Theory of Random Number Generation

Ji Yong-Hyeon

Department of Information Security, Cryptology, and Mathematics
College of Science and Technology
Kookmin University

November 29, 2023

Contents

Chapter 1

Introduction

Summary

- Required Properties for Random Bit Generator
  - **Unpredictability**, **Unbiasedness**, **Independence**
- Components of Cryptographically Secure Random Bit Generator
  - TRNG (Entropy Source) + PRNG (Pseudorandom Number Generator)
- Methods for Evaluating the Security of Random Bit Generator
  - Estimation of entropy for the output sequence from TRNG
  - Statistical randomness tests for the output sequence from RNG
- Types of Random Bit Generators
  - Hardware/Software-based Random Bit Generators
  - Operating System-based Random Bit Generators
  - Various Standard Pseudorandom Number Generators

Functions of RBG (Random Bit Generator)

Provides random numbers required for cryptographic systems An essential element (algorithm) for the operation of cryptographic systems and modules Required Properties: Unpredictability, Unbiasedness, Independence between bits

Ideally, the output should be akin to the results of "coin tossing." Applications of Random Bit Generator

Generation of Key and Initialization Vector (IV) used in symmetric-key cryptography (block/stream ciphers) Generation of various parameters in public-key cryptography: prime number generation, probabilistic public-key cryptography, etc. Generation of various parameters used in cryptographic protocols: nonce, salt, etc.

Chapter 2

Probability Theory

2.1 Introduction

Definition 2.1.

- An **experiment** is the process of observing a phenomenon that has variation in its outcomes.
- The **sample space** S associated with an experiment is the collection of all possible distinct outcomes of the experiment.
- An **event** A, B is the set of elementary outcomes possessing a designated feature. ($A, B \subseteq S$)

Remark 2.1.

- Union: $A \cup B$
- Complement: A^C
- Intersection: $A \cap B$ (simply, AB)
- A, B are mutually disjoint $\iff A \cap B = \emptyset$

2.2 Axioms of Probability

2.2.1 Kolmogorov's Axiom

Kolmogorov's Axiom

Axiom. The probability is a function $\Pr : 2^{\Omega} \to [0,1] \subseteq \mathbb{R}$ satisfies

(A1) $\forall$event A, $0 \leq \Pr[A] \leq 1$.

(A2) $\Pr[\Omega] = 1$.

(A3) (Countable Additivity) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P[A_i]$, where $\{A_1, A_2, \dots\}$ is a countable set.

Remark 2.2. A probability is a function $\Pr : 2^{\Omega} \to [0,1] \subseteq \mathbb{R}$.

Proposition 2.1. *Let $A, B \subseteq \Omega$.*

(1) $\Pr[A] = \Pr[AB^C] + \Pr[AB]$

(2) $\Pr[B] = \Pr[AB] + \Pr[A^C B]$

(3) $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[AB]$

(4) $\Pr[A \cup B] = \Pr[AB^C] + \Pr[AB] + \Pr[A^C B]$

(5) $\Pr[A^C] = 1 - \Pr[A]$

(6) $A \subseteq B \implies \Pr[A] \leq \Pr[B]$

2.2.2 Conditional Probability and Independent

Conditional Probability

Definition 2.2. The **conditional probability** of A given B is denoted by $\Pr[A|B]$ and defined by the formula

$$\Pr[A|B] = \frac{\Pr[AB]}{\Pr[B]} \quad \text{with} \quad \Pr[B] > 0.$$

Equivalently, this formula can be written as **multiplication law of probability**:

$$\Pr[AB] = \Pr[A|B]\Pr[B].$$

Example 2.1.

(1) Start with a *shuffled deck of cards* and distribute all 52 cards to 4 player, 13 cards to each. What is the probability that each player gets an Ace?

(2) Next, assume that you are a player and you get a single Ace. What is the probability now that each player gets an Ace?

Solution.

(1) If any ordering of cards is equally likely, then any position of the four Aces in the deck is also equally likely. There are

$$\binom{52}{4} = \frac{52!}{4!48!}$$

possibilities for the positions (slots) for the 4 aces. Out of these, the number of positions that give each player an Ace 13^4 pick the first slot among the cards that the first player gets, then the second slot among the second player's card, then the third and the fourth slot. Therefore, the answer is $\frac{13^4}{\binom{52}{4}} \approx 0.1055$.

(2) After you see that you have a single Ace, the probability goes up the previous answer need to be divided by the probability that you get a single Ace, which is

$$\frac{13 \cdot \binom{39}{3}}{\binom{52}{4}} \approx 0.4388.$$

Note that

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}.$$

The answer then becomes $\frac{13^4}{13 \cdot \binom{39}{3}} \approx 0.2404$.

□

Independence

Definition 2.3. Two events A and B are **independent** if

$$\Pr[A|B] = \Pr[A]$$

Equivalent conditions are

$$\Pr[B|A] = \Pr[B] \quad \text{or} \quad \Pr[AB] = \Pr[A]\Pr[B]$$

Remark 2.3. $\Pr[A] = \Pr[A|B] = \dfrac{\Pr[AB]}{\Pr[B]} \implies \Pr[AB] = \Pr[A]\Pr[B]$.

Example 2.2. Suppose we roll a dice once. Let the universal set is $U = \{1,2,3,4,5,6\}$.

(1) (Independent but Not Disjoint) Let

$$A = \{1,3,5\} \quad \text{and} \quad B = \{3,6\}.$$

Then $A \cap B = \{3\} \neq \emptyset$, that is, A and B are not disjoint. Note that

$$\Pr[A] = \frac{3}{6} = \frac{1}{2}, \quad \Pr[B] = \frac{2}{6} = \frac{1}{3},$$
$$\Pr[A \mid B] = \frac{\Pr[AB]}{\Pr[B]} = \frac{1/6}{1/3} = \frac{1}{2}, \quad \Pr[B \mid A] = \frac{\Pr[BA]}{\Pr[A]} = \frac{1/6}{1/2} = \frac{1}{3}.$$

Thus, $\Pr[A|B] = \Pr[A]$ and $\Pr[B|A] = \Pr[B]$. That is, A and B are mutually independent.

(2) (Not Independent but Disjoint) Let

$$A = \{1,3,5\} \quad \text{and} \quad B = \{2,4,6\}.$$

Then $A \cap B = \emptyset$, that is, A and B are disjoint. Note that

$$\Pr[A] = \frac{3}{6} = \frac{1}{2}, \quad \Pr[B] = \frac{3}{6} = \frac{1}{2},$$
$$\Pr[A \mid B] = \frac{\Pr[AB]}{\Pr[B]} = \frac{0}{1/2} = 0, \quad \Pr[B \mid A] = \frac{\Pr[BA]}{\Pr[A]} = \frac{0}{1/2} = 0.$$

Thus, $\Pr[A|B] \neq \Pr[A]$ and $\Pr[B|A] \neq \Pr[B]$. That is, A and B are not independent.

Rule of Total Probailtiy

Proposition 2.2. *Let events $A_1, \ldots, A_n$ are satisfies*

(1) $\Pr[A_i] > 0$ *for* $i = 1, \ldots, n$

(2) $A_i \cap A_j = \emptyset$ *for* $i \neq j$

(3) $\bigcup_{i=1}^{n} A_i = \Omega$

Then

$$\begin{aligned}\Pr[B] &= \sum_{i=1}^{n} \Pr[B|A_i] \Pr[A_i] \\ &= \Pr[B|A_1] \Pr[A_1] + \Pr[B|A_2] \Pr[A_2] + \cdots + \Pr[B|A_n] \Pr[A_n].\end{aligned}$$

Proof. $B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n} (B \cap A_i)$. □

2.2.3 Bayes' Theorem

Bayes' Theorem

Theorem 2.3.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

The posterior probability of $\bar{B}$ *is then* $P(\bar{B}|A) = 1 - P(B|A)$.

Remark 2.4.

$$\Pr[B \mid A] = \frac{\Pr[A|B] \cdot \Pr[B]}{\Pr[A]} \iff \text{Posterior} = \frac{\text{Likelihood} \cdot \text{Prior}}{\text{Evidence}}.$$

2.3 Random Variables

Random Variable

Definition 2.4. A **random variable** X is real-valued function on Ω the space of outcomes:

$$X : \Omega \to \mathbb{R}.$$

In other words, a random variable is a number whose value depends upon the outcome of a random experiment.

Remark 2.5. Sometimes, when convenient, we also allow X to have the value ∞ or, more rarely, $-\infty$.

2.3.1 Discrete Random Variables

Discrete Random Variable

Definition 2.5. A **discrete random variable** X has finitely or countably many values

$$x_i \quad \text{for} \quad i = 1, 2, \cdots$$

and

$$p(x_i) = P(X = x_i)$$

with $i = 1, 2, \cdots$, is called the **probability mass function of** X.

Remark 2.6. A probability mass function p has the following properties:

(1) $x = x_i, i \in I \implies p(x) = \Pr[X = x_i]$

(2) $0 \leq p(x) \leq 1, \sum_{x \in X} p(x) = 1.$

(3) $\Pr[a < X \leq b] = \sum_{a < x \leq b} p(x).$

Discrete Probability Distribution

Definition 2.6. The **probability distribution** of a discrete of a random variable X is described as the function

$$f(x_i) = P(X = x_i)$$

which gives the probability for each value and satisfies:

1. $0 \leq f(x_i) \leq 1$ for each value x_i of X
2. $\displaystyle\sum_{i=1}^{k} f(x_i) = 1$

Expectation(Mean) and Standard Deviation of a Probability Distribution

Definition 2.7.

- The **mean** of X or **population mean**

$$E[X] = \mu$$
$$= \sum(\text{Value} \times \text{Probability}) = \sum x_i f(x_i)$$

  Here the sum extends over all the distinct values x_i of X.

- The **Variance and Standard Deviation of X** is given by

$$\sigma^2 = \text{Var}[X] = \sum (x_i - \mu)^2 f(x_i)$$
$$\sigma = \text{sd}[X] = +\sqrt{\text{Var}[X]}$$

- **Alternative Formula for Hand calculation:**

$$\sigma^2 = \sum x_i^2 f(x_i) - \mu^2$$

Example 2.3 (Calculating a Population Variance and Standard Deviation). Calculate the variance and the standard deviation of the distribution of X that appears in the left two columns of below table.

x	$f(x)$	$xf(x)$	$(x-\mu)$	$(x-\mu)^2$	$(x-\mu)^2 f(x)$	$x^2 f(x)$
0	.1	0	-2	4	.4	0
1	.2	.2	-1	1	.2	0.2
2	.4	.8	0	0	.0	1.6
3	.2	.6	1	1	.2	1.8
4	.1	.4	2	4	.4	1.6
Total	1.0	$2.0 = \mu$			$1.2 = \sigma^2$	$5.2 = \sum x^2 f(x)$

$$\text{Var}(X) = \sigma^2 = 1.2 \qquad \sigma^2 = 5.2 - (2.0)^2 = 1.2$$
$$\text{sd}(X) = \sigma = \sqrt{1.2} = 1.095 \qquad \sigma = \sqrt{1.2} = 1.095$$

2.3.2 Bernoulli

Note.

- The sample space $S = \{ \mathrm{S}, \mathrm{F} \}$.
- The probability of success $p = P(S)$, the probability of failure $q = P(F)$.
- $0 \leq p \leq 1$, $q = 1 - p$.

Binomial Distribution

Definition 2.8. The **binomial distribution** with n trails and success probability p is described by the function

$$f(x) = P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}$$

for the possible values $x = 0, 1, \cdots, n$.

Example 2.4 (An Example of the Binomial Distribution). The elementary outcomes of 4 samples, the associated probabilities, and the value of X are listed as follows.

FFFF	SFFF	SSFF	SSSF	SSSS
	FSFF	SFSF	SSFS	
	FFSF	SFFS	SFSS	
	FFFS	FSSF	FSSS	
		FSFS		
		FFSS		

Value of X	0	1	2	3	4
Probability of each outcome	q^4	pq^3	p^2q^2	p^3q	p^4
Number of outcomes	$1 = \binom{4}{0}$	$4 = \binom{4}{1}$	$6 = \binom{4}{2}$	$4 = \binom{4}{1}$	$1 = \binom{4}{4}$

Value x	0	1	2	3	4
Probability $f(x)$	$\binom{4}{0}p^0q^4$	$\binom{4}{1}p^1q^3$	$\binom{4}{2}p^2q^2$	$\binom{4}{1}p^3q^1$	$\binom{4}{4}p^4q^0$

Mean and Standard Deviation of the Binomial Distribution

Definition 2.9.

$$X = X_1 + X_2 + \cdots + X_n \sim B(n, p)$$

- $E[X] = E[X_1] + \cdots + E[X_n] = np$
- $\text{Var}[X] = \text{Var}[X_1] + \cdots + \text{Var}[X_n] = npq$

The binomial distribution with n trials and success probability p has

$$\begin{aligned} \text{Mean} &= np \\ \text{Variance} &= npq = np(1-p) \\ \text{sd} &= \sqrt{npq} \end{aligned}$$

Covariance and Correlation Coefficient of Two Random Variables

Definition 2.10. Let X, Y be a random variables. Then

1. The covariance of them:

$$Cov(X, Y) = E[(X - \mu_1)(Y - \mu_2)]$$

2. The correlation coefficient of them:

$$Corr(X, Y) = E\left[\left(\frac{X - \mu_1}{\sigma_1}\right)\left(\frac{Y - \mu_2}{\sigma_2}\right)\right] = \frac{Cov(X, Y)}{\text{sd}(X)\text{sd}(Y)}$$

Remark 2.7. Note that $-1 \leq Corr(X, Y) \leq 1$ and

$$\begin{aligned} Cov(X, Y) &= E[(X - \mu_1)(Y - \mu_2)] \\ &= E[XY - \mu_2 X - \mu_1 Y + \mu_1\mu_2] \\ &= E[XY] - \mu_2 E[X] - \mu_1 E[Y] + \mu_1\mu_2 \\ &= E[XY] - \mu_1\mu_2. \end{aligned}$$

That is, $Cov(X, Y) = E[XY] - \mu_1\mu_2$.

Proposition 2.4.

(1) $Cov(aX + b, cY + d) = ac \cdot Cov(X, Y)$

(2) $$Corr(aX + b, cY + d) = \begin{cases} Corr(X, Y) & : ac > 0 \\ -Corr(X, Y) & : ac < 0 \end{cases}$$

Proof. (1)

$$\begin{aligned} Cov(aX+b, cY+d) &= E[(aX+b) - (a\mu_x + b) \cdot (cY + d - (c\mu_y + d))] \\ &= E[a(X-\mu_x) \cdot c(Y-\mu_y)] = acE[(X-\mu_x)(Y-\mu_y)] \\ &= ac \cdot Cov(X,Y). \end{aligned}$$

(2) Note that $\sigma_{aX+b} = \sqrt{\text{Var}(aX+b)} = \sqrt{a^2\text{Var}(X)} = |a|\,\sigma_X$. Similarly $\sigma_{cY+d} = |c|\,\sigma_Y$.

$$Corr(aX+b, cY+d) = \frac{Cov(aX+b, cY+d)}{\sigma_{aX+b}\sigma_{cY+d}} = \frac{ac \cdot Cov(X,Y)}{|a|\,\sigma_X |c|\,\sigma_Y} = \frac{ac}{|ac|} Corr(X,Y).$$

$$\text{Hence, } Corr(aX+b, cY+d) = \begin{cases} Corr(X,Y) & \text{if } ac > 0 \\ -Corr(X,Y) & \text{if } ac < 0 \end{cases}$$

□

Distribution of Sum of Two Probability Variables

Proposition 2.5.

(1) $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

(2) $Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$

Two Probability Variables are Independent

Proposition 2.6.

(1) $E[XY] = E[X] \cdot E[Y]$

(2) $Cov(X,Y) = 0, Corr(X,Y) = 0$

(3) $Var(X \pm Y) = Var(X) + Var(Y)$

Proof. (1)

$$\begin{aligned} E[XY] &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i y_j p(x_i, y_j) \\ &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i y_j p_1(x_i) p_2(y_j) \\ &= \sum_{i=1}^{\infty} x_i p_1(x_i) \sum_{j=1}^{\infty} y_j p_2(y_j) \\ &= E[X] \cdot E[Y]. \end{aligned}$$

(2) $Cov(X,Y) = E[XY] - E[X] \cdot E[Y] = 0.$

□

2.3.3 Continuous Random Variables

Probability Density Function

Definition 2.11. The **probability density function** $f(x)$ describes the distribution of probability for a continuous random variable. It has the properties:

(1) The total area under the probability density curve is 1.

(2) $P[a \leq X \leq b]$ = area under the probability density curve between a and b.

(3) $f(x) \geq 0$ for all x.

Remark 2.8. With a continuous random variable, the probability that $X = x$ is **always** 0. It is only meaningful to speak about the probability that X lies in an interval.

Remark 2.9. $p(x)$ is called **probability density function** of continuous random variable X if $p(x)$ satisfies:

(i) $p(x) \geq 0$, $\int_{-\infty}^{\infty} p(x)\,dx = 1$,

(ii) $P(a \leq X \leq b) = \int_{a}^{b} p(x)\,dx$.

Note that

- For any constant c, $\int_{c}^{c} p(x)\,dx = 0$.
- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$.

Expectation of a Continuous Random Variable

Definition 2.12.

- Expectation(or Mean) of a Random Variable X
  - a Discrete Random variable: $E[X] = \sum_{i=1}^{\infty} x_i p(x_i)$
  - a Continuous Random variable: $E[X] = \int_{-\infty}^{\infty} xp(x)\,dx$
- Expectation and Median of a Continuous Random Variable X
  - Expectation($\mu = E[X]$): the balance point of the probability mass.
  - Median: the value of X that divides the area under the curve into halves.

2.3.4 Normal Random Variable

Normal Random Variable

Definition 2.13. A random variable is **Normal with parameter** $\mu \in \mathbb{R}$ **and** $\sigma^2 > 0$ or, in short, X **is** $N(\mu, \sigma^2)$, if its density is the function given below.

$$\text{Density}: \ f(x) = f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$

where $x \in (-\infty, \infty)$.

Proposition 2.7.

(1) $For f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right], \quad \displaystyle\int_{-\infty}^{\infty} f(x)\, dx = 1.$

(2) $EX = \mu.$

(3) $Var(X) = \sigma^2.$

2.3.5 The Normal Approximation to the Binomial

The Normal Approximation to the Binomial

Theorem 2.8. *When np and $np(1-p)$ are both large, say, greater than 15, the binomial distribution is well approximated by the normal distribution having mean $= np$ and sd $= \sqrt{np(1-p)}$. That is,*

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \text{ is approximately } N(0,1).$$

Mean and Standard Deviation of $\overline{X}$

Definition 2.14. The distribution of the sample mean, based on a random sample of size n, has

$$E[\overline{X}] = \mu \qquad (= \text{Population mean})$$

$$\text{Var}[\overline{X}] = \frac{\sigma^2}{n} \qquad \left(= \frac{\text{Population variance}}{\text{Sample size}}\right)$$

$$\text{sd}[\overline{X}] = \frac{\sigma}{\sqrt{n}} \qquad \left(= \frac{\text{Population standard deviation}}{\sqrt{\text{Sample size}}}\right)$$

2.4 Central Limit Theorem

2.4.1 CLT

Central Limit Theorem

Theorem 2.9. *Assume that $X, X_1, X_2, \ldots$ are independent, identically distributed random variables, with finite $\mu = EX$ and $\sigma^2 = Var[X]$. Then,*

$$\lim_{n\to\infty} \Pr\left[\frac{\sum_{i=1}^{n} X_i - \mu n}{\sigma\sqrt{n}} \leq x\right] = \Pr\left[Z \leq x\right],$$

where Z is standard Normal.

2.4.2 Laws of Large Numbers

Weak Law of Large Numbers

Theorem 2.10. *Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean $E[X_i] = \mu$ and variance σ^2. Then, for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \Pr\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right| \geq \varepsilon\right] = 0.$$

Strong Law of Large Numbers

Theorem 2.11. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with a finite first moment, $\mathbb{E}[X_i] = \mu$. Then*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \quad \textit{almost surely as} \quad n \to \infty.$$

2.5 Problem: RBG → RNG

Exercise 2.1 (Uniform Distribution)**.** Consider an algorithm

Step 1: Drive the RBG independently 4 times to generate a 4-bit integer value r.
Step 2: **If** $r < 10$ **then**
return r
else
go to Step 1

Prove that

$$\Pr[\text{ouput} = n] = \frac{1}{10}$$

for $n = 0, 1, 2, \ldots, 9$.

Solution. Let

- $n \leq 2^k = m$ with $n = 10, k = 4$ and $m = 16$
- Output digit $= r \in [0, 9]$.

Then

output	**1st iteration**	**2nd iteration**	$\cdots$	**step iteration**
0	$\Pr[0] = \frac{1}{m}$	$\Pr[0] = \frac{1}{m} \cdot \frac{m-n}{m}$	$\cdots$	$\Pr[0] = \frac{1}{m} \cdot \left(\frac{m-n}{m}\right)^{\textbf{step}-1}$
1	$\Pr[1] = \frac{1}{m}$	$\Pr[1] = \frac{1}{m} \cdot \frac{m-n}{m}$	$\cdots$	$\Pr[1] = \frac{1}{m} \cdot \left(\frac{m-n}{m}\right)^{\textbf{step}-1}$
2	$\Pr[2] = \frac{1}{m}$	$\Pr[2] = \frac{1}{m} \cdot \frac{m-n}{m}$	$\cdots$	$\Pr[2] = \frac{1}{m} \cdot \left(\frac{m-n}{m}\right)^{\textbf{step}-1}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$n-1$	$\Pr[n-1] = \frac{1}{m}$	$\Pr[n-1] = \frac{1}{m} \cdot \frac{m-n}{m}$	$\cdots$	$\Pr[n-1] = \frac{1}{m} \cdot \left(\frac{m-n}{m}\right)^{\textbf{step}-1}$

Thus,

$$\begin{aligned}
\Pr[\text{output} = r] &= \frac{1}{m} + \frac{1}{m} \cdot \frac{m-n}{m} + \cdots + \frac{1}{m} \cdot \left(\frac{m-n}{m}\right)^s + \cdots \\
&= \frac{1}{m} \sum_{s=0}^{\infty} \left(\frac{m-n}{m}\right)^s = \frac{1}{m} \sum_{s=0}^{\infty} \left(1 - \frac{n}{m}\right)^s \\
&= \frac{1}{m} \cdot \frac{1}{1 - \left(1 - \frac{n}{m}\right)} = \frac{1}{m} \cdot \frac{m}{n} \\
&= \frac{1}{n}.
\end{aligned}$$

□

Chapter 3

Markov Chains

3.1 Introduction

Markov Chain

Definition 3.1. Let

$$\langle X_n \rangle_{n \geq 0} := \{X_n : n = 0, 1, 2, \cdots\}$$

be a stochastic process over a countable set S. Let $\Pr[X]$ is the probability of the random variable X. Then $\langle X_n \rangle_{n \geq 0}$ satisfies **Markov property** if

$$\Pr[X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n] = \Pr[X_{n+1} = x_{n+1} \mid X_n = x_n]$$

for all $n \geq 0$ and all $x_0, \ldots, x_{n+1} \in S$. Then $\langle X_n \rangle_{n \geq 0}$ is a **Markov chain**.

Remark 3.1.

(1) The conditional probability of X_{i+1} is dependent only upon X_i, and upon no earlier values of $\langle X_n \rangle$

(2) the state of $\langle X_n \rangle$ in the future is unaffected by its history.

(3) The set S is called the **state space** of the Markov chain.

(4) The conditional probabilities $\Pr[X_{n+1} = y \mid X_n = x]$ are called the **transition probabilities**.

(5) Markov chain having **stationary transition probabilities**, i.e., $\Pr(X_{n+1} = y \mid X_n = x)$, is independent of n.

Example 3.1 (*The general two-state Markov chain*)**.** There are two states 0 and 1 with transitions

- $0 \to 1$ with probability p
- $0 \to 0$ with probability $1 - p$
- $1 \to 0$ with probability q
- $1 \to 1$ with probability $1 - q$.

Thus we have

$$\begin{aligned}\Pr\left[X_{n+1} = 1 \mid X_n = 0\right] &= p,\\ \Pr\left[X_{n+1} = 0 \mid X_n = 1\right] &= q,\end{aligned}$$

and $\Pr[X_0 = 0] = \pi_0(0)$. Since there are only two states, 0 and 1, it follows immediately that

$$\begin{aligned}\Pr\left[X_{n+1} = 0 \mid X_n = 0\right] &= 1 - p,\\ \Pr\left[X_{n+1} = 1 \mid X_n = 1\right] &= 1 - q,\end{aligned}$$

and $\pi_0(1) = \Pr[X_0 = 1] = 1 - \pi_0(0)$. The transition matrix has two parameters $p, q \in [0, 1]$:

$$\begin{bmatrix} T_{00} & T_{01} \\ T_{10} & T_{11} \end{bmatrix} = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}.$$

Note that

- $\Pr[A] = \Pr[B \cap A] + \Pr[B^C \cap A]$
- $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B \mid A]$

Then we observe that

$$\begin{aligned}\Pr\left[X_{n+1} = 0\right] &= \Pr\left[X_n = 0 \wedge X_{n+1} = 0\right] + \Pr\left[X_n = 1 \wedge X_{n+1} = 0\right]\\ &= \Pr[X_n = 0]\Pr[X_{n+1} = 0 \mid X_n = 0]\\ &\quad + \Pr[X_n = 1]\Pr[X_{n+1} = 0 \mid X_n = 1]\\ &= \Pr[X_n = 0] \cdot (1 - p) + \Pr[X_n = 1] \cdot q\\ &= (1 - p) \cdot \Pr[X_n = 0] + q \cdot \big(1 - \Pr[X_n = 0]\big)\\ &= (1 - p - q) \cdot \Pr[X_n = 0] + q.\end{aligned}$$

Now

$$\begin{aligned}\Pr[X_0 = 0] &= \pi_0(0),\\ \Pr[X_1 = 0] &= (1 - p - q)\pi_0(0) + q,\\ \Pr[X_2 = 0] &= (1 - p - q)\Pr[X_1 = 0] + q\\ &= (1 - p - q)^2\pi_0(0) + q\left(1 + (1 - p - q)\right),\\ \Pr[X_3 = 0] &= (1 - p - q)\Pr[X_2 = 0] + q\\ &= (1 - p - q)^3\pi_0(0) + q\left(1 + (1 - p - q) + (1 - p - q)^2\right),\\ &\vdots\\ \Pr[X_n = 0] &= (1 - p - q)^n\pi_0(0) + q\sum_{j=0}^{n-1}(1 - p - q)^j.\end{aligned}$$

In the trivial case $p = q = 0$, it is clear that for all n

$$\Pr[X_n = 0] = \pi_0(0) \quad \text{and} \quad \Pr[X_n = 1] = \pi_0(1).$$

Suppose that $p + q > 0$. By the formula $\displaystyle\sum_{j=0}^{n-1} r^j = \frac{1 - r^n}{1 - r}$ for the sum of a finite geometric progression,

$$\sum_{j=0}^{n-1}(1 - p - q)^j = \frac{1 - (1 - p - q)^n}{p + q}.$$

Thus,

$$\Pr[X_n = 0] = \frac{q}{p + q} + (1 - p - q)^n \cdot \left(\pi_0(0) - \frac{q}{p + q}\right),$$
$$\Pr[X_n = 1] = \frac{p}{p + q} + (1 - p - q)^n \cdot \left(\pi_0(1) - \frac{p}{p + q}\right).$$

Suppose that $p, q \notin \{0, 1\}$. Then

$$0 < p + q < 2 \implies \left|1 - p - q\right| \leq 1.$$

Then

$$\lim_{n\to\infty} \Pr[X_n = 0] = \frac{q}{p + q} \quad \text{and} \quad \lim_{n\to\infty} \Pr[X_n = 1] = \frac{p}{p + q}.$$

Suppose we want to choose $\pi_0(0)$ and $\pi_0(1)$ so that $\Pr[X_n = 0]$ and $\Pr[X_n = 1]$ are independent of n. To do this, we should choose $\pi_0(0) = q/(p + q)$ and $\pi_0(1) = p/(p + q)$. Thus if $\langle X_n \rangle_{n\geq 0}$ start with the initial distribution

$$\pi_0 = \Pr[X_n = 0] = \frac{q}{p + q} \quad \text{and} \quad \pi_0(1) = \frac{p}{p + q},$$

then for all n

$$\Pr[X_n = 0] = \frac{q}{p + q} \quad \text{and} \quad \Pr[X_n = 1] = \frac{p}{p + q}.$$

Example 3.2. Let $n = 2$ and $x_0, x_1, x_2 \in \{0, 1\}$. Then

$$\begin{aligned}
&\Pr[X_0 = x_0, X_1 = x_1, X_2 = x_2] \\
&= \Pr[X_0 = x_0, X_1 = x_1] \cdot \Pr[X_2 = x_2 \mid X_0 = x_0, X_1 = x_1] \\
&= \Pr[X_0 = x_0] \Pr[X_1 = x_1 \mid X_0 = x_0] \cdot \Pr[X_2 = x_2 \mid X_0 = x_0, X_1 = x_1].
\end{aligned}$$

If the Markov property is satisfied, then

$$\Pr[X_2 = x_2 \mid X_0 = x_0, X_1 = x_1] = \Pr[X_2 = 2 \mid X_1 = x_1],$$

which is determined by p and q. In this case

$$\Pr[X_0 = x_0, X_1 = x_1, X_2 = x_2] = \Pr[X_0 = x_0] \Pr[X_1 = x_1 \mid X_0 = x_0] \Pr[X_2 = x_2 \mid X_1 = x_1].$$

Recall that the transition matrix with $p, q \in [0, 1]$:

$$\begin{bmatrix} T_{00} & T_{01} \\ T_{10} & T_{11} \end{bmatrix} = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}.$$

Then

x_0	x_1	x_2	$\Pr[X_0 = x_0, X_1 = x_1, X_2 = x_2]$
0	0	0	$\pi_0(0)(1-p)^2$
0	0	1	$\pi_0(0)(1-p)p$
0	1	0	$\pi_0(0)pq$
0	1	1	$\pi_0(0)p(1-q)$
1	0	0	$(1-\pi_0(0))q(1-p)$
1	0	1	$(1-\pi_0(0))qp$
1	1	0	$(1-\pi_0(0))(1-q)q$
1	1	1	$(1-\pi_0(0))(1-q)^2$

Chapter 4

Examples of Markov chains

Chapter 5

Statistical Inferences

Chapter 6

Statistical Tests for Randomness

Statistical tests for random number generators (RNGs) are crucial in assessing the quality of the RNGs, which purport to produce random sequences. While it is not possible to mathematically prove that a sequence is random, these tests can reveal certain weaknesses. The tests work by evaluating sample outputs from the RNG against attributes expected from truly random sequences. The outcomes of these tests are probabilistic and not definite. If a sequence fails any of the tests, it may be rejected as non-random, or subjected to further testing. Passing all the tests allows the RNG to be accepted as random, or more precisely, not rejected, as it only provides probabilistic evidence of randomness.

6.1 Statistical tests for RNGs

Statistical tests for RNGs measure the quality of a bit generator's randomness. These tests are essential in identifying weaknesses in the generators by applying various statistical methods to the output sequences. For instance, a sequence that passes all the tests is not definitively random but is likely to exhibit characteristics of randomness, whereas a sequence that fails any test is potentially non-random and may require additional testing or rejection.

6.1.1 Golomb's randomness postulates

Golomb's randomness postulates are historical attempts to define necessary conditions for periodic pseudorandom sequences to appear random. They are not sufficient conditions for randomness but were among the first efforts to systematically address the randomness in sequences. These postulates serve as a fundamental basis for more complex tests and are critical in understanding the nature of pseudorandom sequences and their applications.

Definition 6.1. Let

$$s = s_0, s_1, s_2, \dots$$

be an infnite sequence. The subsequence consisting of the first n terms of s is denoted by

$$s^n = s_0, s_1, \cdots, s_{n-1}.$$

Remark 6.1. s is the bit sequence if $s_i \in \{0, 1\}$.

N-Periodic

Definition 6.2. The sequence $s = s_0, s_1, s_2, \dots$ is said to be **N-periodic** if

$$s_i = s_{i+N}$$

for all $i \geq 0$.

Remark 6.2. If s is a N-periodic sequence, then the **cycle** of s is the subsequence s^N.

Run - Gap / Block

Definition 6.3. Let s be a sequence.

- A **run** of s is a subsequence of s consisting of consecutive 0's or 1's.
- A run of 0's is called a **gap**.
- A run of 1's is called a **block**.

Autocorrelation Function

Definition 6.4. Let s be a N-periodic sequence. The **autocorrelation function** of s is the integer-value function $C(t) : \{s_i\} \to \mathbb{Z}$ defined as

$$C(t) = \frac{1}{N} \sum_{i=0}^{N-1} (2s_i - 1) \cdot (2s_{i+t} - 1)$$

for $0 \leq t \leq N - 1$.

Remark 6.3.

Remark 6.4. The autocorrelation function $C(t)$ measure the amount of similarity between the sequence s and a shift of s by t positions. If s is a random N-periodic sequence, then $\left|N \cdot C(t)\right|$ can be expected to be quite small for all vlaue of $t \in (0, N)$.

Golomb's randomness postulates

Definition 6.5. Let s be a N-periodic sequence. **Golomb's randomness postulates** are as follows:

R1 In the cycle s^N of s, the number of 1's differs from the number of 0's by at most 1.

R2 In the cycle s^N, at least half the runs have length 1, at least one-fourth have length 2, at least one-eighth have length 3, and so on, as long as the number of runs so indicated exceeds 1. Moreover, for each of these lengths, there are (almost) equally many gaps and blocks.

R3 The autocorrelation function $C(t)$ is two-valued. That is for some integer K,

$$N \cdot C(t) = \sum_{i=0}^{N-1} (2s_i - 1) \cdot (2s_{i+t} - 1) = \begin{cases} N, & : t = 0, \\ K, & : 1 \leq t \leq N - 1. \end{cases}$$

Pseudo-Noise Sequence (pn-sequence)

Definition 6.6. A binary sequence which satisfies Golomb's randomness postulates is called a **pseudo-noise sequence** (**pn-sequence**).

Remark 6.5. Pseudo-noise sequences arise in practice as output sequences of **maximum-length linear**

The significance of Golomb's randomness postulates lies in their ability to provide a framework for the evaluation of periodic sequences. A periodic sequence with period N is assessed based on the following criteria:

6.1.2 Golomb's randomness postulates

Solomon W. Golomb's randomness postulates are a cornerstone in the field of cryptography and sequence design, primarily due to their foundational role in the analysis and interpretation of sequences as random. These postulates provide a mathematical framework for evaluating the randomness of binary sequences. Specifically, they set forth criteria that a sequence must meet to be considered random.

The postulates are as follows:

1. The number of ones and zeros in the sequence should be approximately equal, which is necessary for the sequence to have no bias.

2. The distance between consecutive ones should follow a geometric distribution. For binary sequences that are infinitely long, this implies that the probability of a 'gap' of length n between ones is $(1/2)^{n+1}$, representing the lack of structure in the sequence.

3. The sequence should be balanced, which means that for any binary substring, the number of occurrences of this substring and its complement should be approximately the same. This property is also referred to as the run property, where a 'run' is a substring of consecutive identical digits.

Golomb's postulates are integral in the design and analysis of pseudo-random number generators (PRNGs), as they serve as a benchmark for the sequence's randomness. A sequence that fulfills these postulates is considered to be a good candidate for cryptographic applications because it exhibits the unpredictability necessary for securing communications.

Application in Cryptography: In cryptography, the randomness of key material is paramount. Golomb's postulates are used to ensure that the generated keys do not exhibit patterns or regularities that could be exploited by adversaries. By applying these postulates to evaluate the randomness of binary sequences, cryptographers can quantify the security level of cryptographic systems.

Mathematical Implications: The postulates form the basis of several statistical tests, such as the runs test and the autocorrelation test. These tests are applied to binary sequences to check for the presence of patterns and correlations that would indicate non-randomness. The theoretical underpinnings of Golomb's postulates also contribute to the field of combinatorics and information theory, where they have implications for the construction of codes and error correction.

In conclusion, Golomb's randomness postulates are not only historically significant but also remain highly relevant in the modern analysis of cryptographic systems. They

Chapter 7

Statistical Tests for Randomness

7.1 Introduction

Statistical tests for randomness play a pivotal role in validating the quality and integrity of random number generators (RNGs). These tests serve as a benchmark against which RNGs are measured to ensure that their output sequences exhibit properties characteristic of true randomness. Despite the intrinsic limitations in proving randomness, these tests can effectively identify non-randomness in sequences, which is vital for applications in cryptography, simulations, and various stochastic modeling scenarios.

7.2 Statistical tests for RNGs

The purpose of statistical tests for RNGs is to analyze sequences for unpredictability, lack of patterns, and uniform distribution—traits expected from ideal random sequences. These tests range from simple frequency analysis to complex tests for serial correlation, and no single test can validate randomness conclusively. Hence, a battery of tests is typically employed, where failure of any test suggests non-randomness, prompting further scrutiny or rejection of the RNG.

7.2.1 Golomb's randomness postulates

Golomb's randomness postulates, rooted in the theory of shift register sequences, lay foundational criteria for assessing the randomness of periodic sequences. These postulates dictate:

1. **The Frequency Postulate**: The number of zeros and ones in a sequence should be approximately the same, reflecting the balance of a sequence.

2. **The Run Postulate**: A sequence should contain runs of various lengths distributed according to expected probabilities. For instance, in a binary sequence, half of the runs should be of length one, one-fourth should be of length two, and so forth.

3. **The Autocorrelation Postulate**: The autocorrelation function of the sequence should rapidly drop to zero as the shift increases, which implies that each bit should be independent of others at a certain distance.

These postulates form the basis for more sophisticated tests, and while they are not sufficient to declare a sequence as random, they are necessary conditions. In particular, they address the uniformity and independence of a sequence, which are critical aspects in cryptographic applications.

Implementing these postulates in practical statistical tests has advanced the analysis of pseudorandom number generators (PRNGs), providing a methodology to assess their suitability for various applications. The evaluation of PRNGs against these postulates often involves chi-squared tests, spectral tests, and other statistical methodologies to detect non-random behavior in generated sequences.