Another Look at Key Randomisation Hypotheses

Subhabrata Samajder and Palash Sarkar
Applied Statistics Unit
Indian Statistical Institute
203, B.T.Road, Kolkata, India - 700108.
{subhabrata.samajder@gmail.com, palash@isical.ac.in}

January 31, 2020

Abstract

In the context of linear cryptanalysis of block ciphers, let p_0 (resp. p_1) be the probability that a particular linear approximation holds for the right (resp. a wrong) key choice. The standard right key randomisation hypothesis states that p_0 is a constant $p \neq 1/2$ and the standard wrong key randomisation hypothesis states that $p_1 = 1/2$. Using these hypotheses, the success probability P_S of the attack can be expressed in terms of the data complexity N. The resulting expression for P_S is a monotone increasing function of N.

Building on earlier work by Daemen and Rijmen (2007), Bogdanov and Tischhauser (2014) argued that p_1 should be considered to be a random variable. They postulated the adjusted wrong key randomisation hypothesis which states that p_1 follows a normal distribution. A non-intuitive consequence was that the resulting expression for P_S is no longer a monotone increasing function of N. A later work by Blondeau and Nyberg (2017) argued that p_0 should also be considered to be a random variable and they postulated the adjusted right key randomisation hypothesis which states that p_0 follows a normal distribution.

In this work, we revisit the key randomisation hypotheses. While the argument that p_0 and p_1 should be considered to be random variables is indeed valid, we consider the modelling of their distributions by normal to be inappropriate. Being probabilities, the support of the distributions of p_0 and p_1 should be subsets of $[0, 1]$ which does not hold for normal distributions. We show that if p_0 and p_1 follow any distributions with supports which are subsets of $[0, 1]$, and $E[p_0] = p$ and $E[p_1] = 1/2$, then the expression for P_S that is obtained is exactly the same as the one obtained using the standard key randomisation hypotheses. Consequently, P_S is a monotone increasing function of N even when p_0 and p_1 are considered to be random variables.

Keywords: linear cryptanalysis, key randomisation hypotheses
MSC: 94A60, 11T71

1 Introduction

Linear cryptanalysis [8, 9] is one of the basic attacks against a block cipher. The first task in mounting a linear cryptanalysis is to obtain a linear approximation of the block cipher which holds with some probability p_0 which is significantly different from the probability p_1 that such a linear approximation holds for a uniform random permutation. Obtaining such a linear approximation is a non-trivial task and often requires a substantial amount of ingenuity. For the actual attack, the cryptanalysis algorithm requires as input plaintext-ciphertext pairs such that a single secret key was used to encrypt the plaintexts to obtain the corresponding ciphertexts. The output of the algorithm is a list of possible values of a subset of bits of the secret key. The attack is considered to be successful if the correct value is in the output list. Such an attack is called a key recovery attack. A weaker form of attack, called distinguishing attack, has also been considered in the literature.

Statistical techniques are used to assess the efficacy of linear cryptanalysis. The goal is to obtain a relation between the success probability P_S, the number N of plaintext-ciphertext pairs and the size of the output list which is expressed in terms of a parameter a called the advantage. The randomness arises from the distribution of the plaintexts which are n-bit strings. In this work, we consider the setting where the plaintexts are sampled uniformly at random with replacement. When the number of plaintexts is less than $2^{n/2}$, then due to the birthday bound, the obtained plaintexts are very likely to be distinct even if they are sampled with replacement.

Statistical analysis requires identifying a suitable test statistic T and obtaining its distributions under two settings: first, when the linear approximation holds for the block cipher, and, second, for a uniform random permutation. A fundamental requirement for obtaining the two distributions of T, and hence, for the overall statistical analysis, is to hypothesise the values of p_0 and p_1. The standard right key randomisation hypothesis postulates that p_0 is a constant p which is different from $1/2$ while the standard wrong key randomisation hypothesis postulates that $p_1 = 1/2$. Using these standard key randomisation hypotheses, the two distributions of T can be obtained and further statistical methods can be used to obtain an expression for P_S in terms of N and a. Further, it can be proved that P_S increases monotonically with N.

Bogdanov and Tischhauser [3] argued that it is not appropriate to consider $p_1 = 1/2$. Rather, p_1 should be considered to be a random variable. They based their arguments on an earlier result on distribution of correlations in uniform random permutations which was stated by O'Connor [10] and proved by Daemen and Rijmen [4] who had also given a normal approximation of the distribution. Based on the normal approximation, Bogdanov and Tischhauser [3] postulated the adjusted wrong key randomisation hypothesis that p_1 follows a normal distribution. Using the adjusted wrong key randomisation hypothesis and the standard right key randomisation hypothesis, an expression for P_S was obtained in terms of N and a. Counter-intuitively, P_S given by this expression does not increase monotonically with N. A later work by Blondeau and Nyberg [2] postulated the adjusted right key randomisation hypothesis that p_0 also follows a normal distribution and obtained an expression for P_S under both the adjusted right and wrong key randomisation hypotheses.

Our Contributions

Our starting point is the observation by Bogdanov and Tischhauser [3] that p_1 should be considered to be a random variable. However, unlike Bogdanov and Tischhauser, we do not use a normal distribution to model p_1. Note that p_1 is a probability and so it should take values only in the interval $[0, 1]$. Using a normal distribution to model p_1 does not ensure this. So, we model p_1 directly using the discrete distribution given in [10, 4] instead of using the normal approximation of this distribution. It is easily proved that the expectation of p_1 under this distribution is 1/2, i.e., $E[p_1] = 1/2$.

More generally, we consider the scenario where p_0 and p_1 are random variables having arbitrary distributions with the constraint that the supports of these distributions are subsets of $[0, 1]$ such that $E[p_0] = p$ and $E[p_1] = 1/2$. Under this setting, we show that the expression for P_S that is obtained is exactly the same as that obtained using the standard right and wrong key randomisation hypotheses. Consequently, P_S is an increasing function of N as is to be expected. The following are the consequences of our work.

1. Even if p_0 and p_1 are considered to be random variables, this has no effect on the relation between the success probability P_S, the data complexity N and the advantage a.

2. The adjusted wrong and right key randomisation hypotheses which postulate modelling p_1 and p_0 using normal distributions are unnecessary and lead to incorrect results.

3. The counter-intuitive behaviour of P_S decreasing with N reported in [3] is an effect of the adjusted wrong key hypothesis rather than being a model of reality.

Related Works

Since the seminal works of Matsui [8, 9], there has been a vast amount of work on linear cryptanalysis. We mention only the papers which are directly related to the present work. Selçuk [14] introduced the notion of advantage a of an attack and used a result on order statistics to obtain an expression for P_S in terms of N and a. The order statistics based approach built on Matsui's original key ranking method and the later work by Junod and Vaudenay [6] on optimal key ranking. Limitations of the order statistics based approach were pointed out in [11] and the alternative hypothesis testing based approach was used to derive the same expressions.

The standard wrong key randomisation hypothesis was formally stated by Harpes et al. [5]. As mentioned earlier, Bogdanov and Tischhauser [3] postulated the adjusted wrong key randomisation hypothesis while the adjusted right key randomisation hypothesis was postulated by Blondeau and Nyberg [2]. Both of these works as well as the present work assumes that the plaintexts are chosen under uniform random sampling with replacement. The setting of choosing plaintexts under uniform random sampling without replacement was considered by Ashur et al. [1]. They considered the adjusted wrong key randomisation and the standard right key randomisation to obtain an expression for P_S. It was pointed out in [13] that the modelling of p_0 and p_1 using normal distributions cannot be theoretically justified and heuristic explanations were forwarded. The work also carried out a comprehensive analysis of P_S under standard/adjusted right/wrong key randomisation hypotheses and both sampling with and without replacement.

Independent works by Matsui [9] and Kaliski and Robshaw [7] considered more than one linear approximations. A long line of works have analysed the setting where more than one linear approximation is available. The notion of adjusted key randomisation hypotheses has been carried over to multiple/multi-dimensional linear approximations [2, 12].

2 Overview of Linear Cryptanalysis of Block Ciphers

Consider a k-bit block cipher with n-bit blocks. The encryption function of such a block cipher is a map $E : \{0,1\}^k \times \{0,1\}^n \mapsto \{0,1\}^n$ where for each $K \in \{0,1\}^k$, we write $E_K(\cdot) = E(K,\cdot)$ and $E_K : \{0,1\}^n \to \{0,1\}^n$ is a bijection. Given a key K, a plaintext P, the ciphertext C is obtained as $C = E_K(P)$.

A key recovery attack on a block cipher is a probabilistic algorithm. The input to the algorithm is a set of N plaintext-ciphertext pairs (P_i, C_i), where $C_i = E_K(P_i)$, $i = 1,\ldots,N$. The goal of the attack is to recover m secret bits of the key. The output of the algorithm is a list of m-bit strings. The attack is successful if the correct m-bit string is in the output list. The probability of this event is called the success probability of the attack. Following [14], an attack is said to have advantage a if the size of the output list is equal to 2^{m-a}, i.e., a fraction 2^{-a} portion of the possible 2^m is produced as candidate keys. The number N of plaintext-ciphertext pairs provided as input to the algorithm is said to be the data complexity of the attack.

The encryption function of an iterated block cipher is obtained by composing round functions where each round function is a bijective map parameterised by a round key. The round keys are obtained by applying a key scheduling algorithm to the secret key K. Suppose the round keys are $k^{(0)}, k^{(1)}, \ldots$ and the round functions are $R^{(0)}_{k^{(0)}}, R^{(1)}_{k^{(1)}}, \ldots$. For $i \geq 1$, let $K^{(i)}$ denote the concatenation of the first i round keys and $E^{(i)}_{K^{(i)}}$ denote the composition of the first i round functions. Suppose the block cipher has $r+1$ rounds, i.e., $C = E^{(r+1)}_{K^{(r+1)}}(P)$, and by B the output after r rounds, i.e., $B = E^{(r)}_{K^{(r)}}(P)$ and $C = R^{(r)}_{k^{(r)}}(B)$. The basic task in the linear cryptanalysis of an iterated block cipher is to obtain a linear relation of the form

$$\langle \Gamma_P, P\rangle \oplus \langle \Gamma_B, B\rangle \quad = \quad \langle \Gamma_K, K^{(r)}\rangle. \tag{1}$$

where $\Gamma_P, \Gamma_B \in \{0,1\}^n$ and $\Gamma_{K^{(r)}} \in \{0,1\}^{nr}$ denote the plaintext mask, the mask to the input of the last round and the key mask respectively. A linear relation of the type (1) usually holds with some probability which is taken over the uniform random choice of the plaintext P.

To compute $\langle \Gamma_B, B\rangle$, the subset of the bits of B corresponding to the support of Γ_B is required. These bits of B are obtained from C by partially decrypting C by one round. This partial decryption of C involves a subset of the bits of the last round key $k^{(r)}$. This subset of bits of the last round key is said to be the target sub-key.

Let the size of the target sub-key be m. There are 2^m possible choices of the target sub-key out of which only one is correct. The purpose of the attack is to identify the correct value.

In the following, assume that K is the secret key used for encryptions. Once K is chosen, it is fixed and though it is secret, there is no randomness in K. So, $z = \langle \Gamma_K, K \rangle$ is an unknown, but, fixed bit. Let P be a plaintext chosen uniformly at random from $\{0,1\}^n$; C be the corresponding ciphertext; and B_κ be the result of partially decrypting C with a choice κ of the target sub-key. Define

$$\begin{array}{rcll} p_0 & = & \Pr[\langle \Gamma_P, P \rangle \oplus \langle \Gamma_B, B_\kappa \rangle = 1] & \text{if } \kappa = \kappa^*; \\ p_1 & = & \Pr[\langle \Gamma_P, P \rangle \oplus \langle \Gamma_B, B_\kappa \rangle = 1] & \text{if } \kappa \neq \kappa^*. \end{array} \tag{2}$$

Let $\epsilon_i = p_i - 1/2$, for $i = 0, 1$. Then ϵ_0 (resp. ϵ_1) is the bias corresponding to the correct (resp. incorrect) choice of the target sub-key.

Let $P_1, \ldots, P_N$, be chosen independently and uniformly at random from $\{0,1\}^n$. For $j = 1, \ldots, N$, let $C_j = E_K(P_j)$. For each choice κ of the target sub-key it is possible for the attacker to partially decrypt each C_j by one round to obtain $B_{\kappa,j}; j = 1, 2, \ldots, N$. For $\kappa \in \{0, 1, \ldots, 2^m - 1\}$, $z \in \{0,1\}$, $j = 1, \ldots, N$, define

$$\begin{array}{rcl} X_{\kappa,z,j} & = & \langle \Gamma_P, P_j \rangle \oplus \langle \Gamma_B, B_{\kappa,j} \rangle \oplus z; \\ X_{\kappa,z} & = & X_{\kappa,z,1} + \cdots + X_{\kappa,z,N}. \end{array}$$

Since each P_j is chosen uniformly at random from $\{0,1\}^n$, from (2), we have that for $j = 1, \ldots, N$,

$$\begin{array}{rcll} \Pr[X_{\kappa,0,j} = 1] & = & p_0 & \text{if } \kappa = \kappa^*; \\ \Pr[X_{\kappa,0,j} = 1] & = & p_1 & \text{if } \kappa \neq \kappa^*. \end{array} \tag{3}$$

From the definition of $X_{\kappa,z,j}$, we have $X_{\kappa,z,j} \oplus X_{\kappa,1\oplus z,j} = 1$, i.e, one of $X_{\kappa,z,j}$ and $X_{\kappa,1\oplus z,j}$ is 1 and the other is 0, so that $X_{\kappa,0} + X_{\kappa,1} = N$. For each choice κ of the target sub-key and each choice of z, define the test statistic

$$T_{\kappa,z} = |W_{\kappa,z}| \quad \text{where} \quad W_{\kappa,z} = \frac{X_{\kappa,z}}{N} - \frac{1}{2}.$$

Then

$$T_{\kappa,1} = |W_{\kappa,1}| = \left| \frac{X_{\kappa,1}}{N} - \frac{1}{2} \right| = \left| \frac{N - X_{\kappa,0}}{N} - \frac{1}{2} \right| = \left| \frac{1}{2} - \frac{X_{\kappa,0}}{N} \right| = |-W_{\kappa,0}| = T_{\kappa,0}.$$

So, the test statistic $T_{\kappa,z}$ does not depend on the value of z and it is sufficient to consider $z = 0$.

To simplify notation, we will write $X_{\kappa,j}$ and X_κ instead of $X_{\kappa,0,j}$ and $X_{\kappa,0}$ respectively; W_κ and T_κ instead of $W_{\kappa,0}$ and $T_{\kappa,0}$ respectively. In terms of this notation, we have the following.

1. From (3),

$$\begin{array}{rcll} \Pr[X_{\kappa,j} = 1] & = & p_0 & \text{if } \kappa = \kappa^*; \\ \Pr[X_{\kappa,j} = 1] & = & p_1 & \text{if } \kappa \neq \kappa^*. \end{array} \tag{4}$$

2. $X_\kappa = X_{\kappa,1} + \cdots + X_{\kappa,N}$ and

$$T_\kappa = |W_\kappa| \quad \text{where} \quad W_\kappa = \frac{X_\kappa}{N} - \frac{1}{2} = \frac{X_{\kappa,1} + \cdots + X_{\kappa,N}}{N} - \frac{1}{2}. \tag{5}$$

This test statistic was considered by Matsui [8].

There are 2^m choices of the target sub-key and so there are 2^m random variables T_κ. The distribution of T_κ is determined from the distribution of the $X_{\kappa,j}$'s and depends on whether κ is correct or incorrect.

The model of key recovery attack that we consider is based on statistical hypothesis testing. The attack proceeds as follows. For each possible value of the target sub-key κ, the cryptanalyst computes the value of the test statistic T_κ as given in (5). A statistical test of hypothesis with H_0: κ is correct, versus H_1: κ is incorrect is applied. The decision is based on comparing T_κ to an a priori determined threshold t. Based on the decision, κ is either retained as a candidate key or is rejected. We refer to [11, 13] for further details of the statistical hypothesis testing based approach.

In the above attack model, the statistical test is applied to each possible value of the target sub-key κ. Consequently, in the analysis of the statistical test, the value of κ is fixed and the randomness arises from the independent and uniform distribution of $P_1, \ldots, P_N$. *In particular, we note that there is no scope to consider κ to be a random variable.*

Before proceeding, we introduce notation on normal distributions. By $\mathfrak{N}(\mu, \sigma^2)$ we will denote the normal distribution with mean μ and variance σ^2. The density function of $\mathfrak{N}(\mu, \sigma^2)$ will be denoted by $\mathfrak{n}(x; \mu, \sigma^2)$. The density function of the standard normal will be denoted by $\phi(x)$ while the distribution function of the standard normal will be denoted by $\Phi(x)$.

Recall that $P_1, \ldots, P_N$ are chosen independently and uniformly at random from $\{0,1\}^n$ (i.e., uniform random sampling with replacement). So, for any κ, $X_{\kappa,1}, \ldots, X_{\kappa,N}$ are independent Bernoulli distributed random variables. From (4), for $\kappa \neq \kappa^*$, $X_{\kappa,j} \sim \mathsf{Ber}(p_1)$ and so $X_\kappa \sim \mathsf{Bin}(p_1, N)$. Similarly, from (4), $X_{\kappa^*,j} \sim \mathsf{Ber}(p_0)$ and so $X_{\kappa^*} \sim \mathsf{Bin}(p_0, N)$. Using the normal approximation for the binomial distribution, we have the following approximate distributions.

$$X_\kappa \sim \left\{ \begin{array}{ll} \mathfrak{N}(Np_0, Np_0(1-p_0)) & \text{if } \kappa = \kappa^*; \\ \mathfrak{N}(Np_1, Np_1(1-p_1)) & \text{if } \kappa \neq \kappa^*. \end{array} \right. \tag{6}$$

The distributions of W_κ and T_κ are obtained from the distribution of X_κ for both $\kappa = \kappa^*$ and $\kappa \neq \kappa^*$.

2.1 Standard Key Randomisation Hypothesis

For obtaining the distributions of X_{κ^*} and X_κ, $\kappa \neq \kappa^*$, it is required to hypothesise the behaviour of p_0 and p_1 respectively. The two standard key randomisation hypotheses are the following.

Standard right key randomisation hypothesis: $p_0 = p$, for some constant p.
Standard wrong key randomisation hypothesis: $p_1 = 1/2$.

The standard wrong key randomisation hypothesis was formally considered in [5], though it was used in earlier works.

Using $p_0 = p$ and $p_1 = 1/2$, from (6), the distribution of X_κ is obtained as follows.

$$X_\kappa \sim \left\{ \begin{array}{ll} \mathfrak{N}(Np, Np(1-p)) & \text{if } \kappa = \kappa^*; \\ \mathfrak{N}(N/2, N/4) & \text{if } \kappa \neq \kappa^*. \end{array} \right. \tag{7}$$

Given the above distribution of X_κ, previous analysis provides an expression for the success probability P_S in terms of the data complexity N and advantage a. Such an expression was first given in [14]. A later work [13] showed that the expression for P_S given in [14] is not complete and provided the following expression for P_S.

$$P_S = \Phi\left(2\sqrt{N}|\epsilon| - \gamma\right) + \Phi\left(-2\sqrt{N}|\epsilon| - \gamma\right). \tag{8}$$

where $\gamma = \Phi^{-1}\left(1 - \frac{2^{m-a-1}}{2^m-1}\right)$.

3 Adjusted Key Randomisation Hypotheses

The rationale for the wrong key randomisation hypothesis is that if the choice κ is wrong, then the block cipher is assumed to behave like a uniform random permutation of $\{0,1\}^n$. The choice of $p_1 = 1/2$ is supposed to reflect this behaviour. Bogdanov and Tischhauser [3] were the first to suggest that the standard wrong key randomisation hypothesis is not proper. Their reasoning was based on an earlier work on distribution of correlations for a uniform random permutation of $\{0,1\}^n$. This distribution was stated by O'Connor [10] and proved by Daemen and Rijmen [4]. The crux of the result on the distribution of correlations for uniform random permutation is that p_1 is not a constant. Rather, it follows the following discrete probability distribution. For integer $x \in \{0,\ldots,2^{n-1}\}$,

$$\Pr\left[p_1 = 1 - \frac{x}{2^{n-1}}\right] \quad = \quad \frac{\binom{2^{n-1}}{x}^2}{\binom{2^n}{2^{n-1}}}. \tag{9}$$

Using a normal approximation of the distribution in (9) given in [4], the following was formally stated in [3].

Adjusted wrong key randomisation hypothesis:

$$\epsilon_1 \sim \mathfrak{N}\left(0, 2^{-n-2}\right), \text{ or, equivalently } p_1 \sim \mathfrak{N}\left(1/2, 2^{-n-2}\right).$$

From (6), for $\kappa \neq \kappa^*$, $X_\kappa \sim \mathfrak{N}(Np_1, Np_1(1-p_1))$ where under the adjusted wrong key randomisation hypothesis $p_1 \sim \mathfrak{N}\left(1/2, 2^{-n-2}\right)$. The standard result on compound of a normal distribution with another normal distribution is the following. If $Y_1 \sim \mathfrak{N}(aY_2, \sigma_1^2)$ and $Y_2 \sim \mathfrak{N}(\mu, \sigma_2^2)$ for constants σ_1 and σ_2, then $Y_1 \sim \mathfrak{N}(a\mu, \sigma_1^2 + a^2\sigma_2^2)$. Since, the variance of X_κ depends on p_1 (and hence is not a constant), this result on compound of normal distributions does not apply to X_κ and p_1. If, however, we make the approximation that $Np_1(1-p_1) \approx N/4$, then the result on compound of normal distributions applies and we obtain the following approximate distribution of X_κ.

$$X_\kappa \quad \sim \quad \mathfrak{N}(N/2, N^2(1/(4N) + 1/2^{n+2})) \quad \text{for } \kappa \neq \kappa^*. \tag{10}$$

The distribution of X_κ provides the distribution of W_κ and T_κ for $\kappa \neq \kappa^*$. Using the adjusted wrong key randomisation hypothesis along with the standard right key randomisation hypothesis and applying the techniques from [14], an expression for P_S was obtained in [3]. Somewhat surprisingly, it was shown that P_S is not monotone increasing with N, i.e., there is a range of values of N such that as N increases in this range, the value of P_S goes down. This is unintuitive since as the number of plaintext-ciphertext pairs increases, the success probability should not go down. Explanations were provided in [3] to justify why such a situation may indeed arise.

A later work [2], introduced a modification of the standard right key randomisation hypothesis as follows.

Adjusted right key randomisation hypothesis:

$$\epsilon_0 \sim \mathfrak{N}\left(\epsilon, \frac{\mathsf{ELP}-4\epsilon^2}{4}\right), \text{ or, equivalently } p_0 \sim \mathfrak{N}\left(p, \frac{\mathsf{ELP}-4\epsilon^2}{4}\right), \text{ where } \epsilon = p - 1/2 \text{ and } \mathsf{ELP} \geq 4\epsilon^2.$$

From (6), $X_\kappa \sim \mathfrak{N}(Np_0, Np_0(1-p_0))$ where under the adjusted right key randomisation hypothesis $p_0 \sim \mathfrak{N}\left(p, \frac{\mathsf{ELP}-4\epsilon^2}{4}\right)$. Again, approximating $p_0(1-p_0)$ by $1/4$ and applying the result on compound of normal distributions, we obtain the following approximate distribution of X_{κ^*}.

$$X_{\kappa^*} \quad \sim \quad \mathfrak{N}(Np, (N^2/4)(1/N + \mathsf{ELP} - 4\epsilon^2)). \tag{11}$$

Remark: The quantities p_0 and p_1 are probabilities and so cannot take values outside $[0,1]$. However, the assumption that these quantities follow normal distributions allows them to take values outside $[0,1]$. So, the normality assumption on the distributions of p_0 and p_1 are heuristics and cannot be theoretically justified. This has been pointed out in [13].

4 Distributions for Wrong and Right Key Choices Revisited

The random variables $X_{\kappa,1},\ldots,X_{\kappa,N}$ are independent Bernoulli distributed random variables. The probability of success is p_0 if $\kappa=\kappa^*$ and is p_1 if $\kappa\neq\kappa^*$. The motivation behind the formulating the adjusted key randomisation hypotheses is that p_0 and p_1 are themselves random variables instead of being constant values. We follow this motivation. Our point of departure from the adjusted key randomisation hypotheses is that we do not consider approximate normal distributions for p_0 and p_1. Being probabilities, the supports of the distributions of p_0 and p_1 must be subsets of $[0,1]$, i.e., these variables take values outside $[0,1]$ with probability 0. To analyse such a scenario we start with the following simple result.

Proposition 1. *Let X and Q be random variables such that $X\sim\mathsf{Ber}(Q)$ and Q follows a distribution whose support is a subset of $[0,1]$ and $E[Q]=\mu$. Then $X\sim\mathsf{Ber}(\mu)$.*

Proof. First suppose that Q follows a discrete distribution taking values $q_0,\ldots,q_\ell\in[0,1]$. Then

$$\begin{aligned}\Pr[X=1] &= \sum_{i=0}^{\ell}\Pr[X=1\wedge Q=q_i]\\ &= \sum_{i=0}^{\ell}\Pr[X=1\mid Q=q_i]\Pr[Q=q_i]\\ &= \sum_{i=0}^{\ell}q_i\Pr[Q=q_i]\\ &= E[Q]=\mu.\end{aligned}$$

The case when Q follows a continuous distribution is tackled in a similar fashion by considering the density function of Q and changing the sum to integral. □

Distribution under wrong key choice: We apply Proposition 1 to analyse the scenario arising from a wrong key choice. For the analysis, we directly apply the distribution given by (9) instead of the normal approximation of the distribution formulated in the adjusted wrong key randomisation hypothesis. For $\kappa\neq\kappa^*$, the random variables $X_{\kappa,1},\ldots,X_{\kappa,N}$ are independent $\mathsf{Ber}(p_1)$ distributed random variables, where p_1 follows the distribution given by (9). We compute $\mu=E[p_1]$ as follows.

$$\begin{aligned}\mu=E[p_1] &= \sum_{x=0}^{2^{n-1}}\left(1-\frac{x}{2^{n-1}}\right)\frac{\binom{2^{n-1}}{x}^2}{\binom{2^n}{2^{n-1}}}\\ &= 1-\frac{1}{2^{n-1}}\sum_{x=0}^{2^{n-1}}x\frac{\binom{2^{n-1}}{x}^2}{\binom{2^n}{2^{n-1}}}\\ &= 1-\frac{1}{2^{n-1}}\cdot\frac{1}{2}\cdot 2^{n-1}\\ &= \frac{1}{2}.\end{aligned}\tag{12}$$

Combining (12) with Proposition 1, we have that the random variables $X_{\kappa,1},\ldots,X_{\kappa,N}$ are independent $\mathsf{Ber}(1/2)$ distributed random variables. Consequently, $X_\kappa\sim\mathsf{Bin}(N,1/2)$ and the approximate normal distribution of X_κ is given by (7). This is exactly the situation which arises out of considering the standard wrong key randomisation hypothesis. So, the distribution of X_κ remains the same irrespective of whether we assume $p_1=1/2$ or p_1 follows a distribution over $[0,1]$ such that $E[p_1]=1/2$.

The approximate distribution of X_κ given by (10) arises due to the following. The approximate normal distribution $\mathfrak{N}(Np_1, Np_1(1-p_1))$ of X_κ given by (6) is heuristically compounded with the approximate normal distribution $\mathfrak{N}(1/2, 2^{-n-2})$ of p_1 given by the adjusted wrong key randomisation hypothesis to obtain the heuristic and approximate normal distribution $\mathfrak{N}(N/2, N^2(1/(4N)+1/2^{n+2}))$ of X_κ given by (10). As seen above, the heuristic considerations are not only theoretically unjustified, they lead to an incorrect distribution for X_κ.

Distribution under right key: Suppose p_0 follows a distribution whose support is a subset of $[0,1]$ and $E[p_0] = p$. Then using Proposition 1, $X_{\kappa^*,j} \sim \mathsf{Ber}(p)$ for $j = 1,\ldots,N$. Using the independence of $X_{\kappa^*,1},\ldots,X_{\kappa^*,N}$, we obtain that $X_{\kappa^*} \sim \mathsf{Bin}(Np, Np(1-p))$. Consequently, the approximate normal distribution of X_{κ^*} is given by (7). Again, this is exactly the situation which arises out of considering the standard right key randomisation hypothesis.

The adjusted right key randomisation hypothesis assumes that p_0 follows a normal distribution. This requires heuristic considerations to compound with the approximate normal distribution $\mathfrak{N}(Np_0, Np_0(1-p_0))$ of X_{κ^*} resulting in the heuristic and approximate normal distribution of X_{κ^*} given by (11). As in the case of distribution under wrong key choice, such heuristic considerations are unnecessary and lead to an incorrect distribution for X_κ.

Success probability: From the above analysis, considering p_0 and p_1 to be random variables (with supports which are subsets of $[0,1]$) leads to the same distribution of X_κ, $\kappa \neq \kappa^*$ and X_{κ^*} as is obtained under the standard key randomisation hypotheses. Consequently, the expression for P_S obtained from these distributions also remain the same as those obtained in the case of the standard key randomisation hypotheses and is given by (8). It has been proved in [13], that the expression for P_S given by (8) is monotone increasing with N for all N. So, the non-intuitive behaviour of the success probability not being monotone increasing with N as reported in [3] is really an outcome of the heuristic considerations involved in the adjusted key randomisation hypotheses rather than being a model of the real world.

5 Conclusion

We have shown that in the setting where the plaintexts are sampled uniformly at random with replacement, the relation between P_S, N and a remain the same as that obtained under the standard key randomisation hypotheses even if p_0 and p_1 are considered to be random variables whose support is a subset of $[0,1]$. Consequently, it follows that the adjusted key randomisation hypotheses are unnecessary and lead to incorrect results.

The adjusted wrong and right key randomisation hypotheses have been used to analyse the situation where plaintexts are sampled without replacement. One possible future work is to extend the results of the present work to cover this situation. This, however, does not seem to be easy. It requires getting a tractable form of the distribution obtained by compounding a hypergeometric distribution with the distribution for p_1 given by (9).

The adjusted key randomisation hypotheses have also been used in the context of multiple/multi-dimensional linear cryptanalysis which models the probability vectors with multi-variate normal distributions. This again is problematic since then the probabilities are allowed to take values outside $[0,1]$. So, another direction of research is to do away with the multi-variate normal distributions postulated by the adjusted key randomisation hypotheses for multiple/multi-dimensional linear cryptanalysis and instead work directly with the correct distributions of the probability vectors. This also presents significant technical difficulties.

References

[1] Tomer Ashur, Tim Beyne, and Vincent Rijmen. Revisiting the wrong-key-randomization hypothesis. *IACR Cryptology ePrint Archive*, 2016:990, 2016.

[2] Céline Blondeau and Kaisa Nyberg. Joint data and key distribution of simple, multiple, and multidimensional linear cryptanalysis test statistic and its impact to data complexity. *Des. Codes Cryptography*, 82(1-2):319–349, 2017.

[3] Andrey Bogdanov and Elmar Tischhauser. On the Wrong Key Randomisation and Key Equivalence Hypotheses in Matsui's Algorithm 2. In *Fast Software Encryption*, pages 19–38. Springer, 2014.

[4] Joan Daemen and Vincent Rijmen. Probability Distributions of Correlation and Differentials in Block Ciphers. *Journal of Mathematical Cryptology JMC*, 1(3):221–242, 2007.

[5] Carlo Harpes, Gerhard G. Kramer, and James L. Massey. A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma. In *Advances in Cryptology - EUROCRYPT '95, International Conference on the Theory and Application of Cryptographic Techniques, Saint-Malo, France, May 21-25, 1995, Proceeding*, pages 24–38, 1995. `http://link.springer.de/link/service/series/0558/bibs/0921/09210024.htm`.

[6] Pascal Junod and Serge Vaudenay. Optimal Key Ranking Procedures in a Statistical Cryptanalysis. In *Fast Software Encryption*, pages 235–246. Springer, 2003.

[7] Burton S Kaliski Jr and Matthew JB Robshaw. Linear Cryptanalysis Using Multiple Approximations. In *Advances in Cryptology–Crypto'94*, pages 26–39. Springer, 1994.

[8] Mitsuru Matsui. Linear Cryptanalysis Method for DES Cipher. In *Advances in Cryptology–EUROCRYPT'93*, pages 386–397. Springer, 1993.

[9] Mitsuru Matsui. The First Experimental Cryptanalysis of the Data Encryption Standard. In Y. G. Desmedt, editor, *Advances in Cryptology–Crypto'94*, pages 1–11. Springer, 1994.

[10] Luke O'Connor. Properties of linear approximation tables. In Bart Preneel, editor, *Fast Software Encryption: Second International Workshop. Leuven, Belgium, 14-16 December 1994, Proceedings*, volume 1008 of *Lecture Notes in Computer Science*, pages 131–136. Springer, 1994.

[11] Subhabrata Samajder and Palash Sarkar. Another look at normal approximations in cryptanalysis. *J. Mathematical Cryptology*, 10(2):69–99, 2016.

[12] Subhabrata Samajder and Palash Sarkar. Success probability of multiple/multidimensional linear cryptanalysis under general key randomisation hypotheses. *Cryptography and Communications*, 10(5):835–879, 2018.

[13] Subhabrata Samajder and Palash Sarkar. Another look at success probability of linear cryptanalysis. *Adv. in Math. of Comm.*, 13(4):645–688, 2019.

[14] Ali Aydın Selçuk. On Probability of Success in Linear and Differential Cryptanalysis. *Journal of Cryptology*, 21(1):131–147, 2008.