ÉCOLE POLYTHECHNIQUE FÉDÉRALE DE LAUSANNE

Statistical testing of the Sloth function

Laboratory for cryptologic algorithms
Bachelor semester project

Fall 2020

N.Kaluderovic, A.Lenstra

Emna Fendri
School of Computer and Communication Sciences
Communication Systems

January 28, 2021

Contents

1 Introduction **3**

2 Definitions **4**
2.1 The Chi-square distribution 4
2.2 The Chi-square statistic 4

3 Chi-square test of independence applied to Sloth **5**
3.1 Defining random variables and reporting data 5
3.2 Hypothesis testing 6
3.3 Proof of Pearson's theorem 6
3.4 Sample data analysis 7
3.4.1 Decision procedure 7
3.4.2 Measure of evidence using p-value 8

4 Test results **9**
4.1 Modulo 4 10
4.2 Modulo 16 11
4.3 Modulo 32 12
4.4 Modulo 128 13

5 Conclusion **14**

6 References **15**

1 Introduction

Sloth is a slow-timed hash function where the resulting hash can be verified quickly.
Sloth takes as input a *seed* and a large *prime number*, and is used for uncontestable random number generation, where the resulting random number is referred to as a *beacon*.
This project consists in testing the existence of a correlation between **the random value produced with a given seed** and **the random value produced with this same seed with one bit flipped**, with all other parameters unchanged.
For this, we realise a Chi-square test of independence based on sample data.

2 Definitions

2.1 The Chi-square distribution

The Chi-square distribution with d degrees of freedom, denoted by χ^2_d, is the distribution of a sum of the squares of d mutually independent standard normal random variables.
We write :

$$\chi^2_d = Z_1^2 + Z_2^2 + Z_3^2 + ... + Z_d^2$$

where $Z_1, ..., Z_d \overset{iid}{\sim} N(0,1)$ and $d \in \mathbb{N}_{>0}$.
We denote by μ and σ^2, the mean and the variance respectively. They are expressed as,

$$\mu = E(\chi^2_d) = E(\sum_{i=1}^{d} Z_i^2) = \sum_{i=1}^{d} E(Z_i^2) = d,$$

by linearity of expectation.

$$\sigma^2 = \text{var}(\sum_{i=1}^{d} Z_i^2) = \sum_{i=1}^{d} \text{var}(Z_i^2) = d\text{var}(Z_i^2) = d(E(Z_i^4) - E(Z_i^2)^2) = d(3-1) = 2d,$$

using that all Z_i are independent and identically distributed and that $E(Z_i^4) = 3$.

The chi-square distribution is widely used in inferential statistics, and in particular, in hypothesis testing. A test is described as Chi-square if the resulting test statistic obtained follows a Chi-square distribution with d degrees of freedom under the null hypothesis. This Chi-square test statistic is also known as the Pearson statistic and is defined as follows.

2.2 The Chi-square statistic

Let $O_1, ..., O_d$ be the number of observations of a random sample of size $N = n_1 + ... + n_d$, falling into the disjoint categories $1, ..., d$ whose expected values are $E_1, ..., E_d$ where $E_i > 0$.
Then the **Chi-square statistic** is

$$T = \sum_{i=1}^{d} \frac{(O_i - E_i)^2}{E_i}.$$

3 Chi-square test of independence applied to Sloth

Let $g(s)$ be the output of *sloth* for the seed s.
We first randomly select the public prime number p of 2048 bits.
For each observation, we randomly select an input $s \in \{0, ..., p-1\}$ and define s' obtained by flipping one bit in the sequence s and keeping all other digits identical.
We obtain a pair (b_0,b_1), where $b_0 = g(s)$ and $b_1 = g(s')$, as shown in Figure 1 below.
For instance, suppose s=10100, for which the generator outputs $g(s) = 13 = b_0$.
Now, take s'=10110 (2nd bit is flipped) for which the generator outputs $g(s') = 2 = b_1$
We obtain the pair $(13, 2)$.

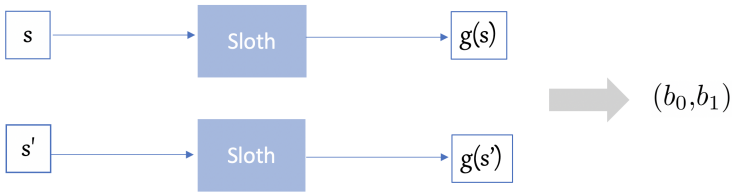


Figure 1: Illustration of 1 observation

3.1 Defining random variables and reporting data

Consider the congruence classes of integers modulo n : $[0]_n, [1]_n, ..., [n-1]_n$.
These are all the possible remainders of the division by n.

→ For inputs s and s', we define the **random variables**:
X: $[i]_n$ such that $b_0 \in [i]_n$.
Y: $[j]_n$ such that $b_1 \in [j]_n$.
$X, Y \in \{[0]_n, [1]_n, ...[n-1]_n\}$.

After gathering the observations, data is expressed in the following contingency table.

We examine the variables X and Y. The variable X represents the rows of the table, so that the rows are labelled: $[0]_n, ..., [n-1]_n$. The variable Y represents the columns of the table so that the columns are labelled: $[0]_n, ..., [n-1]_n$. Let the observed value be O_{ij} that represents the observed number of cases in row i and column j, that is, observed number of cases where X took value $[i]_n$ and Y took the value $[j]_n$, we denote this event by $[i]_n[j]_n$.
(i.e. $b_0 \in [i]_n$ and $b_1 \in [j]_n$ when considering the pair (b_0, b_1)).
Let the row totals be R_i for row i and the column totals C_j for column j, and N,the sample size.

	Variable Y					
Variable X	$[0]_n$	$[1]_n$	$[2]_n$	...	$[n-1]_n$	**Row Total** (R_i)
$[0]_n$	O_{11}	O_{12}	O_{13}	...	O_{1n}	R_1
$[1]_n$	O_{21}	O_{22}	O_{23}	...	O_{2n}	R_2
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
$[n-1]_n$	O_{n1}	O_{n2}	O_{n3}	...	O_{nn}	R_n
Column Total (C_j)	C_1	C_2	C_3	...	C_n	N

3.2 Hypothesis testing

After defining the random variables, we are able to state the null and alternative hypotheses. The null hypothesis is that variables X and Y are independent of each other. The alternative hypothesis is that X and Y are dependent variables.
We state it as :
H_0: X and Y are independent.
H_1: X and Y are dependent.

If the null hypothesis H_0 is true, then by **Pearson Theorem**, the Pearson statistic converges in distribution to the Chi square distribution with d degrees of freedom :

$$T = \sum_{i,j=1}^{n} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \to \chi^2_d,$$

where O_{ij} is the observed number of cases where X took value $[i]_n$ and Y took the value $[j]_n$. The degree of freedom d is $(r-1) \times (c-1)$, where r and c denote the number of rows and columns respectively, in this case they are both equal to n.
E_{ij} is the expected number of observations in row i and column j. The expected values are computed on the basis of the null hypothesis. If variables X and Y are independent of each other, let the probability that $X = X_i = [i]_n$ and $Y = Y_j = [j]_n$ be P_{ij},

$$P_{ij} = P(X_i \cap Y_j) = P(X_i)P(Y_j) = \frac{R_i}{N} \times \frac{C_j}{N}.$$

That is, P_{ij} is the probability of falling in a case which is in row i and column j, under the independence assumption. Thus the expected number of cases in row i and column j is:

$$E_{ij} = P_{ij} \times N = \frac{R_i \times C_j}{N}.$$

3.3 Proof of Pearson's theorem

Each O_{ij} represents the number of time the event $[i]_n[j]_n$ occurred. So, if we consider that the joint distribution $O_{11}, ..., O_{nn}$ is multinomial with denominator N and probabilities $p_{ij} = E_{ij}/N$, then each O_{ij} follows a Binomial distribution $B(N, p_{ij})$, and thus

$$E(O_{ij}) = Np_{ij} = E_{ij}, \qquad \text{var}(O_{ij}) = Np_{ij}(1-p_{ij}) = E_{ij}(1 - E_{ij}/N) \approx E_{ij}.$$

Thus, by the Central Limit Theorem, $Z_{ij} = \frac{(O_{ij}-E_{ij})}{\sqrt{E_{ij}}} \backsim N(0,1)$ assuming large N, and we have that

$$T = \sum_{i,j=1}^{n} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i,j=1}^{n} Z^2_{ij} \sim \chi^2_d.$$

The constraint $\sum_{i,j=1}^{n} O_{ij} = N$ means that only $cr - c - r + 1$ of the Z_{ij} can vary independently, thus reducing the degrees of freedom to $(r-1)(c-1)$.

3.4 Sample data analysis

3.4.1 Decision procedure

In this part, we choose a level α at which we want to test H_0. The most common alpha value is $\alpha = 0.05$, establishing a 95% confidence level.
α is the *probability* of rejecting H_0 when H_0 is actually true.
After setting the alpha value, we compute the critical value cv such that $P(\chi^2_d \geq cv) = 0.05$ (Figure 2), that is the probability that a random variable following a χ^2_d distribution takes a value greater than cv.
If the test statistic T takes a value that is greater than cv, it indicates a strong evidence against the null hypothesis, as there is less than 5% chance the null hypothesis is correct. Therefore, we reject the null hypothesis, and accept the alternative hypothesis.

We set the **decision rule**:
Reject H_0 if $T > cv$, otherwise, accept.

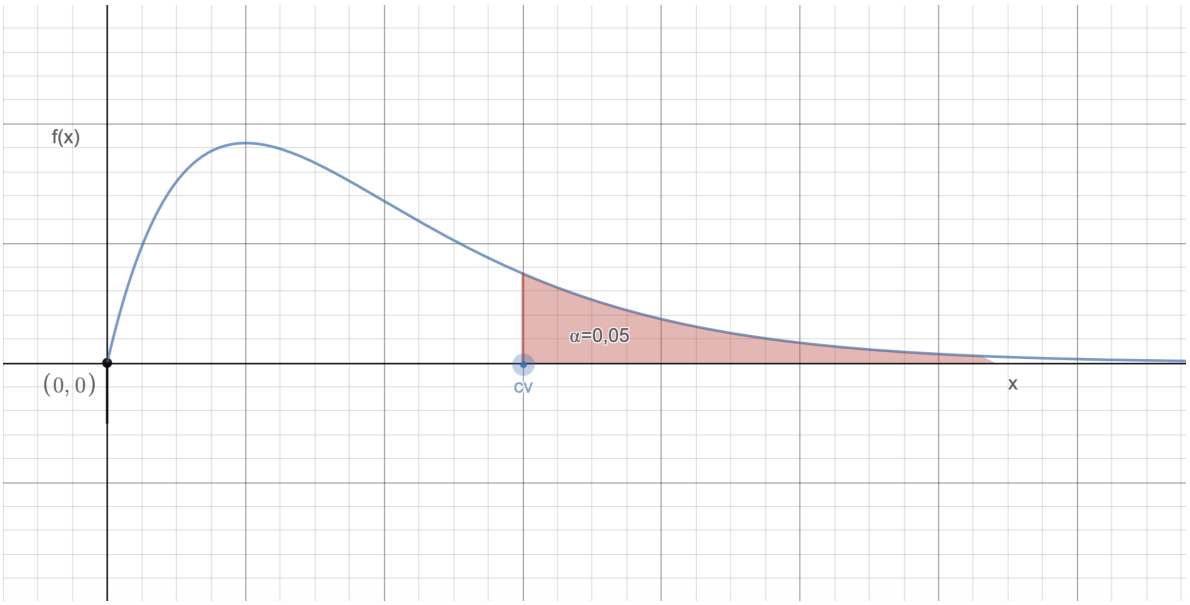


Figure 2: Plot of the Chi-square distribution

3.4.2 Measure of evidence using p-value

The decision procedure as seen above solely consists in either "reject" or "not reject" the null hypothesis. The test statistic T is used to derive the *p-value* that gives more information about the sample data and also measures the evidence against H_0.
The *p-value* is the probability under H_0 of observing a value of the test statistic more extreme than what was actually observed.
Hence, $P(\chi^2_d \geq T) =$*p-value* (Figure 3), where T is the Chi-square statistic.
The smaller the *p-value*, the stronger the evidence against H_0. For this test, we will refer to the commonly used evidence scale shown in Figure 4.

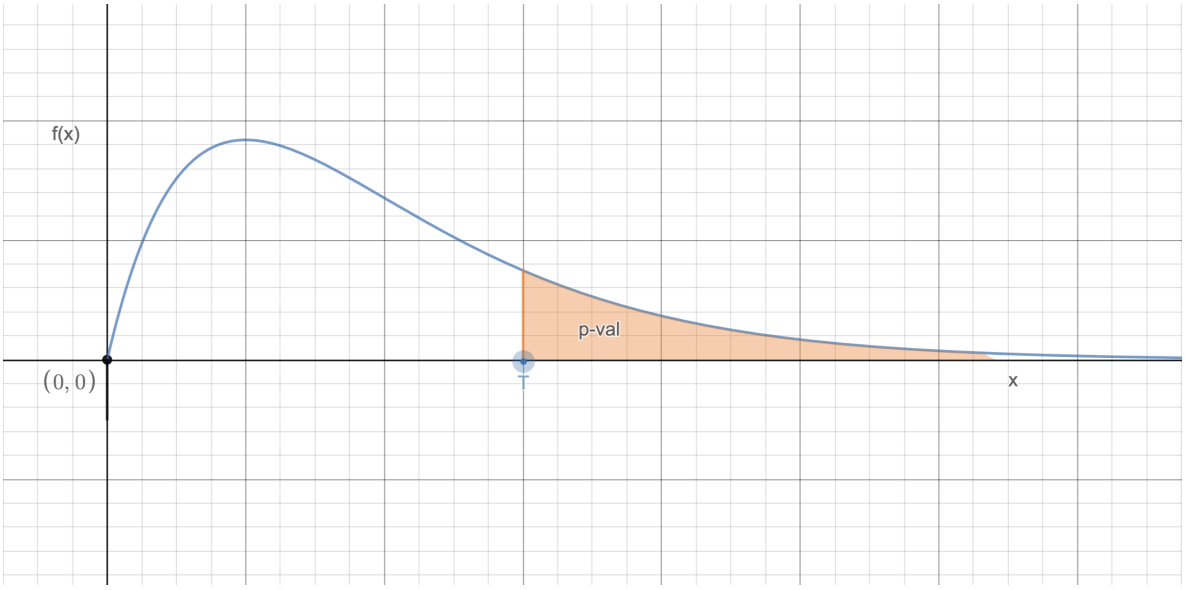


Figure 3: Plot of the Chi-square distribution

p-value	evidence
$< .01$	very strong evidence againt H_0
.01 - .05	strong evidence againt H_0
.05 - .10	weak evidence againt H_0
$> .1$	little or no evidence againt H_0

Figure 4: The evidence scale

4 Test results

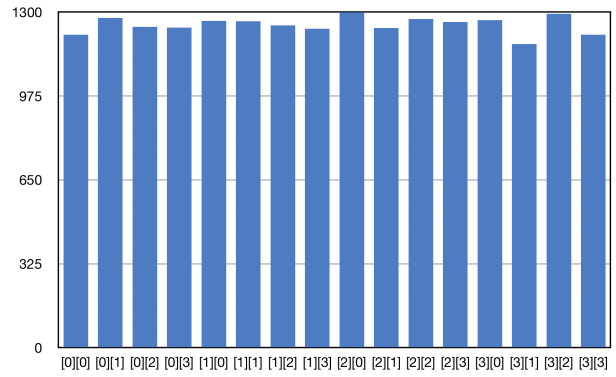
In this section we will present the test results.
We realised 4 tests: by keeping the same prime number p of 2048 bits generated randomly and the same number of iterations, we modified the modulo value n in order to vary the categories. For each n, we also test *sloth* with and without the *xor* operation that is made on the resulted *random number* just before outputting it, to see how this operation affects the result.

- p is generated randomly using the function `mpz_urandomb` provided by GMP library.
- The number of iterations is set to 1. (For 20000 observations, the test took roughly 2 minutes.)
- We generate a different sample for each test.
- In the reported results, we chose to flip the 93^{rd} bit of the binary representation of the seed s for each observation.

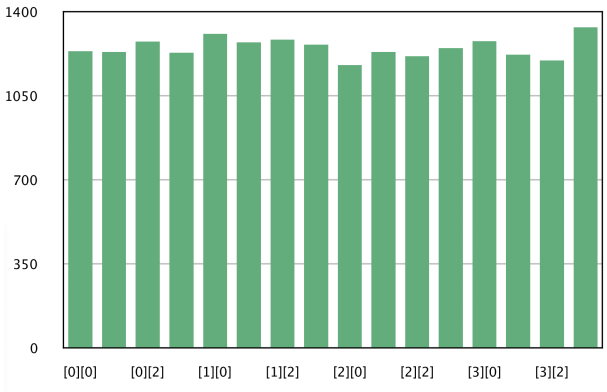
The implementation of the test can be found in `https://github.com/Emna-FENDRI/SlothTest.git`.
After each test, we build the two corresponding **histograms** (with and without the *xor* operation), to visualize the number of observations falling into the n^2 possible categories that we denoted $[i][j]$.
We also **plot** the corresponding χ^2 distribution, its critical value for a 0.05 significance level and the computed test statistics.

4.1 Modulo 4

This first test is done by considering the congruence classes of the generated *beacon*s modulo 4, and the corresponding critical value is **cv= 16.919**.



(a) Without *xor*, for 20000 observations



(b) With *xor*, for 20000 observations

	Test Statistic	**p-val**	**Result**
Without xor	8.745	0.461	No evidence against H_0
With xor	11.651	0.233	No evidence against H_0

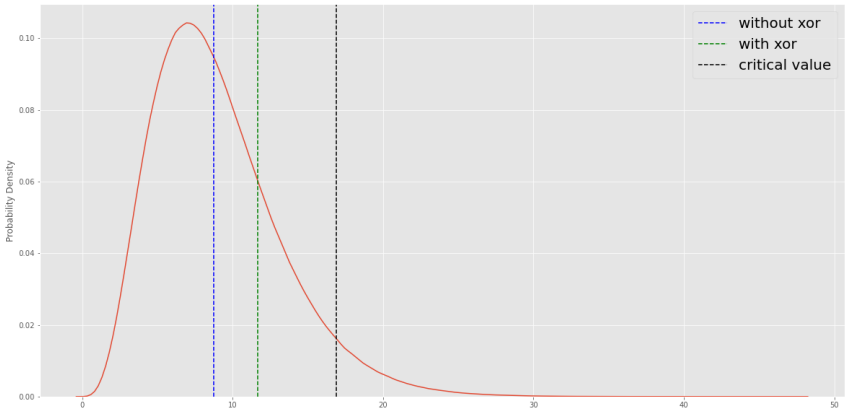
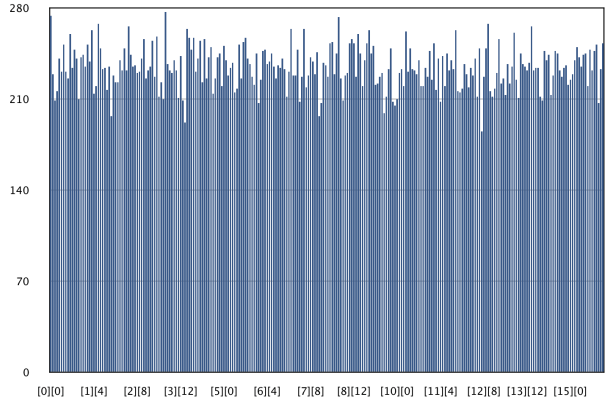


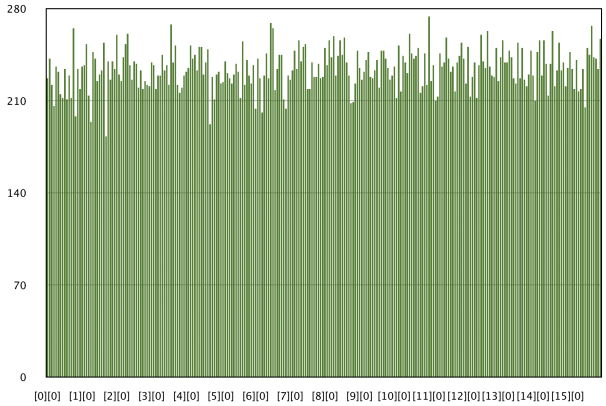
Figure 6: The χ^2_9 probability distribution

4.2 Modulo 16

This second test is done by considering the congruence classes of the generated *beacon*s modulo 16, and the corresponding critical value is **cv= 260.992**.



(a) Without *xor*, for 60000 observations



(b) With *xor*, for 60000 observations

	Test Statistic	**p-val**	**Result**
Without xor	237.940	0,264	No evidence against H_0
With xor	210.511	0,747	No evidence against H_0

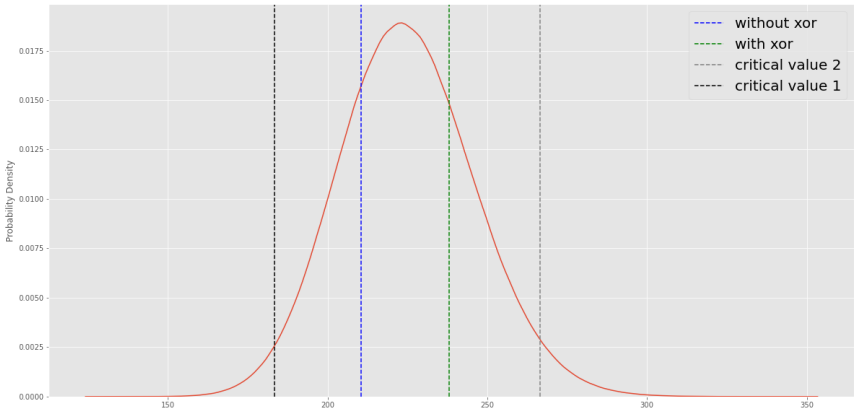
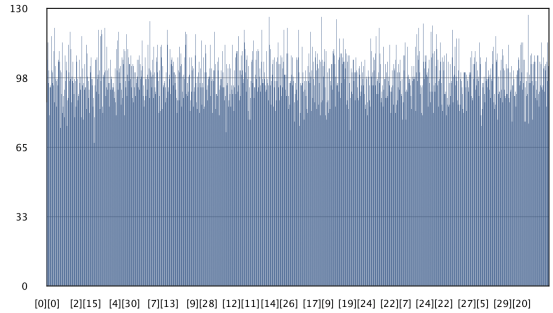


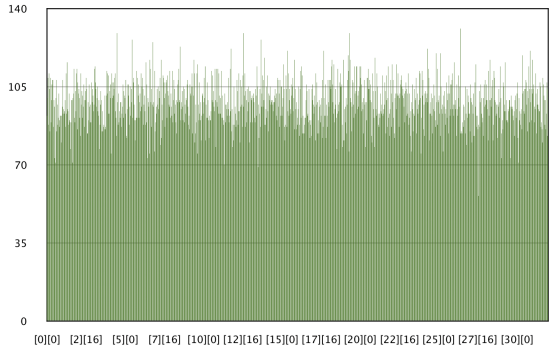
Figure 8: The χ^2_{225} probability distribution

4.3 Modulo 32

This third test is done by considering the congruence classes of the generated *beacon*s modulo 32, and the corresponding critical value is $cv = 1034,2$.
Additionally, due to large degrees of freedom, the χ^2 distribution tends to have a more symmetric shape. So it is interesting to consider a two-tailed test, where each tail covers a probability of 0.025, hence obtaining two critical values: $\mathbf{cv_1 = 875.072}$ and $\mathbf{cv_2 = 1046.93}$, as shown in Figure 10.



(a) Without *xor*, for 100000 observations



(b) With *xor*, for 100000 observations

	Test Statistic	**p-val**	**Result**
Without xor	955.616	0,542	No evidence against H_0
With xor	948.131	0,610	No evidence against H_0

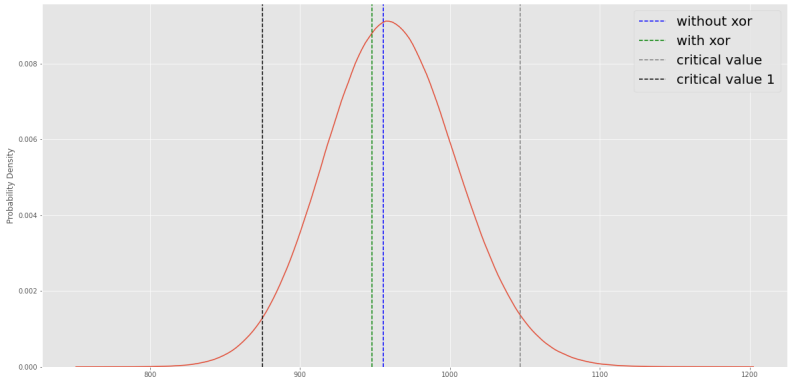
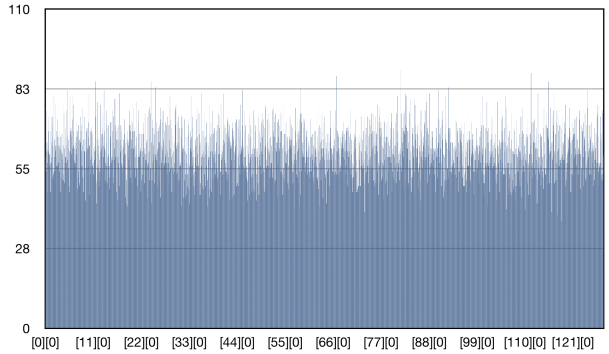


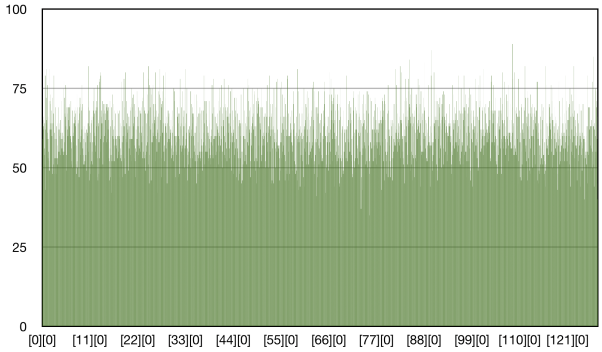
Figure 10: The χ^2_{961} probability distribution

4.4 Modulo 128

This last test is done by considering the congruence classes of the generated *beacon*s modulo 128, and the corresponding critical value is $cv = 16425.3$.
Similarly as above, due to large degrees of freedom, we also consider a two-tailed test, where each tail covers a probability of 0.025, hence obtaining two critical values: $\mathbf{cv_1 = 15777}$ and $\mathbf{cv_2 = 16481}$.



(a) Without *xor*, for 1000000 observations



(b) With *xor*, for 1000000 observations

	Test Statistic	**p-val**	**Result**
Without xor	16043.631	0,737	No evidence against H_0
With xor	16221.472	0,3024	No evidence against H_0

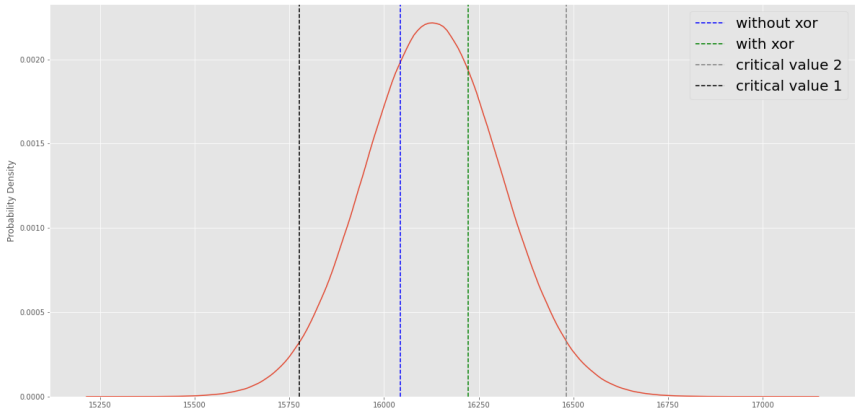


Figure 12: The χ^2_{16384} probability distribution

5 Conclusion

The results obtained in the previous section are very convincing. As we can see from the plots, the test statistics with and without the *xor* clearly fit in the χ^2 distribution, which proves that there is no evidence against H_0, as expected.
Finally, we can conclude that variables X and Y are not correlated. In fact, the *beacon* resulting from a given seed s will not give any information about the *beacon* obtained from s'.

6 References

[Gingrich(1992)] Paul Gingrich. Chapter 10 chi-square test. *Introductory statistics for the social sciences. Department of Sociology and Social Sciences, University of Regina*, 1992.

[Lenstra and Wesolowski(2015)] Arjen K. Lenstra and Benjamin Wesolowski. A random zoo: sloth, unicorn, and trx. Cryptology ePrint Archive, Report 2015/366, 2015. `https://eprint.iacr.org/2015/366`.

[Wasserman(2013)] Larry Wasserman. *All of statistics: a concise course in statistical inference*. Springer Science & Business Media, 2013.

[Davison(2018)] Anthony C. Davison. *Probability and Statistics for SIC*, 2018.