Detection project

Group 20
Jakob Solheim - 10141
Jakob Austrheim - 10046

April 29, 2025

NTNU – Trondheim
Norwegian University of
Science and Technology

Abstract

This project report, prepared by Group 20, solves the detection task in TTT4275 (Estimation, Detection, and Classification). The report explores the application of detection theory, including binary hypothesis testing, in the context of cognitive radios. In addition to hypothesis testing, the Central Limit Theorem was employed to approximate probability density functions when a sufficiently large number of samples were available. Furthermore, the Neyman-Pearson Lemma was applied to determine thresholds that maximize the probability of detection while keeping the probability of false alarm low. The constructed detector ultimately achieved approximately 100% detection rate with a minimal number of false alarms.

Contents

1 Introduction

Wireless communication is an integral part of today's society. With new technology emerging at a rapid pace, the demand for available frequencies to transfer information is growing. However, the spectrum of frequencies used in wireless communication is a scarce resource. Today, the authorities allocate the available frequency ranges to various service operators. To ensure efficient use of the available spectrum a cognitive radio system can be applied. A cognitive radio system is a form of wireless transfer of information consisting of primary users (PU's) and secondary users (SU's). The PU's pay for the right to use certain frequency ranges to transfer their information. However, when the frequencies are unused, SU's are allowed to detect and thereafter utilize the range for their own use. This is called spectrum sensing. An accurate detector is crucial for cognitive radio systems to work as intended. The SU's need to be fast when detecting whether the PU's are idle, while being careful not to intervene with the PU's use of the frequencies, thus reducing their quality of service.

In this report, the making of a binary detector for spectrum sensing will be explored. Firstly, the relevant statistical theory for the report will be presented in Section 2, followed by a presentation of the tasks (Section 3) and their results in Section 4. Finally, in the conclusion (Section 5), the results will be discussed as well as key points learned by performing the tasks.

2 Theory

Signal detection is derived from statistical theory. An introduction to binary hypothesis testing, Neyman-Pearson detectors and estimators is crucial to understanding the tasks performed in this report, as well as a general understanding of probability density functions that the data in this report are governed by.

2.1 Probability density functions and the central limit theorem

The knowledge of probability density functions (PDF's) is crucial to the understanding of statistical theory. There are several that regularly appear when performing statistical analysis. One of these is the Gaussian distribution. It is expressed mathematically by:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

Where μ is the expected value and σ is the standard deviation. A graphical representation of the shape of the Gaussian distribution with $\mu = 0$ and $\sigma = 1$ is shown in Figure 1.

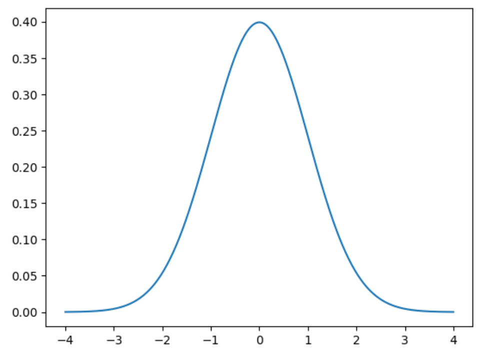


Figure 1: Gaussian distribution with $\mu = 0$ and $\sigma = 1$

A Gaussian distribution with $\mu = 0$ and $\sigma = 1$ is called a standard normal distribution.

The Gaussian distribution is closely related to the central limit theorem (CLT). The central limit theorem states that a sum, or the mean, of independently distributed stochastic variables (SV's) will, given enough samples, be similar to a gaussian distribution [4]. Typically, N $>$ 30 samples will yield a good approximation.

Another important distribution in this report is the χ^2-distribution. A χ^2-distributed SV is defined as a sum of squared, independent standard normally distributed SVs [4]. Its PDF is given by:

$$f(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} exp(-x/2), & \text{for } x \geq 0, \\ 0, & \text{else} \end{cases}$$

Where ν is degrees-of-freedom (dof), while Γ is the gamma function. The expectation value of a χ^2-distributed SV is equal to its dof (ν), while the variance is 2ν. Below is a graphical representation of the χ^2-distribution with $\nu = 3$ (Figure 2). The sum of χ^2-distributed random variables is a χ^2-distributed random variable, with degrees of freedom equal to the sum of the degrees of freedom of the individual distributions.

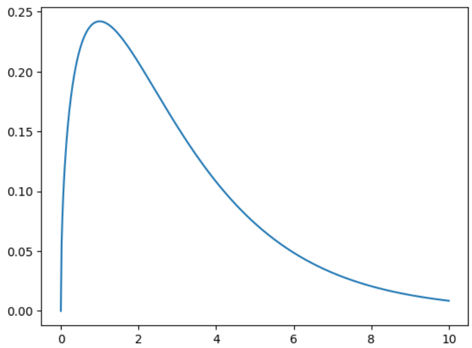


Figure 2: Chi-squared distribution with $\nu = 3$

The cumulative density function (CDF) for a continuous PDF is given as [4]:

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t \tag{1}$$

Q is the tail-probability of the CDF, ergo the area under the graph to the right of a given threshold:

$$Q(x) = \int_{x}^{\infty} f(t)\,\mathrm{d}t \tag{2}$$

2.2 Estimators for expectation value and variance for Gaussian distributions

Estimators are methods used to approximate unknown parameters based on observed data. In statistics, they are commonly applied to estimate moments of PDFs, such as the expectation value and variance. An MVU (Minimum Variance Unbiased) estimator is an estimator that is unbiased and achieves the lowest possible mean squared error (MSE) among all unbiased estimators [2].

For a Gaussian distribution, the MVU estimator of the expectation value is the sample mean, given by:

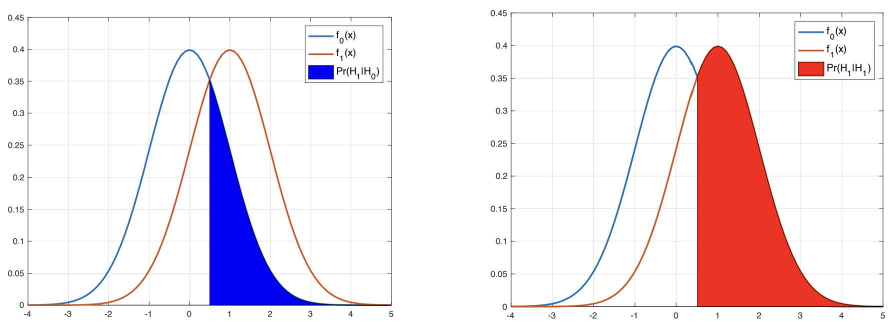
$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

The MVU estimator for the variance is given by:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{4}$$

2.3 Binary hypothesis testing in signal detection

Binary hypothesis testing is a widely used tool in statistical theory. This is used to determine whether some data is best governed by one of two hypotheses [3]. Where in this project its whether a signal is present within a frequency interval or not. One differs between a null hypothesis (H_0) and an alternate hypothesis (H_1), where the goal is to either reject or accept H_1 with as high precision as possible. These hypotheses have their own PDF's, where each represents a partition of the space of possible outcomes. The partition is formed on a threshold, λ, where the null hypothesis space is Ω_0 and the alternative hypothesis space is Ω_1.



(a) Probability of false alarm

(b) Probability of detection

Figure 3: Graphical representation of the probabilities of false alarm 3(a) and detection 3(b). Figures are retrieved from [1]

A threshold has to be set in order to perform a binary hypothesis. The threshold is determined by a tolerated probability of "false alarm" (P_{FA}). If a sample is larger/smaller (depending on the situation) than the given threshold, H_1 is accepted. "False alarm" is one of four outcomes from a binary hypothesis test. The term is used when H_1 is accepted wrongfully, ergo when the sample is larger than the threshold, but is actually governed by the PDF of H_0, denoted $f_0(x)$. It is mathematically expressed by:

$$P_{FA} = \int_{\Omega_1} f_0(x)dx = \int_{\lambda}^{\infty} f_0(x)dx$$

The opposite outcome is detection. This is the probability (PD) of accepting H_1 when the sample data is larger than the threshold, and is actually represented by the PDF of H_1 ($f_1(x)$):

$$P_D = \int_{\Omega_1} f_1(x)dx = \int_{\lambda}^{\infty} f_1(x)dx$$

These probabilities are graphically represented in Figure 3 above. Generally, the P_D, or the "power" of the test, increases with increasing amount of data. A larger number of data points leads to more accurate estimates, making it easier to distinguish between H_0 and H_1.

2.4 Neyman-Pearson detectors

When designing a detector a high probability of detection, as well as a low probability of false alarm, is desired. This is however a difficult compromise to make. Shown by Figure 3, altering the threshold increases both the P_D and the P_{FA} since they both are dependent on the same partition Ω_1.

The Neyman-Pearson Lemma is used to determine the best possible threshold for partitioning the outcome space [3]. It finds a threshold that ensures an acceptable P_{FA}, as well as a large P_D. This is done by minimizing the following cost function:

$$J = P_D + \lambda(P_{FA} - \alpha)$$

Where α is the accepted P_{FA}. By altering this eq, the Neyman-Pearson detector can be expressed as:

$$\Lambda(x) = \frac{f_1(x)}{f_0(x)} > -\lambda = \gamma$$

Where γ is the threshold. The Neyman-Pearson detector is considered to be well performing in a binary detection problem, like detecting the presence of constant signal in noise as in this report.

3 Tasks

Detecting the presence of a primary user in the spectrum can be described as the detection problem:

$$H_0 : x[n] = w[n], n = 0, 1, ..., N - 1$$
$$H_1 : x[n] = s[n] + w[n], n = 0, 1, ..., N - 1$$

where w[n] is additive white noise, and s[n] is the sequence of the primary user.

Throughout this project, a Neyman-Pearson detector was developed to solve the problem above through different tasks. In this section, these tasks will be presented.

3.1 Task 1: Model building

Before designing the detector, it is important to establish a statistical model that explains the way the data x[n] is generated in both hypotheses. To accompany this task, two datasets are provided. Both containing realizations of PU data. In this task, the objective is to build a statistical model by analyzing the histograms of the real and imaginary parts of the data, and assessing whether they are normally distributed according to $s_R(n) \sim \mathcal{N}\left(0, \frac{\sigma_s^2}{2}\right)$ and $s_I(n) \sim \mathcal{N}\left(0, \frac{\sigma_s^2}{2}\right)$.

3.2 Task 2: One-sample-detector:

Developing a NP-detector for a single sample of the spectrum is the objective of task 2. It is given that the test statistic can be simplified to the square of the modulus of the signal:

$$|x[0]|^2 = x_R[0]^2 + x_I[0]^2 > \lambda'$$

Finding λ' gives a decision rule only dependent finding the signal value, as all other values are known.

3.3 Task 3: Performance of the one-sample-detector

In this task, the goal is to confirm that the test statistic obtained in the previous task is distributed according to a χ^2-distribution when scaled appropriately, and calculate the P_D and P_{FA} of the detector. Two datasets are provided, one with no primary user signal present and one with the primary user signal active.

3.4 Task 4: NP detector with data set of K samples

Expanding the detector designed in task 3 to the general case of K samples, while calculating the threshold λ' that maximizes the P_D while keeping the P_{FA} below a set value is the objective of this task.

3.5 Task 5: Performance of a general NP detector

The performance of the detector developed in Task 4 is now to be evaluated. This will be done by computing the distribution of the test statistic, and then using this to plot the ROC-characteristic of the detector.

3.6 Task 6: Approximate performance of a general NP detector:

Using the central limit theorem it is possible to approximate the test statistic as a Gaussian distribution. In this task the expectation value and variance of the test statistic will be found, and used to make a CLT approximation.

3.7 Task 7: Complexity of the detector

Knowing that the test statistic can be approximated as a Gaussian distribution, the task is to find an expression for the number of samples needed to obtain a certain P_D and P_{FA}.

3.8 Task 8: Numerical experiments in PU detection

In the final task, the detector developed throughout the project will be used on a dataset containing 100 realizations of signals observed by the secondary user. This is to see if the detector takes the right decisions.

4 Implementation and results

4.1 Task 1: Model building

The time domain signal s(n) is obtained by applying the IDFT (inverse discrete Fourier transform) to the frequency-domain symbols in the s[k] dataset. This was done by manually using Equation (5) below

$$s[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S[k]exp(\frac{j2\pi nk}{N}), \quad n = 0, 1, \ldots, N-1. \tag{5}$$

Thereafter the real and imaginary parts of s[n] were separated into S_R and S_I respectively, before finally creating histograms for each (Figure 4.

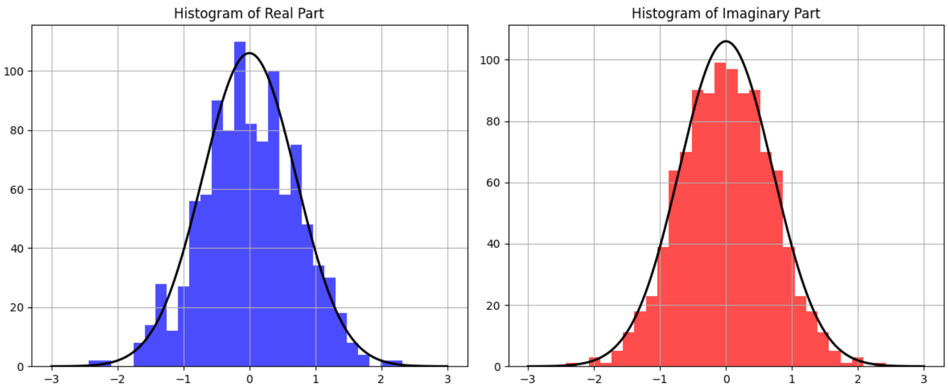


Figure 4: Histograms for s[n] obtained applying IDFT on the realizations of Gaussian PU data s[k]

Furthermore, an estimate of $\mathbb{E}\{s_R[n]s_I[n]\}$ and $\mathbb{E}\{s[n]\} = \mathbb{E}\{s_R[n]\} + j\mathbb{E}\{s_I[n]\}$ was found by calculating the sample mean of both S_R and S_I separately as well as their product $s_R[n]s_I[n]$. This resulted in the values in Table 4.1.

$\mathbb{E}\{s(n)\} = \mathbb{E}\{s_R(n)\} + j\mathbb{E}\{s_I(n)\}$	$\mathbb{E}\{s_R(n)s_I(n)\}$
$0.0148 + j\,1.703$e-15	3.58e-15

Table 1: Expectation value of s(n) and the product of its real and imaginary components respectively

The expectation values acting as real and imaginary components in the expectation value of s(n) above are approximately equal to the expectation value of E$\{S_R\}$ = E$\{S_I\}$ = 0 in Section 4.1. Similarly, the expected value of the product of the real and imaginary parts is also approximately zero. This product should, if the real and imaginary are identically distributed with $\mu = 0$ (as is the case in Section 4.1), be equal to the variance, $\frac{\sigma_s^2}{2}$ (VarX = EX^2 when EX = 0). This is however not the case. Looking at the histogram in Figure 4 it is clear to see that the standard deviation of the measurements is closer to 0.6/0.7. The reason for this might be the amount of provided samples s[k]. With more available realizations the estimate for the expectation values would likely be more accurate.

4.2 Task 2: One-sample-detector

A single sample was used to derive a Neyman-Pearson detector in this task. The detection problem for a single sample is given by:

$$H_0 : x[0] = w[0]$$
$$H_1 : x[0] = s[0] + w[0]$$

In the problem description, it is given that $w[0] \sim \mathcal{CN}(0, \sigma_w)$ and $s[0] \sim \mathcal{CN}(\mu_s, \sigma_s)$. From this, it follows that the PDF of observing a sample x[0] under the null-hypothesis H_0 is:

$$p_0(x[0]) = \frac{1}{\pi(\sigma_w^2)} exp(\frac{|x[0]|^2}{\sigma_w^2})$$

Under H_1, x[0] is assumed to be the sum of two complex Gaussian distributions s[0] and w[0]. This sum results in another complex Gaussian distribution. The mean and variance of x[0] under H_1 are given by:

$$\mathbb{E}\{x[0]\} = \mathbb{E}\{s[0] + w[0]\} = \mathbb{E}\{s[0]\} + \mathbb{E}\{w[0]\} = \mu_s + 0 = \mu_s$$

and,

$$Var\{x[0]\} = Var\{s[0] + w[0]\} = Var\{s[0]\} + Var\{w[0]\} = \sigma_s^2 + \sigma_w^2$$

It follows that the PDF of observing a single sample x[0] under H_1 is:

$$p_1(x[0]) = \frac{1}{\pi(\sigma_w^2 + \sigma_s^2)} exp(\frac{|x[0] - \mu_s|^2}{\sigma_w^2 + \sigma_s^2)})$$

Using this, we can derive the following LRT:

$$\Lambda(x) = \frac{p_1(x[0])}{p_0(x[0])} \begin{cases} \geq \lambda & \Rightarrow H_1 \\ < \lambda & \Rightarrow H_0 \end{cases}$$

$$\Lambda(x) = \frac{\frac{1}{\pi(\sigma_w^2+\sigma_s^2)} exp(\frac{|x[0]-\mu_s|^2}{(\sigma_w^2+\sigma_s^2)})}{\frac{1}{\pi(\sigma_w^2)} exp(\frac{|x[0]|^2}{\sigma_w^2})} = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_s^2} exp(\frac{|x[0] - \mu_s|^2}{\sigma_w^2 + \sigma_s^2} - \frac{|x[0]|^2}{\sigma_w^2})$$

By taking the logarithm on both sides and moving all known terms on one side the decision rule for taking hypothesis H_1 is obtained. The approximation of $\mu_s \approx 0$ is also applied.

$$|x[0]|^2 > \frac{\sigma_s^2(\sigma_s^2+\sigma_w^2)}{\sigma_w^2}(ln(\lambda) - ln(\frac{\sigma_w^2}{\sigma_s^2+\sigma_w^2})) = \lambda'$$

4.3 Task 3: Performance of the one-sample-detector

By doubling the square of the sample magnitude, $|x[0]|^2$, and scaling it with the inverse of the corresponding variance for both hypotheses, it can be seen that both sets of samples follow a χ^2-distribution with two degrees of freedom. This can be seen in Figure 5. When evaluating the performance of the detector, one should consider the probability of false alarm and probability of detection. The expression for P_{FA} was found:

$$P_{FA} = Pr\{|x[0]|^2 > \lambda'|H_0\} = Pr\{\frac{2|x[0]|^2}{\sigma_w^2} > \frac{2\lambda'}{\sigma_w^2}|H_0\} = Q\left(\frac{2\lambda'}{\sigma_w^2}\right)$$

$$= \int_{\frac{2\lambda'}{\sigma_w^2}}^{\infty} \frac{1}{2}exp(-x/2)dx = exp\left(\frac{-\lambda'}{\sigma_w^2}\right)$$

The same approach can be used for P_D:

$$P_D = Pr\{|x[0]|^2 > \lambda'|H_1\} = Pr\{\frac{2|x[0]|^2}{(\sigma_s^2+\sigma_w^2)} > \frac{2\lambda'}{(\sigma_s^2+\sigma_w^2)}|H_1\} = Q\left(\frac{2\lambda'}{\sigma_s^2+\sigma_w^2}\right)$$

$$= \int_{\frac{2\lambda'}{\sigma_s^2+\sigma_w^2}}^{\infty} exp(-x/2)dx = exp\left(\frac{-\lambda'}{\sigma_s^2+\sigma_w^2}\right)$$

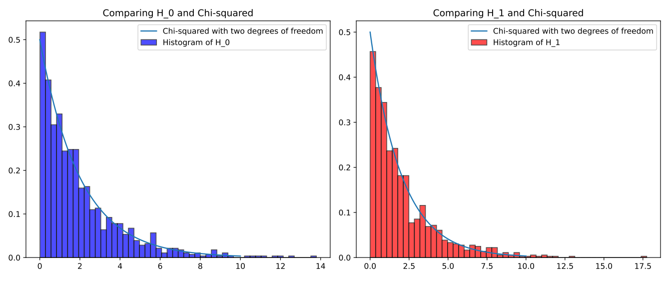


Figure 5: Scaled data when PU is not present (left) and with PU signal present (right) compared to χ^2 (2)

4.4 Task 4: NP detector with data set of K samples

Expanding the detector to the general case of K samples can be done by considering the joint density of K samples when in H_1 and when in H_0

$$\Lambda(\mathbf{x}) = \frac{p_1(x[0], x[1], ..., x[K-1])}{p_0(x[0], x[1], ..., x[K-1])}$$

$$\Lambda(\mathbf{x}) = \frac{\left(\frac{1}{\pi(\sigma_w^2+\sigma_s^2)}\right)^K \exp\left(-\sum_{k=0}^{K-1} \frac{1}{\sigma_w^2+\sigma_s^2}|x[k]|^2\right)}{\left(\frac{1}{\pi(\sigma_w^2)}\right)^K \exp\left(-\sum_{k=0}^{K-1} \frac{1}{\sigma_w^2}|x[k]|^2\right)}$$

Using the same approach as in task 3, a decision rule can be expressed as:

$$\sum_{k=0}^{K-1} |x[k]|^2 \geq \frac{\sigma_w^2(\sigma_w^2+\sigma_s^2)}{\sigma_s^2}\left(\ln\lambda - K\ln\left(\frac{\sigma_w^2}{\sigma_s^2+\sigma_w^2}\right)\right) = \lambda'$$

The simplified test statistic for this LRT is the sum of K squared sample magnitudes. As established in Task 3, each of these are χ^2-distributed. According to the theory presented in Section 2.1, this sum is also χ^2-distributed. Since each squared sample magnitude is distributed with 2 dof, the test statistic in the general case will have 2K.

Next, an equation for maximizing the P_D for a given P_{FA} was found

$$P_D = Pr\{\sum_{n=0}^{K-1}|x[n]|^2 > \lambda'|H_1\} = Pr\{\frac{2}{\sigma_s^2+\sigma_w^2}\sum_{n=0}^{K-1}|x[n]|^2 > \frac{2\lambda'}{\sigma_s^2+\sigma_w^2}|H_1\}$$

Knowing that $T(\mathbf{x}) \sim \chi^2(2K)$, the expression for P_D now becomes:

$$P_D = \int_{\frac{\lambda'}{\sigma_s^2+\sigma_w^2}}^{\infty} \frac{x^{K-1}exp(-\frac{x}{2})}{2^K\Gamma(K)} = Q\left(\frac{2\lambda'}{\sigma_s^2+\sigma_w^2}\right)$$

Furthermore, the expression for P_{FA} is similar:

$$P_{FA} = \int_{\frac{2\lambda'}{\sigma_w^2}}^{\infty} \frac{x^{K-1}exp(-\frac{x}{2})}{2^K\Gamma(K)} = Q\left(\frac{2\lambda'}{\sigma_w^2}\right)$$

The goal is to find a λ' such that P_D is maximized while P_{FA} does not exceed a set value, α_0.

$$Q\left(\frac{\lambda'}{\sigma_w^2}\right) \leq \alpha_0$$

$$\lambda' \leq \sigma_w^2 Q^{-1}(\alpha_0)$$

Using this the threshold λ' can be chosen to ensure a false alarm rate below a predefined value α_0, which is essential for controlling detection performance in practice.

4.5 Task 5: Performance of a general NP detector

As found in part 4 the test statistic is χ^2-distributed with $\nu = 2K$. In Figure 6 the ROC for the test statistic has been plotted. As the graph shows the test statistic becomes better with an increasing ν. This is reasonable as a large ν means that there several independent measurements (large K), resulting in a better performing detector.

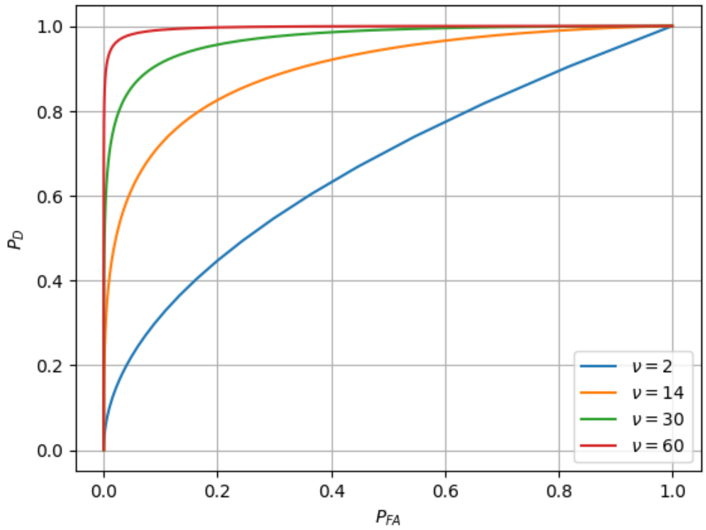


Figure 6: ROC for the test statistic with increasing ν

4.6 Task 6: Gaussian approximation with K samples

With an increasing amount of samples the χ^2-distribution starts to resemble a Gaussian curve, as shown in Figure 7.

Due to the large amount of samples present (N = 1024) the central limit theorem is applied to obtain an approximation of the test statistic $T(\mathbf{x})$. Calculating the first- and second moment of the test statistic is needed to find this approximation.

$$\hat{\mu} = \mathbb{E}\{\sum_{n=0}^{K-1} |x[n]|^2\} = \sum_{n=0}^{K-1} \mathbb{E}\{|x[n]|^2\} = \sum_{n=0}^{K-1} \mathbb{E}\{|x[n]|^2 - \mathbb{E}\{|x[n]|\}^2\} = \sum_{n=0}^{K-1} Var\{|x[n]|\}$$

This value depends on if the sample is from H_0 or H_1.

$$\hat{\mu} = \begin{cases} K\sigma_w^2 & : H_0 \\ K(\sigma_w^2 + \sigma_s^2) & : H_1 \end{cases}$$

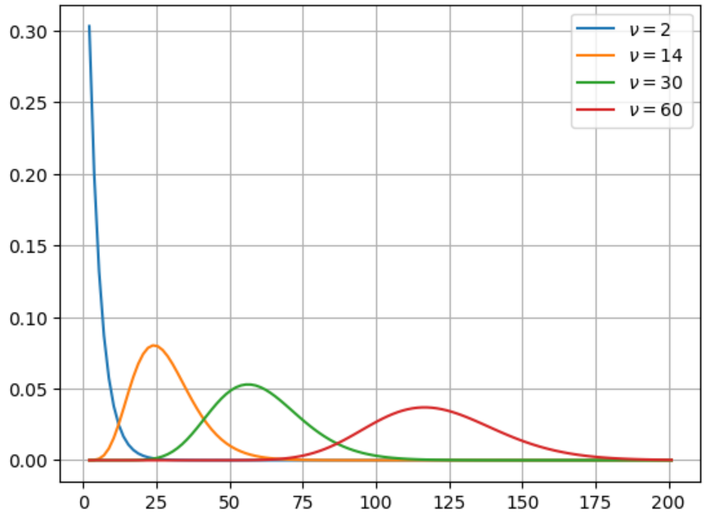


Figure 7: χ^2-distribution with increasing ν

$$\hat{\sigma} = Var\{\sum_{n=0}^{K-1} |x[n]|^2\} = \sum_{n=0}^{K-1} Var\{x_{Re}[n]^2 + x_{Im}[n]^2\} = \sum_{n=0}^{K-1} Var\{x_{Re}[n]^2\} + Var\{x_{Im}[n]^2\}$$

$$= \sum_{n=0}^{K-1} \left(\mathbb{E}\{x_{Re}[n]^4 - \mathbb{E}\{x_{Re}[n]^2\}^2\} + \mathbb{E}\{x_{Im}[n]^4 - \mathbb{E}\{x_{Im}[n]^2\}^2\}\right)$$

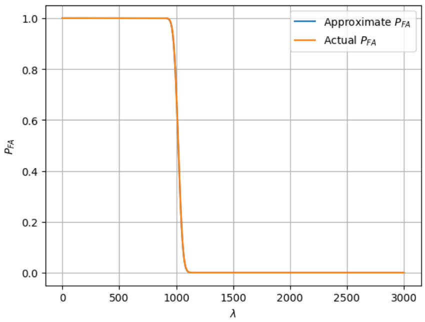
The value of the $\sigma_{0/1}$ variable depends on if the H_0- or H_1 distribution is considered.

$$= 2K(3\left(\frac{\sigma_{0/1}^2}{2}\right)^2 - \sigma_{0/1}^4) = K(\frac{12\sigma_{0/1}^4}{4} - 2\sigma_{0/1}^4) = K\sigma_{0/1}^4$$

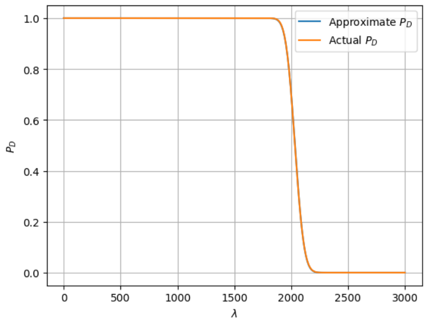
The variance of the test statistic can be expressed as:

$$\hat{\sigma} = \begin{cases} K\sigma_w^4 & : H_0 \\ K(\sigma_w^2 + \sigma_s^2)^2 & : H_1 \end{cases}$$

As seen in Figure 8(a) and Figure 8(b) the approximation is nearly identical with the original distribution of the test statistic. This is however not surprising considering the amount of samples.



(a) The real distribution of the P_{FA} plotted along with its gaussian approximation



(b) The real distribution of the P_D plotted along with its gaussian approximation

4.7 Task 7: Complexity of the detector

As shown in the previous task, the test statistic can be approximated as Gaussian distributed by the central limit theorem.

Using these distributions, an expression for the amount of samples K needed for a given P_D and P_{FA} can be found.

$$P_{FA} = Pr\{\sum_{n=0}^{K-1} |x[n]|^2 < \lambda' | H_0\} = Q\left(\frac{\lambda' - K\sigma_w^2}{\sqrt{K}\sigma_w^2}\right) = \alpha$$

Where Q represents the complementary cumulative distribution function of a standard normal distribution. Furthermore, isolating λ' on one side of the equation gives:

$$\lambda' = \sqrt{K}\sigma_w^2 Q^{-1}(\alpha) - K\sigma_w^2$$

The probability of detection can be written as

$$P_D = Pr\{\sum_{n=0}^{K-1} |x[n]|^2 < \lambda' | H_1\} = Q\left(\frac{\lambda' - K(\sigma_s^2 + \sigma_w^2)}{\sqrt{K}(\sigma_s^2 + \sigma_w^2)}\right)$$

Rewriting this equation and inserting the expression for λ' found previously gives:

$$\frac{\lambda' - K(\sigma_s^2 + \sigma_w^2)}{\sqrt{K}(\sigma_s^2 + \sigma_w^2)} = \frac{\sqrt{K}\sigma_w^2 Q^{-1}(\alpha) - K\sigma_w^2 - K(\sigma_s^2 + \sigma_w^2)}{\sqrt{K}(\sigma_s^2 + \sigma_w^2)} = Q^{-1}(\beta)$$

Solving this equation for K gives

$$K = \left(\frac{\sigma_w^2 Q^{-1}(\alpha) - Q^{-1}(\beta)(\sigma_s^2 + \sigma_w^2)}{\sigma_s^2}\right)^2$$

4.8 Task 8: Numerical experiments in PU detection

The NP-detector using the CLT approximation of the test statistic was applied on a set of 100 signal realizations, with sample size N = 256. The results when using a threshold giving $P_{FA} = 0.1$ and $P_{FA} = 0.01$ can be seen in Table 2 and Figure 8 and 9.

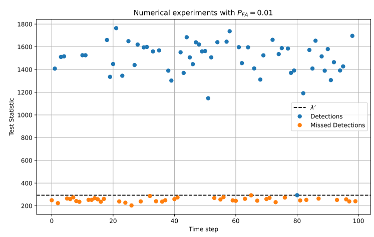


Figure 8: Detection results using $P_{FA} = 0.01$ with Gaussian approximation

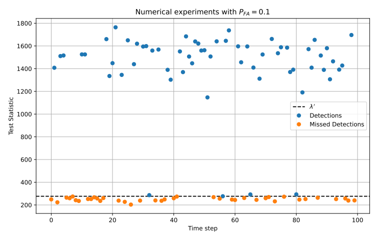


Figure 9: Detection results using $P_{FA} = 0.1$ with Gaussian approximation

P_{FA}	**Detections**	**Missed Detections**
0.1	61	39
0.01	58	42

Table 2: Detection results based on P_{FA} using the Gaussian approximation found in task 6 and 7

When using both thresholds the P_D is approximately 1, meaning that all PUs are likely to be detected. From this it is certain that at least 3 of the detections when using $P_{FA} = 0.1$ are false alarms, as the number of detections decreases by 3 when P_{FA} is decreased to 0.01, while the P_D remains the same. Due to the low probability of false alarms when using $P_{FA} = 0.01$ and that all PUs are likely to be detected, it is fair to be confident in the results provided by the detector. While it is not possible to confirm the absence of false alarms without comparing it to a detector with lower P_{FA}, it is fair to assume that very few of the detections are false alarms due to the low P_{FA}.

Lowering the amount of samples per signal realization to 5 while having $P_{FA} = 0.1$ impacted the probability of detection. The new P_D decreased to approximately 0.951. This is still high considering the amount of samples removed per signal. The signal to noise ratio in the signal is the reason for this, as σ_s^2 is 5 times larger than σ_w^2, making the signal easier to distinguish from the noise. When lowering σ_s^2 to the same value as σ_w^2 and maintaining a sample size of 5 and $P_{FA} = 0.1$, a $P_D \approx 0.683$ could be observed. When using the original sample size and the new signal-to-noise ratio, the probability of detection remained close to 1.

The detector developed in section 4.4 was also used on the dataset. Using both $P_{FA} = 0.1$ and 0.01, the probability of detection was approximately 1. The main distinction between the Gaussian approximation and the detector using the scaled χ^2-distributed test statistic was the number of detections. Using both P_{FA}-values, the number of detections was lowered by 1 when not using the CLT-approximation. It is possible that the approximation for 256 samples was not accurate enough, leading to an extra false alarm.

5 Conclusion

Powerful detectors are crucial to the functioning of cognitive radios, ensuring good quality of service for PU's. Binary hypothesis testing is the foundation for detection problems of this sort. In this report the Neyman-Pearson Lemma was applied to obtain a threshold to partition the outcome space on. The Neyman-Pearson detector is generally a good choice of detector when trying to determine the presence of a constant signal in noise. Firstly, in Section 4.1 the complex-valued time-domain signal s[n] was obtained through inverse discrete Fourier transform of gaussian PU symbols s[k]. The expected value of both the real and imaginary parts, as well as their product, was found. The data was also plotted in histograms. Both expectation values were approximately equal to zero as given in the task 4.1. However, the variance (given by the expectation value of their product) was also zero, contradicting the histograms. The histogram and the analytical values would likely converge with a larger number of samples

A one-sample NP detector was created in Section 4.2, followed by a K-sample detector in the following task (Section 4.3). The test statistic which was used was found to be distributed according to a scaled χ^2-distribution. A ROC was used to study the performance of the detector with an increasing amount of samples in Section 4.5. Not surprisingly, the performance increased as a result of more samples. Furthermore the central limit theorem was applied to obtain an approximation of the test statistic (Section 4.6). The approximation and the real distribution were almost identical with N = 1024 samples. The CLT approximation was then used to provide an expression for the number of samples needed to obtain a given P_D while maintaining a certain P_{FA}.

The numerical experiments that were conducted on the provided datasets in Section 4.8 showed promising results. A high signal-to-noise ratio and large sample size, made for a probability of detection approximately equal to 1 using both detectors and both probabilities of false alarm. It was also shown that using the detectors with a considerably lower sample size could still yield a probability of detection around 0.95, which for some applications might be enough. From the experiments using the lower sample size, the importance of sample size when dealing with lower SNR problems was also highlighted. In real life applications, the signal to be detected is possibly buried in noise, which makes this especially important.

Working with this project we have obtained a practical understanding of theoretical concepts such as hypothesis testing, Neyman-Pearson lemma and ROC's. Especially, we learned that the CLT is great for making good approximations, simplifying calculations, even with relatively few samples. Additionally, we learned that with good SNR the need for many samples is reduced, shown by the performance of the detector in Section 4.8.

References

[1] *Lecture 9 - Hypothesis Testing*. Found on Blackboard (TTT4275 - Estimation, detection and classification).

[2] Tor Andre Myrvoll. *Estimation Theory*. Found on Blackboard (TTT4275 - Estimation, detection and classification).

[3] *Statistical Signal Detection*. Found on Blackboard (TTT4275 - Estimation, detection and classification).

[4] *"Temasider" - TMA4245 (Statistics)*. URL: `https://tma4245.math.ntnu.no/`.