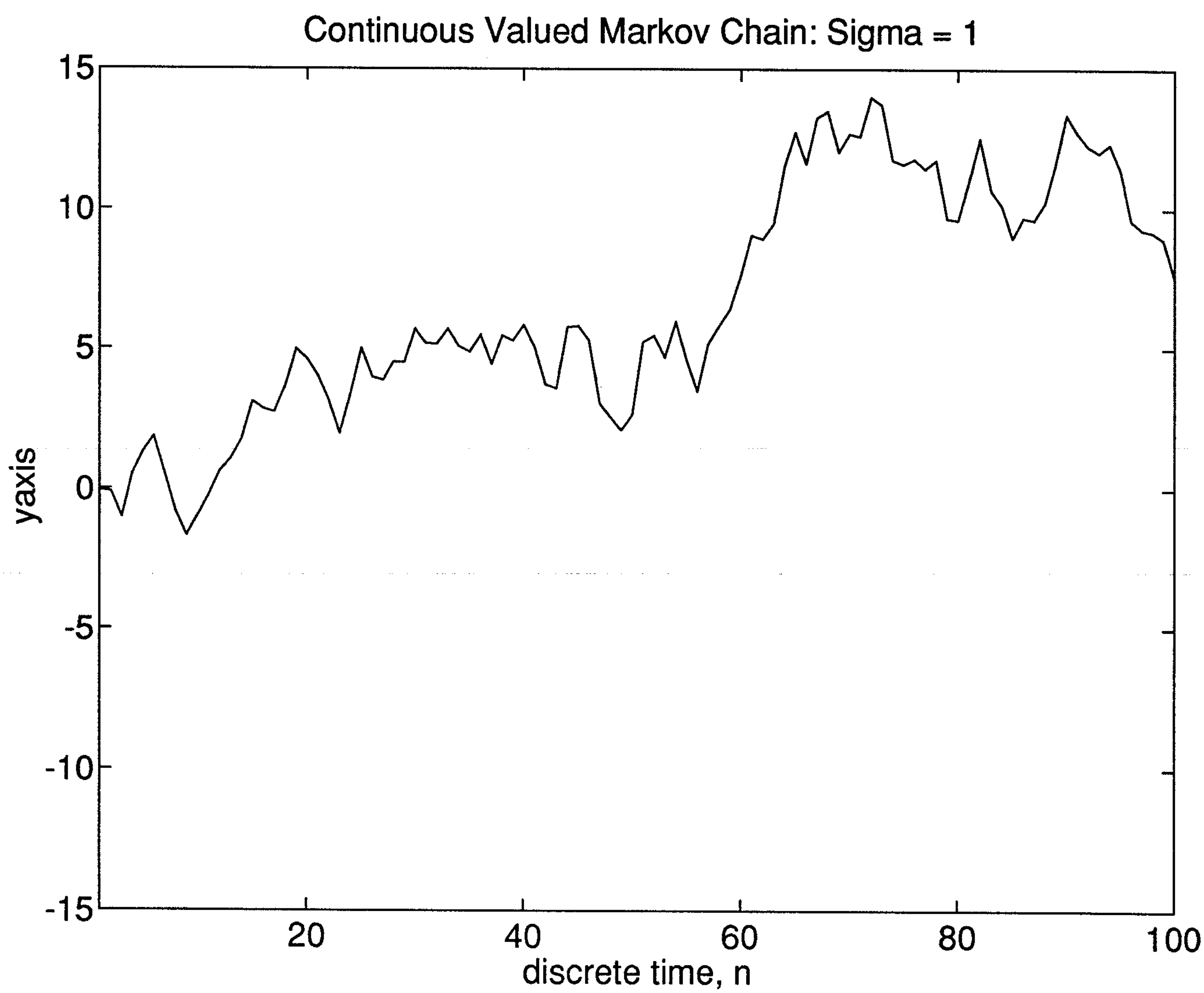
$$p(y_1) = 1/2$$

Define: $K \triangleq N - \sum_{i=2}^{N+1} \delta_{y_i - y_{i-1}}$

K = # of transitions

$N-K$ = # of no transition

$$p(y) = (1-p)^{N-K} p^K (1/2)$$


Continuous Valued Markov Chain: Sigma = 1
15
10
5
0
-5
-10
-15
yaxis
20
40
60
80
100
discrete time, n

Example 2)

Let Y_i and B_i be binary sequences of 0's and 1's

$$P\{B_i = 1\} = p$$
$$P\{B_i = 0\} = 1-p$$

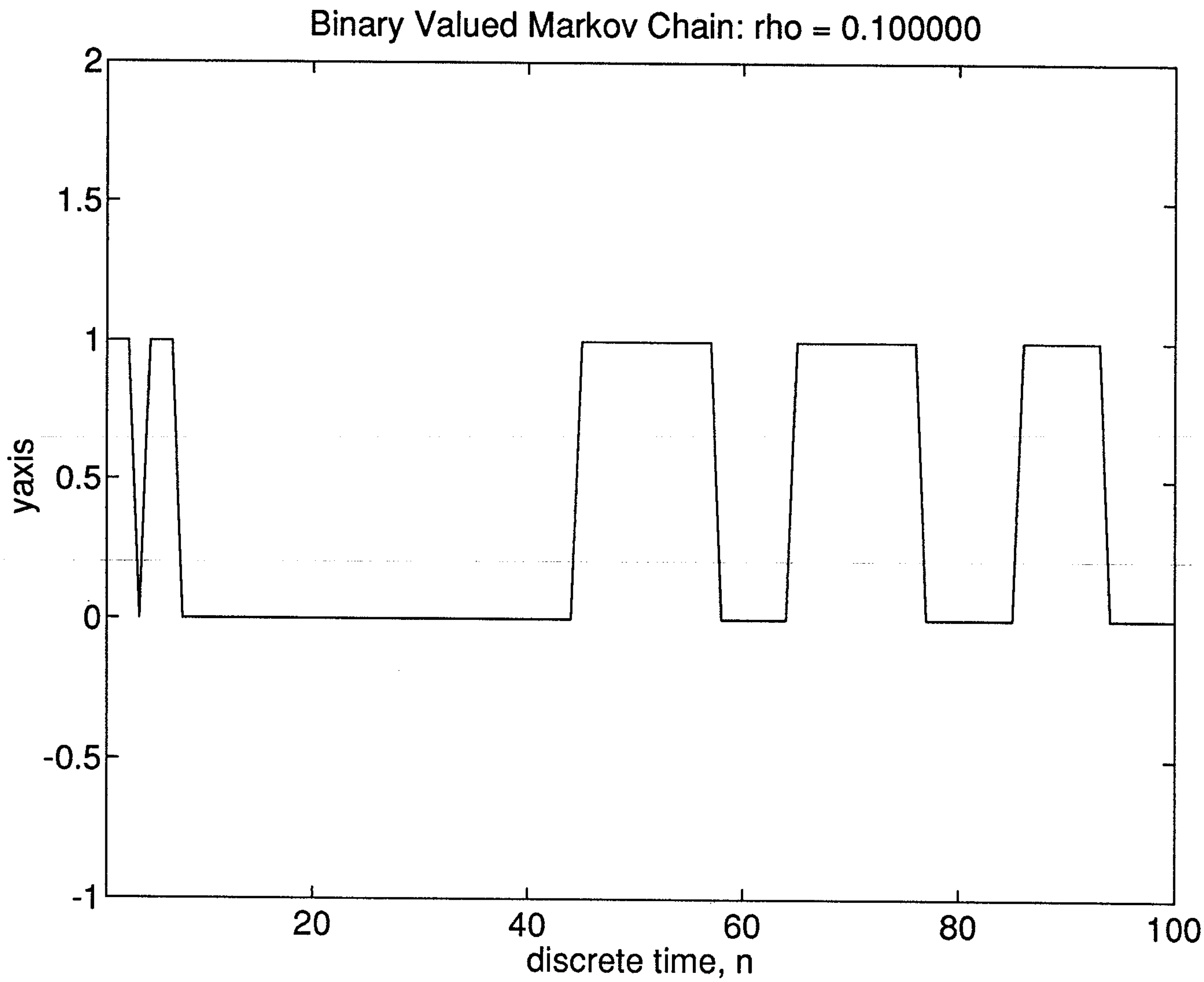
$$P\{Y_1 = 0\} = P\{Y_1 = 1\} = 1/2$$

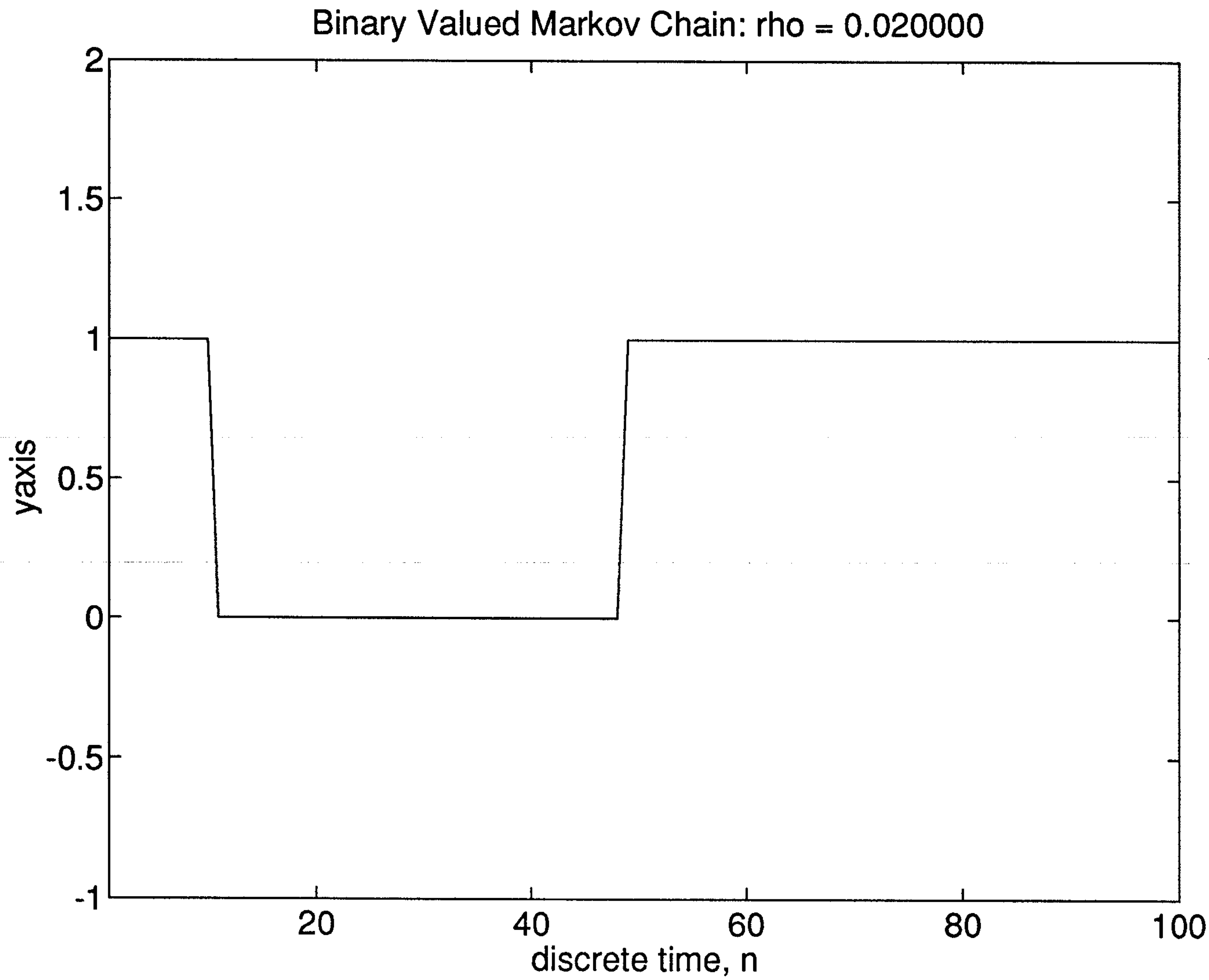
Define

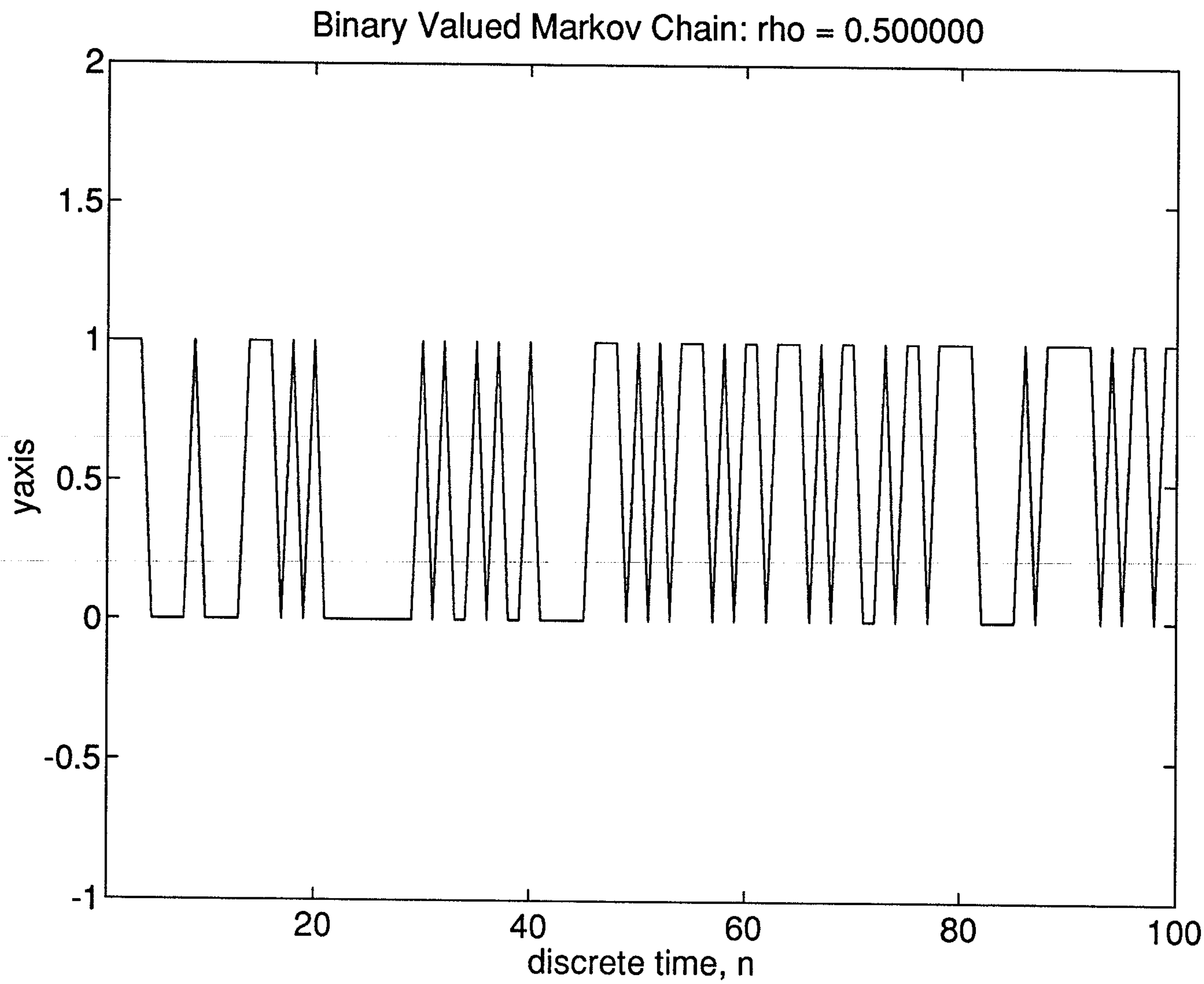
$$Y_{n+1} = \begin{cases} Y_n & B_n = 0 \\ 1 - Y_n & B_n = 1 \end{cases}$$

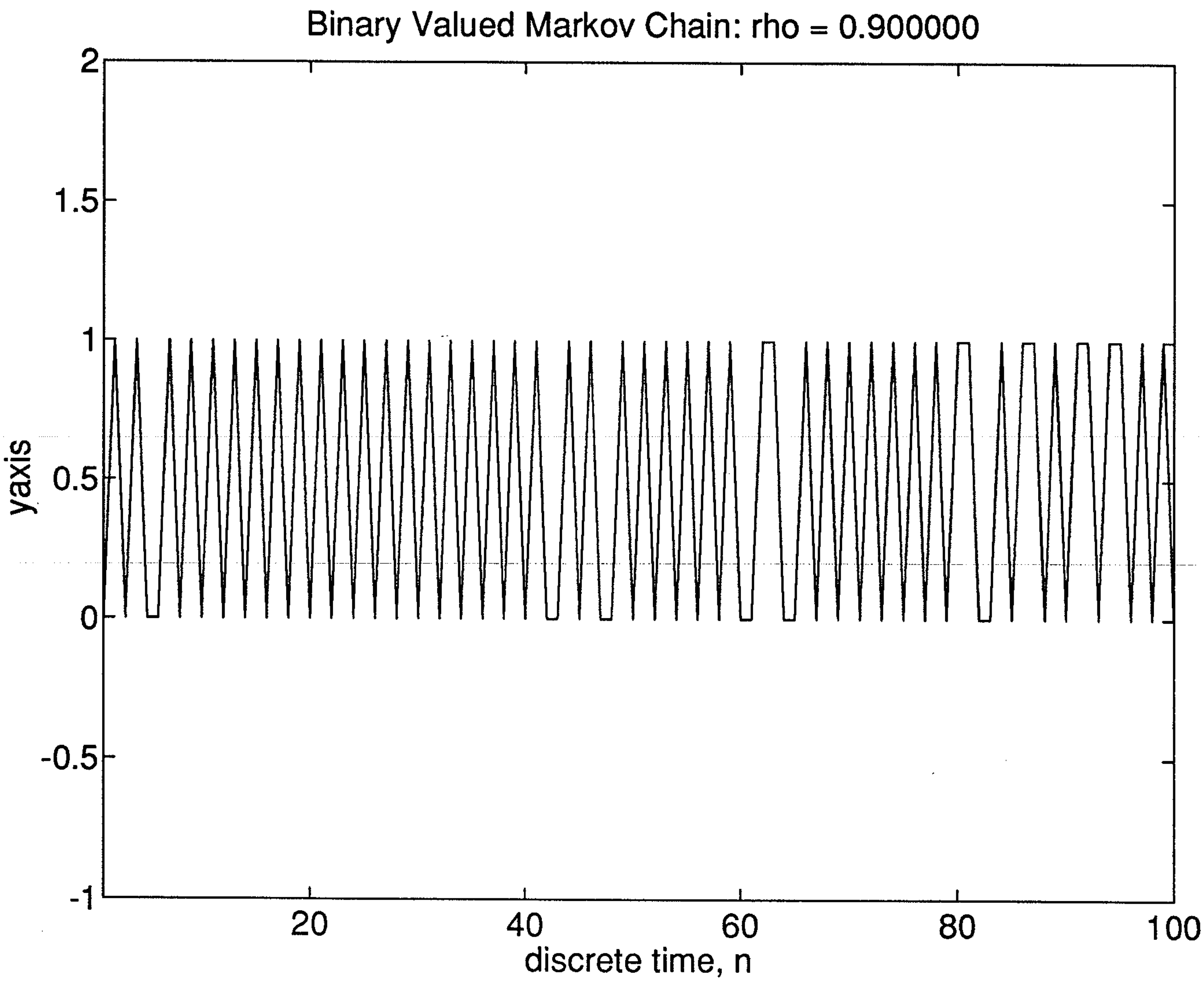
$B_n = 1 \Rightarrow$ transition
$B_n = 0 \Rightarrow$ no transition

$$p(y_n \mid y_i \; i<n) = p(y_n \mid y_{n-1})$$
$$= \begin{cases} 1-p & y_n = y_{n-1} \\ p & y_n = 1 - y_{n-1} \end{cases}$$


Binary Valued Markov Chain: rho = 0.100000
2
1.5
1
0.5
0
-0.5
-1
yaxis
20
40
60
80
100
discrete time, n


Binary Valued Markov Chain: rho = 0.020000
2
1.5
1
0.5
0
-0.5
-1
yaxis
20
40
60
80
100
discrete time, n


Binary Valued Markov Chain: rho = 0.500000
2
1.5
1
0.5
0
-0.5
-1
yaxis
20
40
60
80
100
discrete time, n



Binary Valued Markov Chain: rho = 0.900000
2
1.5
1
0.5
0
-0.5
-1
yaxis
20
40
60
80
100
discrete time, n

<u>Defn</u> X is a homogeneous Markov Sequence if $\forall n$

$$p_n(y_n \mid y_{n-1}) = p(y_n \mid y_{n-1})$$

"Transitions do not depend on time"

<u>Example</u> Continuous y_i from before

$$y_n = \sum_{i=0}^{n} x_i \quad \text{from above}$$

$$p_n(y_n \mid y_{n-1}) \sim N(y_{n-1}, \sigma^2)$$

not a function of $n \Rightarrow$ homogeneous

Stationary? No!

Lecture 13

Parameter Estimation of Markov Chains

$$p_\theta(y) = \left(\prod_{i=2}^{N+1} p_\theta(y_i | y_{i-1})\right) p(y_1)$$

$$\log p_\theta(y) = \sum_{i=2}^{N+1} \log p_\theta(y_i | y_{i-1}) + c$$

$$\hat{\theta}_{ML} = \arg\max_\theta \left\{ \sum_{i=2}^{N+1} \log p_\theta(y_i | y_{i-1}) \right\}$$

Example) y discrete as before

$$p_\rho(y) = (1-\rho)^{N-K} \rho^K (1/2)$$

$$\log p_\rho(y) = (N-K)\log(1-\rho) + K \log \rho + c$$

$$\frac{d}{d\rho} \log p_\rho(y) = \frac{N-K}{1-\rho}(-1) + \frac{K}{\rho} = 0$$

$$(N-K)\rho = K(1-\rho)$$

$$N\rho = K$$

$$\boxed{\rho = \frac{K}{N}}$$

Generalization

$p(x_n \mid x_i \; i<n) = p(x_n \mid x_{n-1}, x_{n-2}, \dots x_{n-p})$

Defn X_n is Markov with order p

Define $Y_n = \begin{bmatrix} X_n \\ X_{n-1} \\ \vdots \\ X_{n-p+1} \end{bmatrix}$

Then Y_n is Markov with order $p = 1$.