Ceng 471 Cryptography

Shannon’s Theory
Perfect Secrecy
Entropy
Birthday Paradox
and
Classical Cryptosystems

Asst. Prof. Dr. Serap Şahin

Izmir Institute of Technology

Shannon’s Theory

- In 1949, Claude Shannon published a paper entitled “Communication Theory of Secrecy Systems” in the Bell Systems Technical Journal.
- This paper had a great influence on the scientific study of cryptography.
- We will discuss several of Shannon’s ideas.

Perfect Secrecy

- There are two basic approaches to discussing the security of a cryptosystem.

1. **computational security** ; This measure concerns the computational effort required to break a cryptosystem.
   - We might define a cryptosystem to **be *computationally secure*** if the best algorithm for breaking it requires at least *N* operations, where *N* is some specified, very large number.
   - The problem is that **no known practical cryptosystem can be proved to be secure under this definition**.
   - Another approach is to provide evidence of computational security by **reducing the security of the cryptosystem to some well-studied problem** that is thought to be difficult. Cryptosystems of this type are sometimes termed "***provably secure,"*** but this approach only provides a proof of security relative to some other problem, not an absolute proof of security.

Perfect Secrecy

2. **unconditional security;** This measure concerns the security of cryptosystems when there is no bound placed on the amount of computation that Oscar is allowed to do. A cryptosystem is defined to be ***unconditionally secure*** if it cannot be broken, even **with infinite computational resources**.

The unconditional security of a cryptosystem obviously cannot be studied from the point of view of computational complexity, since **we allow computation time to be infinite**.

The appropriate framework in which to study unconditional security is **probability theory**. We need only elementary facts concerning probability; the main definitions are reviewed now.

Perfect Secrecy

DEFINITION 2.1 *Suppose* X *and* Y *are random variables. We denote the* ***probability*** *that* X *takes on the value x by* ***p(x)****, and the probability that* Y *takes on the value y by p(y).*

The ***joint probability p(x, y)*** *is the probability that* X *takes on the value x and* Y *takes on the value y.*

The ***conditional probability p(x|y)*** *denotes the probability that* X *takes on the value x given that* Y *takes on the value y.*

Perfect Secrecy

- *The **random variables*** X *and* Y *are said to be **independent** if **p(x, y) = p(x)p(y)** for all possible values x of* X *and y of* Y. **Joint probability can be related to conditional probability by the formula**

$$p(x, y) = p(x|y)p(y).$$

- *Interchanging x and y, we have that*

$$p(x, y) = p(y|x)p(x).$$

- *From these two expressions, we immediately obtain the following result, which is known as **Bayes' Theorem***.
- *THEOREM 2.1 (Bayes' Theorem) If p(y) > 0, then*

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)}.$$

Perfect Secrecy

- ***COROLLARY 2.2***

X *and* **Y** *are independent variables if and only if $p(x|y) = p(x)$ for all x, y.*

- ***Definition:*** **Posterior Probability** *is the conditional probability that is assigned after the relevant evidence is taken into account.*

Perfect Secrecy

- We assume that **a particular key is used for only one encryption**.
- Let us suppose that there is **a probability distribution on the plaintext space**, $\mathcal{P}$
- We denote the *a* ***priori*** **probability** that plaintext *x* occurs by $p_{\mathcal{P}}(x)$
- We assume that the key ***K*** is chosen (by Alice and Bob) using some fixed probability distribution (**often a key is chosen at random, so all keys will be equal probable**).
- Denote the probability that key ***K*** is chosen by $p_{\mathcal{K}}(K)$
- We make the reasonable assumption that the key **K** and the plaintext **x** are independent events.

Perfect Secrecy

- The two probability distributions on $\mathcal{P}$ and $\mathcal{K}$ induce a probability distribution on $\mathcal{C}$.
- Indeed, it is not hard to compute the probability $p_{\mathcal{C}}(y)$ that y is the ciphertext that is transmitted. For a key , $K \in \mathcal{K}$ define

$$C(K) = \{e_K(x) : x \in \mathcal{P}\}$$

- That is, ***C*(*K*)** represents the set of possible ciphertexts if ***K*** is the key. Then, for every $y \in \mathcal{C}$ we have that

$$p_{\mathcal{C}}(y) = \sum_{\{K : y \in C(K)\}} p_{\mathcal{K}}(K) p_{\mathcal{P}}(d_K(y))$$

Perfect Secrecy

- We also observe that, for any $y \in \mathcal{C}$ and $x \in \mathcal{P}$, compute the conditional probability, $p_{\mathcal{C}}(y|x)$ (i.e., the probability that *y* is the ciphertext, given that *x* is the plaintext) to be

$$p_{\mathcal{C}}(y|x) = \sum_{\{K:x=d_K(y)\}} p_{\mathcal{K}}(K)$$

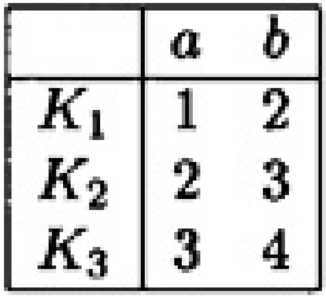
- It is now possible to compute the conditional probability $p_{\mathcal{P}}(x|y)$ (i.e., the probability that *x* is the plaintext, given that *y* is the ciphertext) using Bayes' Theorem. The following formula is obtained:

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)}. \quad \Rightarrow \quad p_{\mathcal{P}}(x|y) = \frac{p_{\mathcal{P}}(x) \displaystyle\sum_{\{K:x=d_K(y)\}} p_{\mathcal{K}}(K)}{\displaystyle\sum_{\{K:y\in C(K)\}} p_{\mathcal{K}}(K) p_{\mathcal{P}}(d_K(y))}$$

Perfect Secrecy

Example 1:

- Let $\mathcal{P} = \{a, b\}$ with $p_{\mathcal{P}}(a) = 1/4, p_{\mathcal{P}}(b) = 3/4$.
- Let $\mathcal{K} = \{K_1, K_2, K_3\}$ with $p_{\mathcal{K}}(K_1) = 1/2, p_{\mathcal{K}}(K_2) = p_{\mathcal{K}}(K_3) = 1/4$.
- Let $\mathcal{C} = \{1, 2, 3, 4\}$ and suppose the encryption functions are defined to be $e_{K_1}(a) = 1, e_{K_1}(b) = 2$; $e_{K_2}(a) = 2, e_{K_2}(b) = 3$; and $e_{K_3}(a) = 3, e_{K_3}(b) = 4$.
- This cryptosystem can be represented by the following ***encryption matrix***:



	a	b
K_1	1	2
K_2	2	3
K_3	3	4

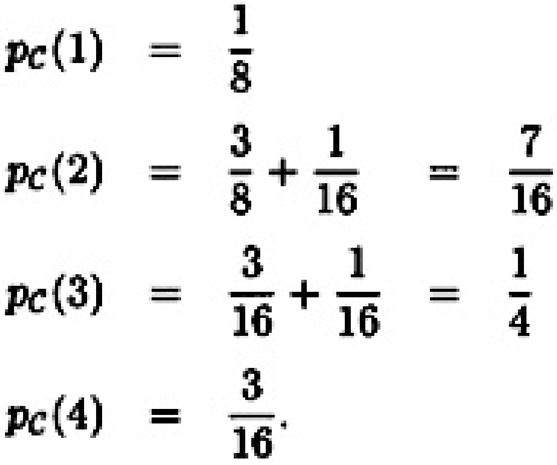
Perfect Secrecy

Example 1 :

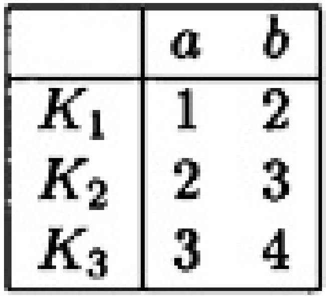
- We now compute the probability distribution . We obtain

Pc(1)=Pp(a).PK(K1)=1/4 . 1/2=1/8

Pc(2)=Pp(b).PK(K1) +Pp(a).PK(K2)= 3/4 . 1/2 + 1/4.1/4 =3/8 + 1/16 = 7/16



$$
\begin{aligned}
p_C(1) &= \frac{1}{8} \\
p_C(2) &= \frac{3}{8} + \frac{1}{16} = \frac{7}{16} \\
p_C(3) &= \frac{3}{16} + \frac{1}{16} = \frac{1}{4} \\
p_C(4) &= \frac{3}{16}.
\end{aligned}
$$



	a	b
K_1	1	2
K_2	2	3
K_3	3	4

Perfect Secrecy

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)}. \Rightarrow p_{\mathcal{P}}(x|y) = \frac{p_{\mathcal{P}}(x) \sum_{\{K:x=d_K(y)\}} p_{\mathcal{K}}(K)}{\sum_{\{K:y\in C(K)\}} p_{\mathcal{K}}(K)p_{\mathcal{P}}(d_K(y))}$$

Example 1 :

- Now we can compute the **conditional probability distributions on the plaintext**, given that a certain ciphertext has been observed. We have:

Pc(a|1)=(Pp(a).PK(K1))/(Pc(1))
=(1/4 . 1/2)/(1/8)=1

Pc(a|2)=(Pp(a).PK(K2))/(Pc(2))
=(1/4 . 1/4)/(7/16)=1/7

$$\begin{array}{lll lll}
p_{\mathcal{P}}(a|1) & = & 1 & p_{\mathcal{P}}(b|1) & = & 0 \\
p_{\mathcal{P}}(a|2) & = & \frac{1}{7} & p_{\mathcal{P}}(b|2) & = & \frac{6}{7} \\
p_{\mathcal{P}}(a|3) & = & \frac{1}{4} & p_{\mathcal{P}}(b|3) & = & \frac{3}{4} \\
p_{\mathcal{P}}(a|4) & = & 0 & p_{\mathcal{P}}(b|4) & = & 1.
\end{array}$$

	a	b
K_1	1	2
K_2	2	3
K_3	3	4

$\mathcal{P} = \{a, b\}$ with $p_{\mathcal{P}}(a) = 1/4, p_{\mathcal{P}}(b) = 3/4$.
$e_{K_1}(a) = 1, e_{K_1}(b) = 2;$ $e_{K_2}(a) = 2, e_{K_2}(b) = 3;$
$e_{K_3}(a) = 3, e_{K_3}(b) = 4$

Perfect Secrecy

- ***DEFINITION 2.2*** *A cryptosystem has perfect secrecy if* $p_{\mathcal{P}}(x|y) = p_{\mathcal{P}}(x)$ *for all* $x \in \mathcal{P}, y \in \mathcal{C}$. *That is, the a posteriori probability that the plaintext is x, given that the ciphertext y is observed, is identical to the a priori probability that the plaintext is x.*
- In the Example 1, the perfect secrecy property is satisfied for the ciphertext 3, but not for the other three ciphertexts.

$$\begin{aligned}
p_{\mathcal{P}}(a|1) &= 1 & p_{\mathcal{P}}(b|1) &= 0 \\
p_{\mathcal{P}}(a|2) &= \frac{1}{7} & p_{\mathcal{P}}(b|2) &= \frac{6}{7} \\
p_{\mathcal{P}}(a|3) &= \frac{1}{4} & p_{\mathcal{P}}(b|3) &= \frac{3}{4} \\
p_{\mathcal{P}}(a|4) &= 0 & p_{\mathcal{P}}(b|4) &= 1.
\end{aligned}$$

$\mathcal{P} = \{a, b\}$ with $p_{\mathcal{P}}(a) = 1/4, p_{\mathcal{P}}(b) = 3/4$

Perfect Secrecy

- The following theorem gives the formal statement and proof using probability distributions.
- ***THEOREM 2.3;*** *Suppose the* 26 *keys in the* **Shift Cipher** *are used with equal probability* 1/26. *Then for any plaintext probability distribution, the* **Shift Cipher** *has perfect secrecy*.
- **PROOF** Recall that $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26}$ and for $0 \leq K \leq 25$, the encryption rule e_K is $e_K(x) = x + K \bmod 26$ $(x \in \mathbb{Z}_{26})$.

First, we compute the distribution $p_\mathcal{C}$. Let $x \in \mathbb{Z}_{26}$; then

$$p_\mathcal{C}(y) = \sum_{K \in \mathbb{Z}_{26}} p_\mathcal{K}(K) p_\mathcal{P}(d_K(y))$$

$$= \sum_{K \in \mathbb{Z}_{26}} \frac{1}{26} p_\mathcal{P}(y - K) \quad = \frac{1}{26} \sum_{K \in \mathbb{Z}_{26}} p_\mathcal{P}(y - K).$$

Perfect Secrecy

- Now, for fixed *y*, the values *y* - *K* mod 26 comprise a permutation of $\mathbb{Z}_{26}$ and $p_{\mathcal{P}}$ is a probability distribution. Hence we have that

$$\sum_{K \in \mathbb{Z}_{26}} p_{\mathcal{P}}(y - K) = \sum_{y \in \mathbb{Z}_{26}} p_{\mathcal{P}}(y)$$

- Consequently,

$$= 1.$$

$p_{\mathcal{C}}(y) = \frac{1}{26}$ for any $y \in \mathbb{Z}_{26}$.

Next, we have that

$$p_{\mathcal{C}}(y|x) = p_{\mathcal{K}}(y - x \bmod 26)$$
$$= \frac{1}{26}$$

Perfect Secrecy

- For every *x, y*, since for every *x, y* the unique key *K* such that $e_K(x) = y$ is $K = y - x$ mod 26.
- Now, using Bayes' Theorem, it is trivial to compute

$$\begin{aligned} p_{\mathcal{P}}(x|y) &= \frac{p_{\mathcal{P}}(x)p_{\mathcal{C}}(y|x)}{p_{\mathcal{C}}(y)} \\ &= \frac{p_{\mathcal{P}}(x)\frac{1}{26}}{\frac{1}{26}} \\ &= p_{\mathcal{P}}(x), \end{aligned}$$

so we have perfect secrecy.

So, the **Shift Cipher** is "unbreakable" provided that a new random key is used to encrypt every plaintext character.

Perfect Secrecy

- ***THEOREM 2.4*** *Suppose* $(\mathcal{P},\mathcal{C},\mathcal{K},\mathcal{E},\mathcal{D})$ *is a cryptosystem where* $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{P}|$. *Then the cryptosystem provides perfect secrecy if and only if every key is used with equal probability* $1/|\mathcal{K}|$, *every* $x \in \mathcal{P}$, *and every* $y \in \mathcal{C}$, *there is a unique key K such that* $e_K(x) = y$.
- **PROOF** Suppose the given cryptosystem provides perfect secrecy. As observed above, for each $x \in \mathcal{P}$ and $y \in \mathcal{C}$, there must be at least one key K such that $e_K(x) = y$. So we have the inequalities:

$$|\mathcal{C}| = |\{e_K(x) : K \in \mathcal{K}\}| \le |\mathcal{K}|.$$

Perfect Secrecy

- But we are assuming that $|\mathcal{C}| = |\mathcal{K}|$. Hence, it must be the case that $|\{e_K(x) : K \in \mathcal{K}\}| = |\mathcal{K}|$.
- That is, there do not exist two distinct keys K_1 and K_2 such that

  $$e_{K_2}(x) = y$$
- Hence, we have shown that for any $x \in \mathcal{P}$ and $y \in \mathcal{C}$, there is exactly one key K such that $e_K(x) = y$.

Entropy

- We now want to look at what happens as more and more plaintexts are encrypted using the *same* key, and how likely a cryptanalyst will be able to carry out a successful ciphertext-only attack, given sufficient time.
- The basic tool in studying this question is the idea of **entropy**, a concept from information theory introduced by Shannon in 1948.
- **Entropy can be thought of as a mathematical measure of information or uncertainty, and is computed as a function of a probability distribution**.

Entropy

- Suppose we have a random variable **X** which takes on a finite set of values according to a probability distribution *p*(**X**).
- What is the information gained by an event which takes place according to distribution *p*(**X**)?
- Equivalently, if the event has not (yet) taken place, what is the uncertainty about the outcome?
- This quantity is called the entropy of **X** and is denoted by *H*(**X**).

Entropy

- Example 2; Suppose our random variable **X** represents the toss of a coin.

The probability distribution is ***p*(*heads*) = *p*(*tails*) = 1/2**.

It would seem reasonable to say that the information, or **entropy, of a coin toss is one bit**, since we could encode ***heads* by 1 and *tails* by 0,** for example.

In a similar fashion, **the entropy of *n* independent coin tosses is *n*, since the *n* coin tosses can be encoded by a bit string of length *n***.

Entropy

- Example 3; Suppose we have a random variable **X** that takes on three possible values x_1, x_2, x_3 with probabilities 1/2, 1/4, 1/4 respectively.
- The most efficient "encoding" of the three possible outcomes is to encode x_1, as 0, to encode x_2 as 10 and to encode x_3 as 11. Then the average number of bits in an encoding of **X** is

$$\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 2 = \frac{3}{2}$$

Entropy

- In these examples suggest that an event which occurs with probability **2^{-n}** can be encoded as a bit string of **length *n***.
- More generally, we could imagine that an event occurring with **probability *p*** might be encoded by **a bit string of length** approximately - **$\log_2 p$**.
- Given an arbitrary **probability distribution $p_1, p_2, \ldots, p_n$** for a random variable **X**, we take the weighted average of the quantities - **$\log_2 p_i$** to be our measure of information. This motivates the following formal definition.

Entropy

- ***DEFINITION 2.3*** *Suppose* **X** *is a random variable which takes on a finite set of values according to a probability distribution p*(**X**).

  Then, the entropy of this probability distribution is defined to be the quantity

$$H(\mathbf{X}) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

- *If the possible values of* **X** are x_i, 1 ≤ *i* ≤ *n, then we have*

$$H(\mathbf{X}) = -\sum_{i=1}^{n} p(\mathbf{X} = x_i) \log_2 p(\mathbf{X} = x_i)\,.$$

$$H(\mathbf{X}) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

- **REMARK** Observe that $\log_2 p_i$ is undefined if $p_i = 0$. Hence, entropy is sometimes defined to be the relevant sum over all the non-zero probabilities.
- when computing the entropy of a probability distribution p_i, the sum is taken over the indices i such that $p_i \neq 0$.
- Also, we note that the choice of **two** as the base of the logarithms is arbitrary: another base would only change the value of the entropy by a constant factor.
- Note that if $p_i = 1/n$ for $1 \leq i \leq n$, then $H(\mathbf{X}) = \log_2 n$. Also, it is easy to see that $H(\mathbf{X}) \geq 0$, and $H(\mathbf{X}) = 0$ if and only if $p_i = 1$ for some i.

Entropy

- Example 1 (continue); We compute as follows:

$\mathcal{P} = \{a, b\}$ with $p_{\mathcal{P}}(a) = 1/4, p_{\mathcal{P}}(b) = 3/4$.

$$\begin{aligned} H(\mathbf{P}) &= -\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4} \\ &= -\frac{1}{4}(-2) - \frac{3}{4}(\log_2 3 - 2) \\ &= 2 - \frac{3}{4}(\log_2 3) \\ &\approx 0.81. \end{aligned}$$

Similar calculations yield *H*(**K**) = 1.5 and $H(\mathbf{C}) \approx 1.85$

Entropy

- Example 4: Suppose there are four messages:

Message	Probability	Bits	example
x1	1/2	1	0
x2	1/4	2	10
x3	1/8	3	110
x4	1/8	3	111

- The entropy is: $H(X) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{8}\log_2\frac{1}{8}$

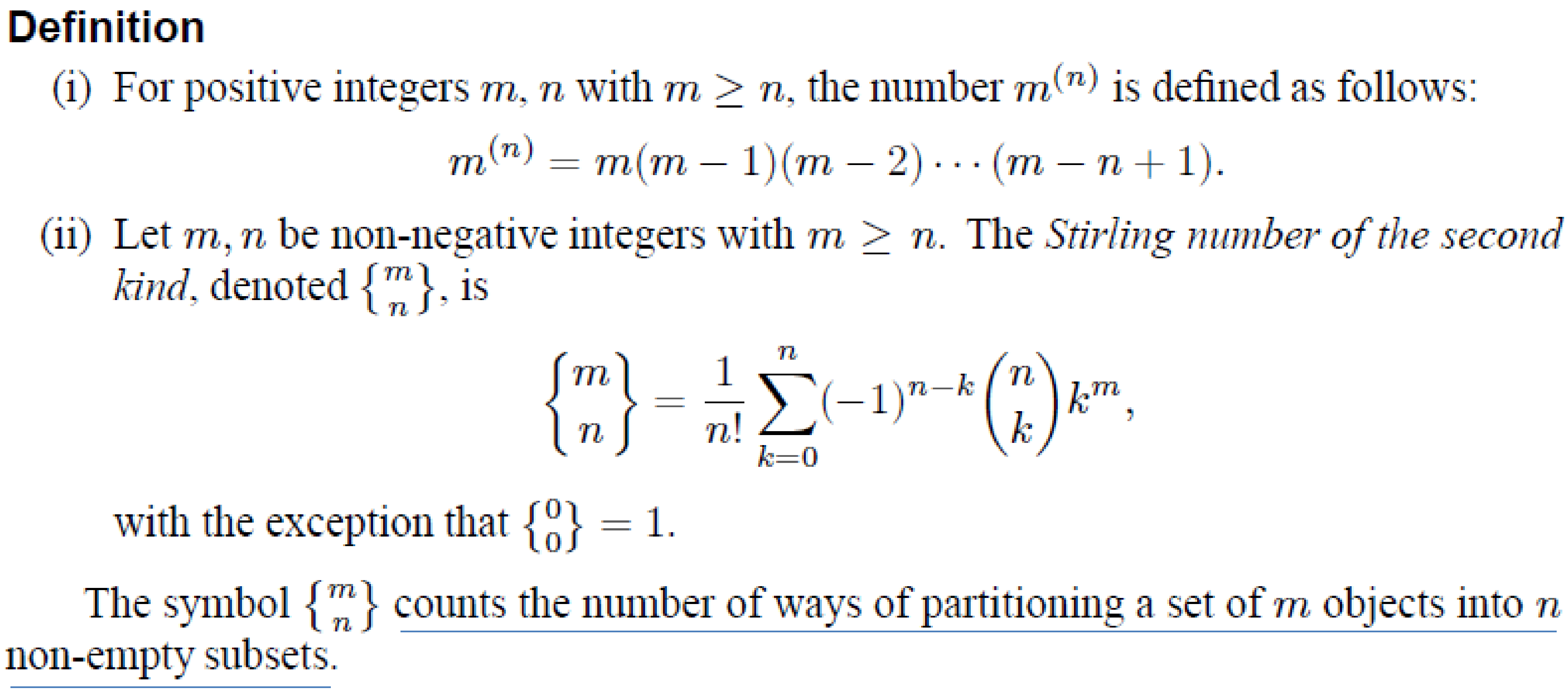
$$= \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} = \frac{7}{4} = 1.75$$

- The average length of the bit string to reveal which x_i was sent is;

$\frac{1}{2}(1) + \frac{1}{4}(2) + 2\frac{1}{8}(3) = \frac{1}{2} + \frac{5}{4} = \frac{7}{4} = 1.75$ the same H(X).

In cryptography, we measure the entropy of cipher text and keys, as well as of plaintext.

Birthday Problems



Definition

(i) For positive integers m, n with $m \geq n$, the number $m^{(n)}$ is defined as follows:

$$m^{(n)} = m(m-1)(m-2)\cdots(m-n+1).$$

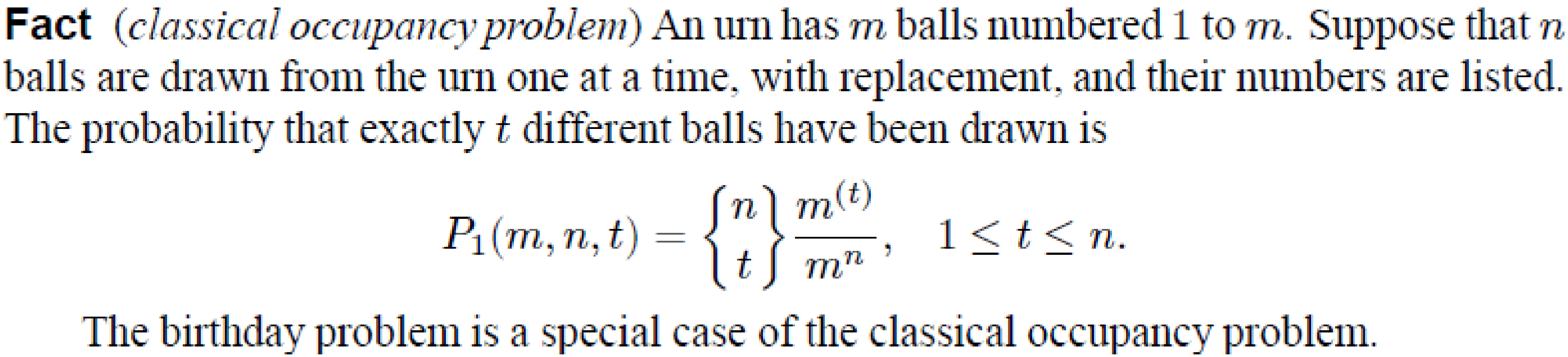
(ii) Let m, n be non-negative integers with $m \geq n$. The *Stirling number of the second kind*, denoted $\left\{ {m \atop n} \right\}$, is

$$\left\{ {m \atop n} \right\} = \frac{1}{n!}\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}k^m,$$

with the exception that $\left\{ {0 \atop 0} \right\} = 1$.

The symbol $\left\{ {m \atop n} \right\}$ counts the number of ways of partitioning a set of m objects into n non-empty subsets.

Birthday Problems

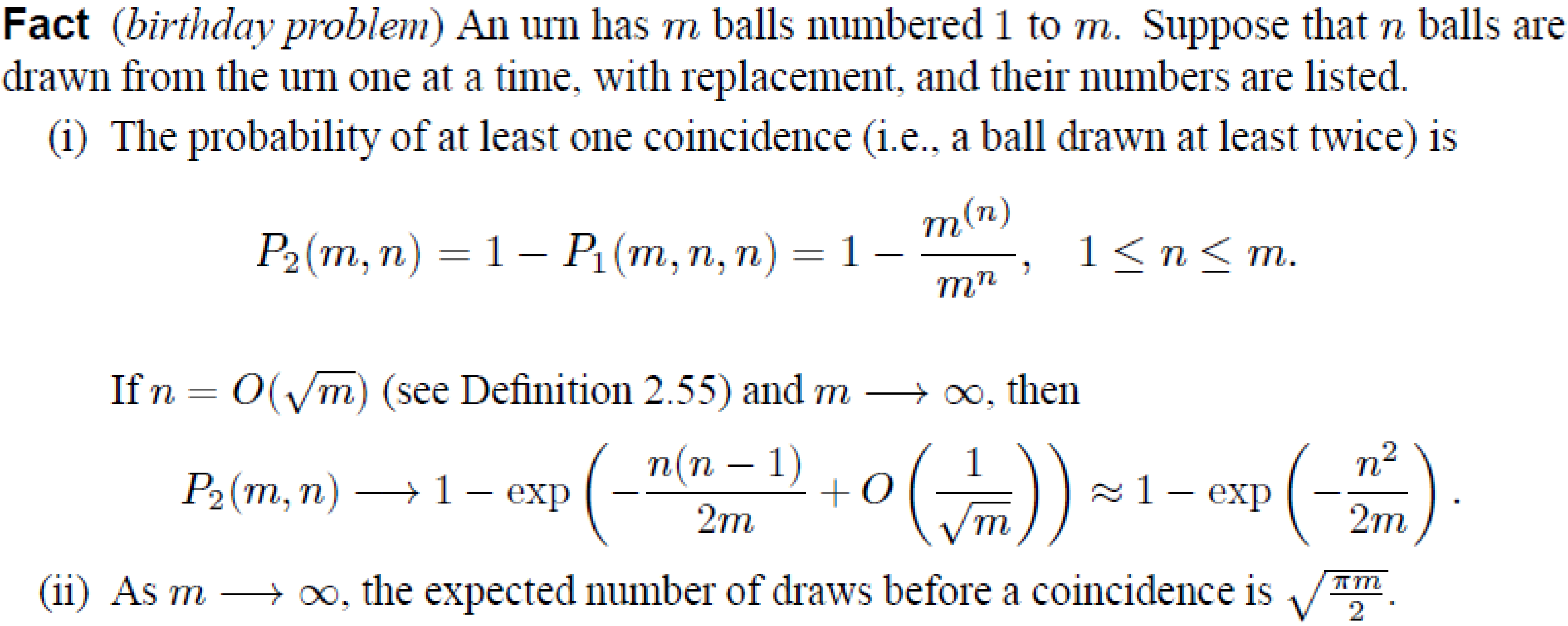


Fact (*classical occupancy problem*) An urn has m balls numbered 1 to m. Suppose that n balls are drawn from the urn one at a time, with replacement, and their numbers are listed. The probability that exactly t different balls have been drawn is

$$P_1(m,n,t) = \left\{ {n \atop t} \right\} \frac{m^{(t)}}{m^n}, \quad 1 \le t \le n.$$

The birthday problem is a special case of the classical occupancy problem.

Birthday Problems



Fact (*birthday problem*) An urn has m balls numbered 1 to m. Suppose that n balls are drawn from the urn one at a time, with replacement, and their numbers are listed.

(i) The probability of at least one coincidence (i.e., a ball drawn at least twice) is

$$P_2(m,n) = 1 - P_1(m,n,n) = 1 - \frac{m^{(n)}}{m^n}, \quad 1 \le n \le m.$$

If $n = O(\sqrt{m})$ (see Definition 2.55) and $m \longrightarrow \infty$, then

$$P_2(m,n) \longrightarrow 1 - \exp\left(-\frac{n(n-1)}{2m} + O\left(\frac{1}{\sqrt{m}}\right)\right) \approx 1 - \exp\left(-\frac{n^2}{2m}\right).$$

(ii) As $m \longrightarrow \infty$, the expected number of draws before a coincidence is $\sqrt{\frac{\pi m}{2}}$.

- The probability that at least 2 people in a room of 23 people have the same birthday is $P_2(365,23) \approx 0.507$, therefore it is referred to as the ***birthday surprise*** **or** ***birthday paradox*.**

Samples from Classical Cryptography

Basic Information

- Plaintext will be written in lowercase letters.
- CIPHERTEXT will be written in capital letters.
- The letters of alphabet are assigned numbers as follows:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Spaces and punctuation are omitted.

Shift Ciphers – Julius Caesar

- He shifted each letter by 3 places so *a* become *D* , *b* become *E*.

gaul is divided into three parts

⇩

JDXOLVGLYLGHGLQWRWKUHHSDUWV

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

q	r	s	t	u	v	w	x	y	z
16	17	18	19	20	21	22	23	24	25

Gaul; an ancient region of western Europe that included what is now northern Italy and France and Belgium and part of Germany and the Netherlands

Shift Ciphers – Julius Caesar

- Label the letters as integer from 0 to 25. The key is an integer K, with $0 \le K \le 25$.
- The encryption process is: $y \rightarrow x+K \pmod{26}$
- The decryption process is: $x \rightarrow y-K \pmod{26}$

Shift Ciphers – Julius Caesar

Types of attack work:

- **Ciphertext only**; Eve has only the cipher text
  - An exhaustive search, since there are only 26 possible keys.
  - If message is longer than a few letters; try to find some words of 4 or 5 letters that are shifts of each other.
  - If message is sufficiently long, is to do a frequency count for the various letters. The letter *e* occurs most frequently in English texts. Suppose the letter *L* appears most frequently in the ciphertext. Since *e=4* and *L=11*, a reasonable quest is that *K=11-4=7*.

Shift Ciphers – Julius Caesar

Types of attack work:

- **Known plaintext**; if you know just one letter of the plaintext along with the corresponding letter ciphertext, you can deduce the key. For instance; If you know $t(=19)$ encrypt to $D(=3)$ then the key $K \equiv 3-19 \equiv -16 \equiv 10 \pmod{26}$.
- **Chosen plaintext**; Chose the letter *a* as the plaintext. The ciphertext gives the key. For example if the ciphertext is *H* then the key is 7.
- **Chosen ciphertext**; Chose a letter *A* as the ciphertext. The plaintext is the negative of the key. For example, if the plaintext is *h,* the key is $-7 \equiv 19 \pmod{26}$.

Affine Ciphers

The shift ciphers may be generalized and strengthened as follows:

- Chose two integers α and β with $\gcd(\alpha, 26)=1$, and consider the function (called an affine function),

$$y \rightarrow \alpha.x+\beta \pmod{26}$$

For example; let $\alpha=9$ and $\beta=2$ so we are working $9.x+2$. Take a plaintext letter x; such as *h(=7).*

It is excrypted to $9.7+2\equiv 65\equiv 13 \pmod{26}$, which is letter *N* as y.

Affine Ciphers

- Using the same function;

  affine = CVVWPM

  How do we decrypt?

  $9.x+2=y \Rightarrow x=1/9\ .\ (y-2)\ (\text{mod } 26)$

  $x=9^{-1}.\ (y-2)\ (\text{mod } 26)$

  $x=3.\ (y-2)\ (\text{mod } 26)$

  $x=3y - 6\ (\text{mod } 26) \equiv 3y+20\ (\text{mod } 26)$

  so, the letter *V(=21)* is mapped to $3.21+20 \equiv 83 \equiv 5 (\text{mod } 26)$, which is letter *f*.

Affine Ciphers

Attacks;

- **Ciphertext only;** α should be gcd(α, 26)=1 therefore there are very few possibility for α. Hence, for an exhaustive search though all keys would take no longer than the corresponding search in the case of shift cipher, however it would be very easy to do on a computer.
- **Known plaintext;** with a luck, knowing two letters of the plaintext and corresponding letters of the ciphertext suffices to find the key.
  - For example; suppose the plaintext starts with *if* and corresponding ciphertext is *PQ.* In numbers, this means that *8(=i)* maps to *15=(P)* and *5* maps to *16*. Therefore we have the equations;

    $8.\alpha+\beta\equiv 15 \pmod{26}$

    $5.\alpha+\beta\equiv 16 \pmod{26}$, then subtracting yields $3.\alpha\equiv -1 \equiv 25 \pmod{26}$ and $\alpha=17$.

    Using the first equation $8.17+\beta\equiv 15 \pmod{26}$, which yields $\beta=9$.

The Vigenere Cipher

- The variation of the shift cipher was invented in the VI.century. Vigenere's encryption methods were more sophisticated. In 20th century, Babbage and Kasiski had shown how to attack it. In the 1920s, Friedman developed additional methods for breaking this and related ciphers.
- The **key** for the encryption is a **vector**.
  - First choose a **key length**. For example 6.
  - Then, choose a **vector** of this size whose entries are integers from 0 to 25. For example; key **k=(21, 4, 2, 19, 14, 17)**. Often the key corresponds a word. In our example this is ***vector***.
- **The security of the system depends on the fact that neither the keyword nor its length is known.**

The Vigenere Cipher

- To encrypt the message using ***k*** vector in our example; **k=(21, 4, 2, 19, 14, 17)**.
- We take first letter of the plaintext and shift by 21,
- Then, take the second letter of the plaintext and shift 4,
- Then, take the third letter of the plaintext and shift 2, and so on..

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

q	r	s	t	u	v	w	x	y	z
16	17	18	19	20	21	22	23	24	25

Plaintext	h	e	r	e	i	s	h	o	w	i	t	w	o	r	k	s
Key	21	4	2	19	14	17	21	4	2	19	14	17	21	4	2	19
Ciphertex	C	I	T	X	W	J	C	S	Y	B	H	N	J	V	M	L

The Vigenere Cipher

Attacks:

- **Known plaintext;** if enough characters are known since the key is simple obtained by the subtracting the plaintext from the ciphertext mod 26.
- **Chosen plaintext** or **chosen ciphertext**; this attack using the plaintext *aaaaaa*.. will yield the key immediately, while a chosen ciphertext attack with *AAAAAA*.. yields the negative of the key.
- **Frequency analysis;** the fact that in most English texts the frequencies of letters are not equal. These frequencies are tabulated by Beker and Piper.

a	b	c	d	e	f	g	h	i	j	k	l	m
.082	.015	.028	.043	.127	.022	.020	.061	.070	.002	.008	.040	.024

n	o	p	q	r	s	t	u	v	w	x	y	z
.067	.075	.019	.001	.060	.063	.091	.028	.010	.023	.001	.020	.001

The Vigenere Cipher

- **Frequency analysis;**
  - In this example the letter ***e*** appears as both ***I*** and ***X***. If the longer key is used in that case ***e*** can appears as more other letters.
  - In the mean time the ***J*** in the ciphertext is not come from only ***s*** or ***o***.
  - If the key length has enough length the frequency analysis is become hard.

Plaintext	h	e	r	e	i	s	h	o	w	i	t	w	o	r	k	s
Key	21	4	2	19	14	17	21	4	2	19	14	17	21	4	2	19
Ciphertex	C	I	T	X	W	J	C	S	Y	B	H	N	J	V	M	L

The Vigenere Cipher

- **Finding the key length;**
  - Write the ciphertext on a long strip of paper, and again on another long strip.
  - Put one strip above the other, but displaced by a certain number of places (the potential key length).
  - Make a * each time a letter and the one below it are the same, and count the total number of coincidences.

														*		
		V	V	H	Q	W	V	V	R	H	M	U	S	G	J	G
V	V	H	Q	W	V	V	R	H	M	U	S	G	J	G	T	H

T	H	K	I	H	T	S	S	E	J	C	H	L	S	F	C	B
K	I	H	T	S	S	E	J	C	H	L	S	F	C	B	G	V

				*												
G	V	W	C	R	L	R	Y	Q	T	F	S	V	G	A	H	..
W	C	R	L	R	Y	Q	T	F	S	V	G	A	H	W	K	..

We have two coincidencies. If we have continue to entire ciphertext, we would have counted 14 of them.

The Vigenere Cipher

- **Finding the key length;**

If we do this for different displacements we obtain the following data:

Displacement	1	2	3	4	5	6
Coincidences	14	14	16	14	24	12

We have the most coincidences for a shift of 5. This is the best guess for the length of the key. We will learn later, why this is.

The Vigenere Cipher

- **Finding the key: First method**

  Let suppose that the key length is 5.

  Let look at the 1st , 6th , 11th,... letters from the cyphertext and see which letter occurs most frequently:

A	B	C	D	E	F	G	H	I	J	K	L	M
0	0	7	1	1	2	9	0	1	8	8	0	0
N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
3	0	4	5	2	0	3	6	5	1	0	1	0

The most frequent is *G*, though *J,K,C* are close behind. Therefore we decide that *G=e* and the first element of the key is ***2=c***.

Then look at the 2nd , 7th, 12th,... letters and see that G occurs 10 times and S occurs 12 times and other letters are far behind. If G=e, then S=q, which should not occur 12 times in the plaintext. Therefore, S=e and the second element of the key is ***14=o***.

The Vigenere Cipher

- **Finding the key: First method**

  Then look at the 3rd , 8th, 13th ,… letters and see the frequencies.

  The initial guess is L=e, but there are problems such that R=k and E=x have too high. The best choice is *H=e* and third key element is ***3=d***.

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	0	3	3	1	3	5	1	0	4	10	0
N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
2	1	2	3	5	3	0	2	8	7	1	0	1

Then look at the 4th, 9th, 14th ,… letters and yields ***4=e*** as the fourth element of the key.

Finally, the 5th, 10th, and 15th,… letters and yields ***18=s*** as the fifth element of the key.

{2, 14, 3, 4, 18}={c, o, d, e, s}

The Vigenere Cipher

- Once a potential key found, test it by using it to decrypt.
- But, there is already exist an open question:
  - **Why the procedure given earlier finds the key length?**

The Vigenere Cipher

Why the procedure given earlier finds the key length?

- The letters are shift elements in vectors. This is define the encryption process.
- The displacement word is defined as slide one strip of paper to the right or left relative to the other strip.

1. Put the frequencies of English letters in a vector: A_0=(.082, .015, .028,…,.020,.001)
2. Let A_i be the result of shifting A_0 by i spaces to the right. A_2=(.020,.001,.082, .015, .028,…)
3. The product of A_0. A_0= $(.082)^2$+ $(.015)^2$ +…=.066
4. Of course A_i. A_i=.066, however A_i. A_j< .066, when i≠j, ranging from .031 to .045:

|i − j|	0	1	2	3	4	5	6	7	8	9	10	11	12	13
A_i. A_j	.066	.039	.032	.034	.044	.033	.036	.039	.034	.034	.038	.045	.039	.042

The Vigenere Cipher

Why the procedure given earlier finds the key length?

5. A_i. A_j= A_j. A_i so i – j and j – i give the same product so the product is only depends on |i – j|.
   – For instance; i – j = 17 corresponds to a shift 17 in one direction or 9 in other direction, so i – j=9 will give the same product.
6. The product of A_0. A_0 is higher than the other products is that the large numbers in the vectors are paired with large numbers and the small numbers are paired with small.
7. **Lets assume that the distribution of letters in the plaintext closely matches that of English, as expressed by the vector A_0.**

Substitution Ciphers

- This is so popular cryptosystem and commonly used in puzzle section on the weekend newspapers.
- Its principle is that; each letter in alphabet is replaced by another (or possibly the same) letter.
- A permutation of the alphabet is chosen and applied to the plaintext.
  - The Shift and Affine ciphers are examples of substitution cipher but the Vigenere and Hill ciphers are not, since they permute blocks of letters rather than one letter at a time.
- The whole substitution ciphers can be broken by frequency counts. However the process is complicated.

Substitution Ciphers

As an example; Thomas Jefferson wants to send a message to Ben Franklin. Clearly he does not want the British read the text if they intercept it, so he encrypts using a substitution cipher. Fortunately, Ben Franklin knows the permutation being used, so he can simply reverse permutation to obtain original message.

Let suppose that you are working for the Government Code and Cypher School in England in 1776, and the following message intercepted message to decrypt.

LWNSOZBNVWBAYBNVBSQWVWOHWDIZWBRBBNPBPOOUWRPAWXAW
PBWZWMYPOBNPBBNWJPAWWRZSLWZQJBNWIAXAWPBSALIBNXWA…

A frequency count yields the following; there are 520 letters in the text:

W	B	R	S	I	V	A	P	N	O	..
76	64	39	36	36	35	34	32	30	16	..

Substitution Ciphers

- The approximate frequency rate of most common letters in English text;

e	t	a	o	i	n	s	h	r
.127	.091	.082	.075	.070	.067	.063	.061	.060

W	B	R	S	I	V	A	P	N	O	..
76	64	39	36	36	35	34	32	30	16	..

- The letter W can represent e according to above frequency counts!
- But a simple frequency count is not enough to decide which is which.

Substitution Ciphers

- We used most frequent letters and analysis their counts as pairs in a counting diagram.
  - The entry 1 in the W row and N column means that the combination WN appears 1 time in the text.
- We have already decided that W=e and if we extended the table to include low-frequency letters, we would see that W contacts many of these letters, which is another characteristic of e.

- The vowels ***a, i, o*** tend to avoid each other.
- If you check the ***R row***, you see that ***R*** does not precede ***S,I, A, N*** very often.
- But a look at the ***R column*** shows ***R*** follows ***S,I, A*** fairly often.

So, *R* should not be one of the *a,i,o*.

Continuing, you should see that most likely possibilities for the ***a,i,o*** are ***S, I, P*** as vowels in some order.

Counting Diagram

	W	B	R	S	I	V	A	P	N
W	3	4	12	2	4	10	14	3	1
B	4	4	0	11	5	5	2	4	20
R	5	5	0	1	1	5	0	3	0
S	1	0	5	0	1	3	5	2	0
I	1	8	10	1	0	2	3	0	0
V	8	10	0	0	2	2	0	3	1
A	7	3	4	2	5	4	0	1	0
P	0	8	6	0	1	1	4	0	0
N	14	3	0	1	1	1	0	7	0

- The letter ***n*** has the property that around 80% of the letters that precede it are vowels. Since we have identified ***W, S, I, P*** as vowels, we see that ***R*** and ***A*** are the most likely candidates.
- The letter ***h* often appears before *e*** and rarely after it. This tell us that ***N=h***.
- The most common digram is ***th***, therefore ***B=t.***
- Other frequent letters are ***r*** and ***s*** remain, they should ***V*** and one of ***A, R***.
- Since ***r*** pairs more with vowels and ***s*** pairs more with consonants, we see that ***V*** must be ***s*** and ***r*** be represented by either ***A*** or ***R.***
- The combination ***rn*** should appear more than ***nr,*** and ***AR*** is more frequent than ***RA***, so the guess is that ***A=r*** and ***R=n***.

	W	B	R	S	I	V	A	P	N
W	3	4	12	2	4	10	14	3	1
B	4	4	0	11	5	5	2	4	20
R	5	5	0	1	1	5	0	3	0
S	1	0	5	0	1	3	5	2	0
I	1	8	10	1	0	2	3	0	0
V	8	10	0	0	2	2	0	3	1
A	7	3	4	2	5	4	0	1	0
P	0	8	6	0	1	1	4	0	0
N	**14**	3	0	1	1	1	0	7	0

e	**t**	**a**	**o**	**i**	**n**	**s**	**h**	**r**
.127	.091	.082	.075	.070	.067	.063	.061	.060

W	**B**	**R**	**S**	**I**	**V**	**A**	**P**	**N**	**O**	**..**
76	64	39	36	36	35	34	32	30	16	..

Substitution Ciphers

- The analysis is continue with ***S=o*** (note that ***to*** is much more common than ***ot***), ***I=i*** and ***P=a*** are the most likely choices.
- We have therefore determined reasonable guesses for 382 of the 520 characters in the text.

L	W	N	S	O	Z	B	N	W	V	W	B	A	Y	B	N	V	B	S
	e	h	o			t	h	e	s	e	t	r		t	h	s	t	o

Q	W	V	W	O	H	W	D	I	Z	W	R	B	B	N	P	B	P	..
	e	s	e			e		i		e	n	t	t	h	a	t	a	..

⇩

we hold these truths to be self evident that all men are created …