AI 5090: Stochastic Processes

Lecture 14

Scribe: Kavali Sofia Sagar & Siddharth Shrivastava

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

21 February 2025

Contents

1 Recap

1.1 Mean, Auto-correlation & Auto-covariance

Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, let $\{X_t : t \in \mathcal{T}\}$ be a random process defined over $\mathscr{F}$.

Definition 1 (Mean)**.** *The mean of the random process $\{X_t : t \in \mathcal{T}\}$ is a function $M_X : \mathcal{T} \to [-\infty, \infty]$ defined as:*

$$M_X(t) = \mathbb{E}[X_t], \qquad t \in \mathcal{T}.$$

Definition 2 (Autocorrelation and Autocovariance)**.** *The* **autocorrelation** *and* **autocovariance** *of the process $\{X_t : t \in \mathcal{T}\}$ are functions $R_X, C_X : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$, defined as:*

$$R_X(t,s) = \mathbb{E}[X_t X_s], \quad C_X(t,s) = \mathit{Cov}(X_t, X_s), \quad t, s \in \mathcal{T}.$$

1.2 Stationary Process

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{X_t : t \in \mathbb{R}^+\}$ be a random process.

Definition 3 (Stationary Process)**.** *A process $\{X_t : t \geq 0\}$ is said to be (strictly) stationary if all finite dimensional distributions (FDDs) are translation invariant, i.e., for any $n \in \mathbb{N}$, $\mathbf{t} \in \mathbb{R}^n_+$, and $h \in \mathbb{R}_+$,*

$$F_{\mathbf{t}} = F_{\mathbf{t}+h}.$$

Here, $\mathbf{t} + h$ is a vector with each coordinate incremented by h with respect to the corresponding coordinate in $\mathbf{t}$.

1.3 Weakly Stationary Process

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{X_t : t \in \mathbb{R}_+\}$ be a random process defined w.r.t. $\mathscr{F}$.

Definition 4 (Weakly Stationary Process)**.** *$\{X_t : t \in \mathbb{R}_+\}$ is said to be weakly stationary (or wide-sense stationary) if for all $t_1, t_2 \in \mathbb{R}_+$ and $h \in \mathbb{R}_+$:*

1. $M_X(t_1) = M_X(t_2)$.
2. $C_X(t_1, t_2) = C_X(t_1 + h, t_2 + h)$.

Some remarks are in order.

- A process is weakly stationary *iff* it has constant mean, and $C_X(t, t+h) = C_X(0, h)$ for all $t, h \in \mathbb{R}_+$.
- Every stationary process with finite variance is wide-sense stationary.

1.4 IID Process

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{X_t : t \in \mathbb{R}_+\}$ be a random process defined w.r.t. $\mathscr{F}$.

Definition 5 (IID Process)**.** *$\{X_t : t \in \mathbb{R}_+\}$ is said to be an IID process with the common CDF F if for any $n \in \mathbb{N}$ and $\mathbf{t} \in \mathbb{R}_+^n$,*

$$F_{\mathbf{t}}(\mathbf{x}) = \prod_{i=1}^{n} F(x_i), \quad \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

That is, for any $n \in \mathbb{N}$ and $\mathbf{t} = (t_1, \ldots, t_n)$, the random variables $(X_{t_1}, \ldots, X_{t_n})$ are IID.

We recall the following result on IID processes.

Lemma 1. *Suppose that $\{X_t : t \in \mathbb{R}_+\}$ is an IID Process. Then, the following hold.*

- *The FDDs of $\{X_t : t \in \mathbb{R}_+\}$ are consistent.*
- *$\{X_t : t \in \mathbb{R}_+\}$ is strictly stationary.*
  That is, every IID process is stationary.

2 Random Walk

Let $X_1, X_2, \ldots$ be $\mathbb{N}$-valued IID process. Let $S_0 := 0$, and for each $n \in \mathbb{N}$, let

$$S_n = \sum_{i=1}^{n} X_i.$$

Then, $\{S_n\}_{n \in \mathbb{N} \cup 0}$ is the random walk process.
Now, mean of $\{S_n\}_{n \in \mathbb{N} \cup 0}$ process is:

$$M_S(t) = \mathbb{E}[S_t] = t\,\mathbb{E}[X_1], \qquad t \in \mathbb{N} \cup \{0\},$$

and the autovariance function is given by

$$C_S(t, s) = \min(t, s) \cdot \mathrm{Var}(X_1), \qquad t, s \in \mathbb{N} \cup \{0\}.$$

Clearly, we can observe that $\{S_n\}_{n \in \mathbb{N} \cup 0}$ process is not wide sense stationary, since the mean is not constant across the random variables in the process and also auto-covariance is not a function only of the difference of time indexes of random variables.

Question 1. *Determine the joint PMF of $S_1, \ldots, S_n$, assuming that $X_1, X_2, \ldots$ are jointly discrete.*
Observe that, we need to find the joint pmf of $S_1, \ldots, S_n$ which implies that $X_1, \ldots, X_n$ are jointly discrete random variables. We have

$$\begin{aligned}
p_{S_1,\ldots,S_n}(s_1, \ldots s_n) &= \mathbb{P}(S_1 = s_1, \ldots, S_n = s_n) \\
&= \mathbb{P}(X_1 = s_1, X_2 = s_2 - s_1, \ldots, X_n = s_n - s_{n-1}) \\
&= \mathbb{P}(X1 = s_1) \cdot \mathbb{P}(X_2 = s_2 - s_1) \cdots \mathbb{P}(X_n = s_n - s_{n-1}) \qquad (\because X_1 \perp\!\!\!\perp X_2, \ldots \perp\!\!\!\perp X_n) \\
&= \mathbb{P}(X_1 = s_1) \cdot \prod_{i=2}^{n} \mathbb{P}(X_i = s_i - s_{i-1}) \\
&= p_{X_1}(s_1) \cdot \prod_{i=2}^{n} p_{X_i}(s_i - s_{i-1}).
\end{aligned}$$

Question 2. *If $X_1, X_2, \ldots$ are IID and $\mathbb{R}$-valued with a common PDF f_X, determine the joint pdf of $S_1, \ldots S_n$.*
Strategy: to use Jacobian transformation to transform the joint PDF of $X_1, \ldots, X_n$ to joint PDF of $S_1, \ldots, S_n$.
Now $X_1 = s_1, X_2 = s_2 - s_1, \ldots, X_n = s_n - s_{n-1}$. Therefore, Jacobian matrix ($\mathbf{J}$)

$$\mathbf{J} = \begin{bmatrix} \frac{\partial s_1}{\partial x_1} & \frac{\partial s_1}{\partial x_2} & \cdots & \frac{\partial s_1}{\partial x_n} \\ \frac{\partial s_2}{\partial x_1} & \frac{\partial s_2}{\partial x_2} & \cdots & \frac{\partial s_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial s_n}{\partial x_1} & \frac{\partial s_n}{\partial x_2} & \cdots & \frac{\partial s_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Thus, we see that $|\det(\mathbf{J})| = 1$. Using the Jacobian transformation formula, we get

$$\begin{aligned}
f_{S_1,S_2,\ldots,S_n}(s_1, s_2, \ldots, s_n) &= \frac{f_{X_1,X_2,\ldots,X_n}(s_1, s_2 - s_1, \ldots, s_n - s_{n-1})}{|\det(\mathbf{J})|} \\
&= f_{X_1,X_2,\ldots,X_n}(s_1, s_2 - s_1, \ldots, s_n - s_{n-1}) \\
&= f_X(s_1) \cdot \prod_{i=2}^{n} f_X(s_i - s_{i-1}) \qquad (\because X_1, \ldots, X_n \overset{\text{iid}}{\sim} f_X).
\end{aligned}$$

3 Gaussian Process

A continuous-time process, say $\{X_t : t \in \mathbb{R}\}$, is said to be a Gaussian process if for any $n \in \mathbb{N}$ and $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n$, the random variables $(X_{t_1}, \ldots, X_{t_n})$ are jointly continuous, and in particular jointly Gaussian, with the joint PDF

$$f_{\mathbf{t}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(K_{\mathbf{t}})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{t}})^\top K_{\mathbf{t}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{t}})\right), \quad \mathbf{x} \in \mathbb{R}^n,$$

where $\boldsymbol{\mu}_{\mathbf{t}}$ and $K_{\mathbf{t}}$ are given by

$$\boldsymbol{\mu}_{\mathbf{t}} = \begin{bmatrix} \mathbb{E}[X_{t_1}] \\ \vdots \\ \mathbb{E}[X_{t_n}] \end{bmatrix} = \begin{bmatrix} M_X(t_1) \\ \vdots \\ M_X(t_n) \end{bmatrix}, \qquad K_{\mathbf{t}} = \begin{bmatrix} C_X(t_1, t_1) & C_X(t_1, t_2) & \cdots & C_X(t_1, t_n) \\ C_X(t_2, t_1) & C_X(t_2, t_2) & \cdots & C_X(t_2, t_n) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_n, t_1) & C_X(t_n, t_2) & \cdots & C_X(t_n, t_n) \end{bmatrix}.$$

Lemma 2. *A wide sense stationary Gaussian process is stationary, i.e., $\forall n \in \mathbb{N}, \mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n, h \in \mathbb{R}, f_{\mathbf{t}+h} = f_{\mathbf{t}}$.*

Proof. If $\{X_t : t \in \mathbb{R}\}$ is a WSS Gaussian process, then it follows that

- $M_X(t)$ is a constant function.
- Auto-covariance function is translation invariant, i.e., $C_X(t, s) = C_X(t + h, s + h)$ for all $t, s, h \in \mathbb{R}$.

We know that the finite-dimensional joint PDFs of $\{X_n\}_{n=1}^{\infty}$ are specified as follows: for any $n \in \mathbb{N}$ and $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n$,

$$f_{\mathbf{t+h}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(K_{\mathbf{t+h}})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{t+h}})^\top K_{\mathbf{t+h}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{t+h}})\right), \quad \mathbf{x} \in \mathbb{R}^n$$

Because the mean function is constant, we have

$$\boldsymbol{\mu}_{\mathbf{t+h}} = \boldsymbol{\mu}_{\mathbf{t}}.$$

Furthermore, because the auto-covariance function is shift invariant, we must have

$$\begin{aligned} K_{\mathbf{t+h}} &= \begin{bmatrix} C_X(t_{1+h}, t_{1+h}) & C_X(t_{1+h}, t_{2+h}) & \cdots & C_X(t_{1+h}, t_{n+h}) \\ C_X(t_{2+h}, t_{1+h}) & C_X(t_{2+h}, t_{2+h}) & \cdots & C_X(t_{2+h}, t_{n+h}) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_{n+h}, t_{1+h}) & C_X(t_{n+h}, t_{2+h}) & \cdots & C_X(t_{n+h}, t_{n+h}) \end{bmatrix} \\ &= \begin{bmatrix} C_X(t_1, t_1) & C_X(t_1, t_2) & \cdots & C_X(t_1, t_n) \\ C_X(t_2, t_1) & C_X(t_2, t_2) & \cdots & C_X(t_2, t_n) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_n, t_1) & C_X(t_n, t_2) & \cdots & C_X(t_n, t_n) \end{bmatrix} \\ &= K_{\mathbf{t}}. \end{aligned}$$

Thus, it follows that

$$\begin{aligned} f_{\mathbf{t+h}}(\mathbf{x}) &= \frac{1}{\sqrt{(2\pi)^n \det(K_{\mathbf{t}})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{t}})^\top K_{\mathbf{t}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{t}})\right) \\ &= f_{\mathbf{t}}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n, \end{aligned}$$

thereby proving that a WSS Gaussian process is stationary.

4 Independent Increments of Random Walk

Let $X_1, X_2, \ldots$ be $\mathbb{N}$-valued IID process. Let $S_0 := 0$, and for each $n \in \mathbb{N}$, let

$$S_n = \sum_{i=1}^{n} X_i.$$

Then, $\{S_n\}_{n \in \mathbb{N} \cup 0}$ is the random walk process.

Claim 1. *$\{S_n\}_{n=0}^{\infty}$ has independent increments, i.e., $\forall n \in \mathbb{N}$ and $0 \leq t_1 < t_2 < \ldots < t_n$,*

$$S_{t_2} - S_{t_1}, S_{t_3} - S_{t_2}, \ldots, S_{t_n} - S_{t_{n-1}}$$

are independent random variables.

Proof.

$$\begin{aligned} S_{t_2} - S_{t_1} &= \sum_{i=t_1+1}^{t_2} X_i = g_1(X_{t_1+1}, X_{t_1+2}, \ldots, X_{t_2}), \\ S_{t_3} - S_{t_2} &= \sum_{i=t_2+1}^{t_3} X_i = g_2(X_{t_2+1}, X_{t_2+2}, \ldots, X_{t_3}), \\ &\vdots \\ S_{t_n} - S_{t_{n-1}} &= \sum_{i=t_{n-1}+1}^{t_n} X_i = g_{n-1}(X_{t_{n-1}+1}, X_{t_{n-1}+2}, \ldots, X_{t_n}). \end{aligned}$$

Because $X_1, X_2, \ldots, X_n$ are independent of each other, and functions $g_1, g_2, \ldots, g_{n-1}$ are each Borel-measurable, using the property that Borel-measurable functions of indepedent random variables are also independent, we get that $S_{t_2} - S_{t_1}, S_{t_3} - S_{t_2}, \ldots, S_{t_n} - S_{t_{n-1}}$ are independent random variables.

Claim 2. *$\{S_n\}_{n=0}^{\infty}$ has stationary increments, i.e., $\forall n \in \mathbb{N}$, $0 \le t_1 < t_2 < \ldots < t_n$, and $h \in \mathbb{R}$,*

$$(S_{t_2} - S_{t_1}, S_{t_3} - S_{t_2}, \ldots, S_{t_n} - S_{t_{n-1}}) \overset{d}{=} (S_{t_2+h} - S_{t_1+h}, S_{t_3+h} - S_{t_2+h}, \ldots, S_{t_n+h} - S_{t_{n-1}+h}),$$

where $\overset{d}{=}$ stands for "is distributed identically as".

Proof. Assume $\{X_n\}_{n=0}^{\infty}$ are discrete random variables. Now, Consider two random variables $S_{t_2} - S_{t_1}$ and $S_{t_2+h} - S_{t_1+h}$.

$$\begin{aligned}
p_{S_{t_2}-S_{t_1}}(s) &= \mathbb{P}(S_{t_2} - S_{t_1} = s) \\
&= \mathbb{P}\left(\sum_{i=t_1+1}^{t_2} X_i = s\right) \\
&= \sum_{\substack{(x_1,x_2,\ldots,x_{t_2-t_1}): \\ x_1+x_2+\ldots x_{t_2-t_1}=s}} \mathbb{P}(X_{t_1+1} = x_1, X_{t_1+2} = x_2, \ldots, X_{t_2} = x_{t_2-t_1}) \\
&= \sum_{\substack{(x_1,x_2,\ldots,x_{t_2-t_1}): \\ x_1+x_2+\ldots x_{t_2-t_1}=s}} \mathbb{P}(X_{t_1+1} = x_1) \cdot \mathbb{P}(X_{t_1+2} = x_2) \cdots \mathbb{P}(X_{t_2} = x_{t_2-t_1}) \qquad (\because X_1 \perp\!\!\!\perp X_2, \ldots \perp\!\!\!\perp X_n) \\
&= \sum_{\substack{(x_1,x_2,\ldots,x_{t_2-t_1}): \\ x_1+x_2+\ldots x_{t_2-t_1}=s}} \mathbb{P}(X_{t_1+1+h} = x_1) \cdot \mathbb{P}(X_{t_1+2+h} = x_2) \cdots \mathbb{P}(X_{t_2+h} = x_{t_2-t_1}) \qquad (\because X_1, X_2, \ldots \text{ are IID}) \\
&= \sum_{\substack{(x_1,x_2,\ldots,x_{t_2-t_1}): \\ x_1+x_2+\ldots x_{t_2-t_1}=s}} \mathbb{P}(X_{t_1+1+h} = x_1,\ X_{t_1+2+h} = x_2, \ldots, X_{t_2+h} = x_{t_2-t_1}) \\
&= \mathbb{P}\left(\sum_{i=t_1+1+h}^{t_2+h} X_i = s\right) \\
&= \mathbb{P}(S_{t_2+h} - S_{t_1+h} = s) \\
&= p_{S_{t_2+h}-S_{t_1+h}}(s).
\end{aligned}$$

Similarly, $(S_{t_3} - S_{t_2})$ and $(S_{t_3+h} - S_{t_2+h})$ have same distribution, ..., $(S_{t_n} - S_{t_{n-1}})$ and $(S_{t_n+h} - S_{t_{n-1}+h})$ have same distribution. We also note that that increments are independent of each other, as

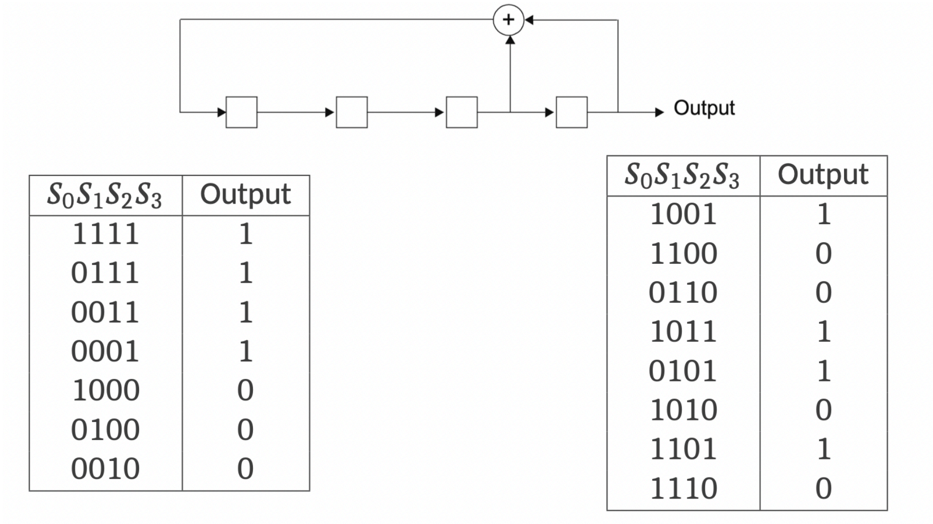
$$\begin{aligned}
p_{(S_{t_2}-S_{t_1}),\ldots,(S_{t_n}-S_{t_{n-1}})}(s_1, \ldots s_{n-1}) &= \mathbb{P}((S_{t_2} - S_{t_1}) = s_1, \ldots, (S_{t_n} - S_{t_{n-1}}) = s_{n-1}) \\
&= \mathbb{P}((S_{t_2} - S_{t_1}) = s_1) \cdots \mathbb{P}((S_{t_n} - S_{t_{n-1}}) = s_{n-1}) \\
&= \mathbb{P}((S_{t_2+h} - S_{t_1+h}) = s_1) \cdots \mathbb{P}((S_{t_n+h} - S_{t_{n-1}+h}) = s_{n-1}) \\
&= \mathbb{P}((S_{t_2+h} - S_{t_1+h}) = s_1, \ldots, (S_{t_n+h} - S_{t_{n-1+h}}) = s_{n-1}) \\
&= p_{(S_{t_2+h}-S_{t_1+h}),\ldots,(S_{t_n+h}-S_{t_{n-1}+h})}(s_1, \ldots s_{n-1})
\end{aligned}$$

Hence, it follows that

$$(S_{t_2} - S_{t_1}, S_{t_3} - S_{t_2}, \ldots, S_{t_n} - S_{t_{n-1}}) \overset{d}{=} (S_{t_2+h} - S_{t_1+h}, S_{t_3+h} - S_{t_2+h}, \ldots, S_{t_n+h} - S_{t_{n-1+h}}).$$

5 Binary Pseudo Random Number Generator

In this section, we shall study pseudo-random number generators (PRNGs) and explore how their outputs, though deterministic and inherently non-random, exhibit statistical properties remarkably similar to those of random sequences. Specifically, we will focus on binary pseudo-random number generators (binary PRNGs), whose outputs closely mimic sequences produced by an independent and identically distributed (IID) Bernoulli source.
It is crucial to emphasize that binary PRNGs generate sequences based purely on deterministic algorithms, meaning the same initial conditions or "seed" will always yield the same output sequence. Despite their deterministic nature, the generated sequences display desirable statistical properties, including an approximately equal distribution of 0s and 1s and minimal autocorrelation. These properties closely resemble the statistical behavior expected from IID Bernoulli random



$S_0S_1S_2S_3$	Output
1111	1
0111	1
0011	1
0001	1
1000	0
0100	0
0010	0

$S_0S_1S_2S_3$	Output
1001	1
1100	0
0110	0
1011	1
0101	1
1010	0
1101	1
1110	0

Figure 1: A four stage, binary linear feedback shift register; with its truth table

variables, making binary PRNGs valuable for applications that require repeatability, such as simulations, cryptographic algorithms, and testing scenarios.

To clearly illustrate these concepts, we will discuss a practical example involving a binary linear feedback shift register (LFSR), one of the simplest and most widely used binary PRNGs. A binary LFSR consists of shift registers and exclusive OR (XOR) gates to generate binary sequences cyclically. Binary LFSRs can produce sequences with maximal periods and favorable statistical properties by carefully choosing the feedback connections. This example demonstrates how such deterministic devices effectively produce binary sequences whose characteristics convincingly approximate randomness. Each block in the shift register, depicts a flip-flop and the feedback is collected through EX-OR gate. The output sequence (when the initial state of the 4 flip flops is $1, 1, 1, 1$) as shown in Figure 1:

$$1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0$$

- Number of zeros in one period = Number of ones in one period (which is required for uniform binary random number generator)
- Period: the smallest number of iterations after which the output starts repeating = 15 (which is not desired in binary random number generator, where after a certain iterations numbers starts repeating and therefore become deterministic)

The max number of states in a period for n-stage binary shift register is $2^n - 1$. Therefore to increase the number of states in a period such that no states are repeating (being not deterministic) itself, we can increase the number of stages in shift register.

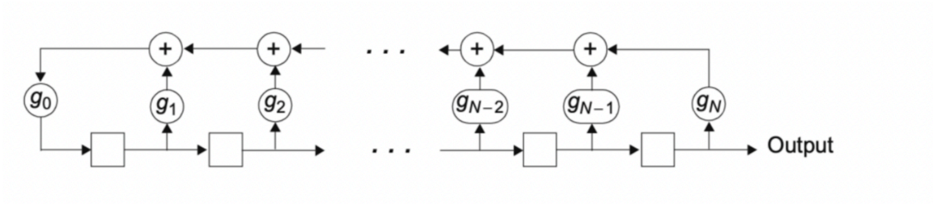


Figure 2: N-stage, binary linear feedback shift register

Concerning Figure 2, we can construct a configuration of an N-stage binary linear shift register with a $2^n - 1$ period. In the figure 2, g_i can be either 0 or 1, representing whether the corresponding feedback loop is present. Thus, in a computer using above mentioned binary linear feedback shift register with large period $2^n - 1$, we can generate a binary sequences which resemble those of an IID Bernoulli source. Thus there is nothing as random number in computer, and hence the name pseudo random generator. We will discuss more about PRNGs in the forthcoming lecture.