

Estimate The Spectral Envelope Of Bacterial DNA By Using Modified Daniell Kernel

Tahir R. Dikheel

Ali Ghanim Abood

University of AL-Qadisiyah

Corresponding Author : Ali Ghanim Abood

Abstract : In this article, we estimated the spectral envelope density function using the modified Daniell Kernel function for a non-stationarity time series, we solved the non-stationary problem by dividing the spectral envelope into a number of local spectral envelopes to reduce the effect of the trend, the time series used in the this article was the sequence of the nitrogenous bases of Escherichia coli bacteria's DNA. The problem of the article is dealing with categorical data in a non-stationary time series.

Keywords: spectral envelope, Modify Daniell kernel, SNR, Kullback-Leibler divergence, Tree-Based Adaptive Segmentation

Introduction: We can analyze the DNA from a statistical standpoint, so in order to understand this article, we have to understand what statistics are The statistic is: "the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data" as Cowden-Croxtton defined it. Statistics branches out into many sciences such as regression analysis, stochasticity, mathematical statistics, etc.

One of these sciences is called time series, which is a collection of observations generated sequentially through specific parameter, this parameter is most often time; for this reason, this science is called time series, but it is never required that it be always time, in some time series it is place instead of time.

DNA is a number of nucleotides over a specific sequence, and this is similar to a time series, so time series analysis is the appropriate science for DNA analysis.

Most time series data are numbers, and numbers are easy to deal with because time series analysis depends on equations and functions, but there are other types of data that are categorical, as in DNA whose data is of types nucleotides which is A, C, G & T, this requires replacing each category with a number, this is one of the problems in the article as we will see through the article.

Time series can be analyzed within the time domain, but there is another method which is to analyze within the frequency domain. this is done in several ways such as spectral analysis or wavelet analysis, but in this article we will rely on the spectral envelope analysis method.

The series must be stationary, and not to have a trend, otherwise time series analysis will be complicated, this is the main problem in the article. Therefore, there are methods that are used to get rid of the trend. the method adopted in this article is to divide the time series into local series to find local spectrum envelopes to reduce the effect of the trend.

Stoffer in 1993 is considered the closest study to my research, in addition to Stoffer in 2002 and Jeong in (2012).

THE ARTICLE PROBLEM

the problems of this artical is dealing with non-stationary time series, because analyzing non-stationary time series is complex, we will explain it in the second chapter. Also dealing with categorical data in statistical analysis is one of the difficult dilemmas in statistical modeling, since statistical methods are designed mainly to deal with numbers. THE

ARTICLE OBJECTIVE

The article's goal is to analyze a categorical time series within the frequency domain, we will estimate the spectral envelope function, it is worth noting that estimating the spectral envelope function requires a weight function, the weight function that will be used in this article is modified Daniell, the efficiency is measured by calculating the SNR which is mean Signal-to-Noise ratio.

THE FOURIER TRANSFORM

Some phenomena chang over time in a certain pattern, this called a periodic time series, and the fundamental change on time series can be seen in the pattern, but outside the pattern rang, it is a repetition of a pattern in that rang, any repeat in this time series called frequency which can be one of the domains in series analysis.

Analyzing a time series based on frequency requires a periodic signal, this would be a problem if the signal being analyzed was non-periodic, so the Fourier transform is integration that makes make the signal periodic. the Fourier

transform mechanism starts by assuming the aperiodic signal is a periodic signal, and the period time approaches infinity as following:

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

Where $F(\omega)$ Fourier transform, ω the frequency, t the time, $f(t)$ the signal at t time.

We can write the orthogonal trigonometric sine and cosine function as:

$$x_t = \sum_{k=0}^{n/2} \left(a_k \cos\left(\frac{2\pi kt}{n}\right) + b_k \sin\left(\frac{2\pi kt}{n}\right) \right) \quad t = 1, \dots, n \quad (1)$$

This equation are call Fourier series, we can find the Fourier coefficients a_k and b_k by :

$$a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n x_t \cos\left(\frac{2\pi kt}{n}\right) & k = 0 \text{ and } k = n/2 \text{ if } n \text{ is even} \\ \frac{2}{n} \sum_{t=1}^n x_t \cos\left(\frac{2\pi kt}{n}\right) & k = 1, 2, \dots, \frac{n-1}{2} \text{ if } n \text{ is odd} \end{cases}$$

$$b_k = \frac{2}{n} \sum_{t=1}^n x_t \sin\left(\frac{2\pi kt}{n}\right) \quad k = 1, 2, \dots, \frac{n-1}{2}$$

Now we can write the above formula in split-complex formula as following:

$$x_t = \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i\omega_k t} & \text{if } n \text{ is odd} \\ \sum_{k=-n/2+1}^{n/2} c_k e^{i\omega_k t} & \text{if } n \text{ is even} \end{cases}$$

when:

$$\omega_k = \frac{2\pi k}{n} \quad (2)$$

and the Fourier coefficient c_k can be founded by:

$$c_k = \frac{1}{n} \sum_{t=1}^n x_t e^{i\omega_k t}$$

By the equations (1) and (2) we get:

$$x_t = \sum_{k=0}^{n/2} (a_k \cos(\omega_k t) + b_k \sin(\omega_k t))$$

$$= \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i\omega_k t} & \text{if } n \text{ is odd} \\ \sum_{k=-(n/2)+1}^{n/2} c_k e^{i\omega_k t} & \text{if } n \text{ is even} \end{cases}$$

Now we defined a function $f(t)$ as follow:

$$x_t = f(t), \quad -\frac{p}{2} \leq t \leq \frac{p}{2}$$

$$x_{t+jp} = x_t \quad j = \pm 1, \pm 2, \dots$$

that mean x_t are periodic with p , so [11]:

$$x_t = \sum_{k=-\infty}^{\infty} c_k e^{ik\omega_0 t} \quad (3)$$

The estimating of spectral density in equation (3) will not be consistent because of the autocorrelation δk be flexible the higher value of $|k|$ as the following:

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-T}^T \delta_k w(k) \cos(k\omega)$$

The accuracy of the autocorrelation function is low because of the lack of observations, thus the effect of the tails of the autocorrelation function increases, so we have to truncate some values of N which T will be the point of truncate, since this truncation causes the neglect part of the information, so we need a weight function for weighting the value of the autocorrelation function [4].

Modify Daniell is one of these weight function, this rectangular spectral window smooths the model and reduces some of the variability that we saw in the sample spectrum, when we make the smoothing window wider we will reduce the variability even further and it will make the forecasting difficult, but otherwise we will make the bias bigger, and we can defined modify Daniell as following [8]:

$$w_T(K) = \frac{1}{2T+1} \quad \text{for } -T \leq k \leq T$$

THE SPECTRAL ENVELOPE

It is a technique which depend on a frequency-based principal component applied to a multivariate time series, now when we suppose that $x_t, t=0, \pm 1, \pm 2, \dots$ is a stationary categorical time series with finite state-space which is $c = \{c_1, c_2, \dots, c_k\}$ values, then we suppose the numeric value α_j for any c_j . so α is a vector of the real values, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)'$, and $p_j = Pr(x_t = c_j) > 0$, and $h(x_t)$ be stationary time series with the real value.

Now we defined Y_t as the flowing:

$$Y_t = \begin{cases} z_j & \text{if } x_t = c_j \text{ for } j = 1, 2, 3, \dots, k-1 \\ o & \text{if } x_t = c_k \end{cases}$$

When z_j is a vector of k items, all its items are zero except j th row are one, and o are $k \times 1$ vector, all its items are zero, then we come collusion $h(x_t) = \alpha' Y_t$, all so $h(x_t) = \alpha_j$

In our application the probability space are the alphabet of the DNA which is $\{A, C, G, T\}$, so we can design Y_t as following [9]:

$$Y_t = (1, 0, 0)', \text{ when } X_t = A$$

$$Y_t = (0, 1, 0)', \text{ when } X_t = C$$

$$Y_t = (0, 0, 1)', \text{ when } X_t = G$$

$$Y_t = (0, 0, 0)', \text{ when } X_t = T$$

So the goal is chose the best value to α so that maximize the power at each frequency ω as follow:

$$\lambda(\omega) = \max_{\alpha} \frac{f_{xx}(\omega; \alpha)}{\sigma^2(\alpha)}$$

where $\lambda(\omega)$ is the power of the frequency, $f_{xx}(\omega; \alpha)$ is the spectral density and $\sigma^2(\alpha) = var\{h(x_t)\}$.

Suppose the vector process y_t has a continuous spectral density denoted by $f_{yy}(\omega)$ and for each ω there are $k \times k$ complex-valued Hermitian matrix, as we have $h(x_t) = \alpha' Y_t$ suggest $f_{xx}(\omega; \alpha) = \alpha' f_{yy}(\omega) \alpha$. now when $f_{yy}^{re}(\omega)$ are the real part of $f_{yy}(\omega)$ and $f_{yy}^{im}(\omega)$ the imaginary part of $f_{yy}(\omega)$, and as $f_{yy}^{im}(\omega)$ is skew-symmetric, so $f_{yy}^{im}(\omega)' = -f_{yy}^{im}(\omega)$ and $x = x^{re} + ix^{im}$, so $\alpha' f_{yy}(\omega) \alpha = \alpha' f_{yy}^{re}(\omega) \alpha$

$$\lambda(\omega) = \frac{\alpha' f_{yy}^{re}(\omega) \alpha}{\alpha' V \alpha}$$

where V is the variance-covariance matrix of y_t , where $p = (p_1, p_2, \dots, p_k)'$, $V = D - pp'$, and D IS $K \times K$ diagonal matrix, $D = diag\{p_1, p_2, \dots, p_k\}$. by assumption $p_j > 0$, $j = 1, \dots, k$; so the rank(V) = $k - 1$ with the null space of being spanned by I_k for any $k \times (k - 1)$ full rank matrix Q whose columns are linearly independent of I_k . And $Q'VQ$ is a $(k - 1) \times (k - 1)$ positive symmetric matrix.

When $f_{yy}(\omega)$ is a consistent each $j = 1, \dots, J$ the largest root of $f_{yy}^{re}(\omega)$ is distinct, then:

$$\left\{ \frac{\eta_n[\hat{\lambda}(\omega_j) - \lambda(\omega_j)]}{\lambda(\omega_j)}, \eta_n[\hat{\alpha}(\omega_j) - \alpha(\omega_j)]; j = 1, \dots, J \right\} \quad (4)$$

converges jointly in distribution to independent zero-mean, normal distribution as $n \rightarrow \infty$

It can be noted that the value of η_n in the equation (4) depends on the type of the estimator.

In this article the smoothed periodogram matrix

$$I_n(\omega_j) = \hat{f}_{xx} = \sum_{l=-m}^m h_l I_n(\omega_j + l/n)$$

When we use the smoothed periodogram matrix with weight h_l then $\eta_n^{-2} = \sum_{l=-m}^m h_l^2$

applying the following approximations, the peak search can be aided in estimating the smooth spectral envelope, using first order Taylor expansion, we have:

$$\log \hat{\lambda}(\omega) \approx \log \lambda(\omega) + \frac{\hat{\lambda}(\omega) - \lambda(\omega)}{\lambda(\omega)}$$

hence $E[\log \hat{\lambda}(\omega)] \approx \log \lambda(\omega)$ and $var[\log \hat{\lambda}(\omega)] \approx \eta_n^{-2}$ because $\eta_n(\log \hat{\lambda}(\omega) - \log \lambda(\omega))$ is standard normal [10].

TIME SERIES STATIONARITY

A nonstationary time series is one that has a trend, and the trend is any systematical change in the level of a time series, the presence of a trend within the series will cause us problems in time series analysis which is the inability to calculate the mean, variance and autocorrelations, therefore the trend must be eliminated or its impact reduced, the method of dividing a time series into a set of segments one of the methods followed, because the effect of the trend decreases with each segment.

The algorithm that will be followed in the partitioning process is a Tree-Based Adaptive Segmentation, the series is divided by two segmentations which is the level one, then each segmentation are divided by two in level two till k -

times in level k which is the deeper level, when T is the length of the entire series then length of each block are $T/2^k$, we will denote the block $B(k, l)$, l is the l -th block in k level, and N_k is the length of blocks in k level, then if $D(k, l)$ be between two adjacent blocks, $B(K + 1, 2l)$ and $B(K + 1, 2l - 1)$ compute the estimates of the distances $D(k, l)$.

Let's assume $k \times 1$ a vector value of piecewise stationary process, $\{Y_{s,T}\}_{s=0}^{T-1}$ for $T \geq 1$, is given as:

$$Y_{s,t} = \sum_{b=1}^B Y_{s,b} I(s/T, U_b)$$

where $Y_{s,b}$ are stationary processes with continuous $k \times k$ spectral matrices $f_{s,b}(\omega)$. And $U_b = [u_{b-1}, u_b) \subset [0, 1)$ is the interval, and $I(s/T, U_b)$ is an indicator which be equal to 1 if $s/b \in U_b$, and 0 otherwise...

now let rescale time in each block, so [9]:

$$\{Y_{s,b}: s/t \in U_b\} \rightarrow \{Y_{t,b}: t = 0, \dots, M_b - 1\}$$

and the number of observations in segment b is M_b , and $\sum_{b=1}^B M_b = T$, this rescaling of time represents a simple time shift to the origin where $Y_{s,b} \rightarrow Y_{t,b}$ for $s/t \in U_b$ with $t = s - \sum_{i=1}^{b-1} M_i$

It can say that a categorical time series, $\{x_{s,T}\}$, on nonzero marginal and a finite state-space is piecewise stationary when the corresponding $k \times 1$ point process, $\{Y_{s,T}\}$, is piecewise stationary, to be sure that more observations considered within each stationary segment (or block) when it sampling the process $x_{s,T}$, it assumed that the lower bound, M , for the number of observations in each block, b , satisfies $M \rightarrow \infty$ as $T \rightarrow \infty$.

It can define the local spectral envelope as follows when $x_{s,T}$ is a piecewise stationary categorical time series. the local analogue of the optimality criterion:

$$\lambda_b(\omega) \underset{\alpha \propto I_k}{sub} = \frac{\alpha' f_{yy}^{re}(\omega) \alpha}{\alpha' v_b \alpha}$$

for $b = 1, 2, \dots, B$ where v_b is the variance-covariance matrix of $Y_{t,b}$ and $\lambda_b(\omega)$ the local spectral envelope and the corresponding eigenvector $\alpha_b(\omega)$ to be the local optimal scaling of block b and frequency ω .

now it can be show some asymptotic $T \rightarrow \infty$ results for estimators of the local spectral envelope and the corresponding local scaling vectors.

Next, let Q which its columns are linearly independent of I_k , $Q = [I_{k-1} | 0]$, and let \hat{V}_b be the sample variance-covariance matrix obtained from the data in segment b , $\{Y_{s,T}: s/T \in U_b\}$, or equivalently, $\{Y_{s,b}: t = 0, 1, \dots, M_b - 1\}$, so:

$$Y_{s,b} \stackrel{def}{=} \hat{Q} Y_{s,b}$$

This process effected by removing the k -th element from $Y_{t,b}$ so that it is now a $(k-1) \times 1$ vector, we denote in this case:

$$\hat{V}_b \stackrel{def}{=} \hat{Q} \hat{V}_b \hat{Q} \quad \text{and} \quad \hat{f}_{Y,b}(\omega) \stackrel{def}{=} \hat{Q} \hat{f}_{Y,b}(\omega) \hat{Q}$$

It can see that \hat{V}_b and $\hat{f}_{Y,b}(\omega)$ become the upper $(k-1) \times (k-1)$ blocks of the previously defined \hat{V}_b and $\hat{f}_{Y,b}(\omega)$ matrices respectively, in addition it can use the same convention for the population values V_b , and $f_{Y,b}(\omega)$.

In order to more clarification while maintaining generality it determined the buoyant cover of the local sample, $\hat{\lambda}_b(\omega)$ to assimilate the largest eigenvalue of $\hat{g}_b^{re}(\omega)$ where:

$$\hat{g}_b = \hat{V}_b^{-1/2} \hat{f}_{Y,b} \hat{V}_b^{-1/2}$$

The local sample optimal scaling $\hat{\alpha}_b(\omega)$ defined by $\hat{\alpha}_b(\omega) = \hat{V}_b^{-1/2} \hat{u}_b(\omega)$, where $\hat{u}_b(\omega)$ is the eigenvector of $\hat{g}_b^{re}(\omega)$ related with the root hat $\hat{\lambda}_b(\omega)$ we fixed the scale corresponding to k -th at zero, in addition, let $\hat{u}_b(\omega)$ be normalized so $\hat{u}_b' \hat{u}_b(\omega) = 1$, we saw that with the first nonzero entry of $\hat{u}_b(\omega)$ considered positive.

In order to let the application of a general theory to obtain asymptotic distributions for the estimates of the local spectral density $f_{Y,b}(\omega)$ we assume in this section that $Y_{t,b}$ is fixed stationary for each of block b , and all local cumulant spectra of all orders be exist for each series $Y_{t,b}$, the issu of assuming the existence of all local cumulant spectra is not concern only the categorical case because the elements of $Y_{t,b}$ can be only take two values: zero or one, instead of entering excessive notation.

The local periodogram of the data $\{Y_{s,T}: s/T \in U_b\}$ in black b is given by:

$$I_b(\omega) = d_b(\omega) d_b^*(\omega)$$

where

$$d_b(\omega) = M_b^{-1/2} \sum_{t=0}^{M_b-1} Y_{t,b} e^{-2\pi i \omega t}$$

is the finite Fourier transform of the data $\{Y_{s,T}: s/T \in U_b\}$.
where [6]:

$$\hat{f}_{Y,b} = (2m+1)^{-1} \sum_{i=-m}^m I_b(\omega + i/M_b)$$

Although increasing the blocks reduces the effect of the trend, it causes complexity in the algorithm and calculations, so there is a way to reduce the number of blocks without increasing the effect of the trend which is to merge two adjacent blocks when they show similar behavior in the spectral envelope, and we'll use an algorithm Kullback-Leibler divergence to measure the amount of convergence in behavior between blocks, as follow:

$$I(p(X), q(x)) = \sum \left(\log \frac{p(x)}{q(x)} \right) p(x) \geq 0$$

where $p(x)$ and $q(x)$ denote the probability density functions of random variable x [13].

Finally, we calculated the amount of its efficiency by calculating the SNR which is Signal-to-noise ratio where the higher the ratio, the lower the efficiency of the function, and vice versa. and the SNR is given as [5]:

$$SNR = 10 \log \frac{P_s^2}{P_n^2}$$

THE SIMULATION

In the simulation and calculation of results we will use R version 4.1.3.

First we tested our algorithm before used it in a real data, we used the following equation to generate it:

$$X_1(t) = 2 \cos\left(\frac{2\pi t}{10}\right) + \cos\left(\frac{2\pi t}{3}\right) + 0.3\epsilon_1(t)$$

$$X_2(t) = \cos\left(\frac{2\pi t}{3}\right) + 0.01\epsilon_2(t)$$

where $\epsilon_1(t)$ and $\epsilon_2(t)$ are Gaussian white noise and with unit variance, we repeat the experiment 500 times to reach the stability in the results, we set the deepest level at $K = 4$ to get best segmentation of the data set simulated [10].

It can have obtained that the results as follows table;

Table (1)

T=512	T=256	T=128	T=64	T=32	ε2	ε1
0.8609303	0.6884488	0.5388679	0.4298197	0.3474296	0.1	0.1
0.8769899	0.6673244	0.5412174	0.4310931	0.3566414	0.2	
0.8727881	0.6688786	0.5398473	0.4300496	0.3412425	0.3	
0.85884	0.6861094	0.5402371	0.4279381	0.3430284	0.4	
0.8687784	0.6899997	0.5295313	0.4311372	0.3557941	0.5	
1.753989	1.359883	1.078214	0.8580769	0.7134018	0.1	0.2
1.723663	1.347287	1.064123	0.8769	0.6977458	0.2	
1.770104	1.364689	1.067418	0.8724363	0.714317	0.3	
1.767159	1.359966	1.052843	0.8790027	0.7002129	0.4	
1.740164	1.360493	1.055073	0.8641199	0.6998285	0.5	
2.644265	2.06656	1.597485	1.269268	1.057132	0.1	0.3
2.66691	2.066572	1.600587	1.295924	1.08713	0.2	
2.633649	2.093187	1.592947	1.286477	1.055445	0.3	
2.652459	2.039123	1.606712	1.344623	1.066385	0.4	
2.643174	2.022034	1.590978	1.283201	1.041396	0.5	
3.515744	2.752029	2.113182	1.706787	1.388652	0.1	0.4
3.471733	2.701873	2.124478	1.72714	1.400291	0.2	
3.511962	2.750788	2.16706	1.697334	1.378431	0.3	
3.527932	2.71972	2.158698	1.706705	1.403356	0.4	
3.519508	2.718683	2.147391	1.688534	1.387266	0.5	
4.362178	3.40992	2.652292	2.147682	1.768232	0.1	0.5
4.372519	3.435937	2.689711	2.198115	1.720829	0.2	
4.349968	3.352632	2.694096	2.172029	1.816397	0.3	
4.362505	3.404721	2.636933	2.14441	1.727951	0.4	
4.462218	3.425728	2.691139	2.189037	1.718527	0.5	

This table shows the results of the SNR of the simulation

In the table (1), we calculated the SNR value when $\epsilon_1(t) = 0.1, 0.2, 0.3, 0.4, 0.5$ and $\epsilon_2(t) = 0.1, 0.2, 0.3, 0.4, 0.5$, also at different lengths of the time series T , when $T = 32, 64, 128, 256, 512$ we noticed that the SNR value increases with the increase of the $\epsilon_1(t)$ value and the length of the chain, at the same time it is not affected by the $\epsilon_2(t)$ value.

ESCHERICHIA COLI DNA

In this article, the application made to the DNA of Escherichia coli strain Iso00225 chromosome, this data has been collected from the website NCBI.

It is one of the most important types of bacteria that commonly found in the intestines of mammals. it was discovered by German-Austrian pediatrician and biologist Theodor Escherich.

In general, the effects of this bacterium are harmless and may even be beneficial, as it is an important source of vitamin K, However, there are many strains that may cause disease in mammals that are infected especially in humans like Urinary tract infection or internal bleeding, this will be the length of the string $T = 8192$.

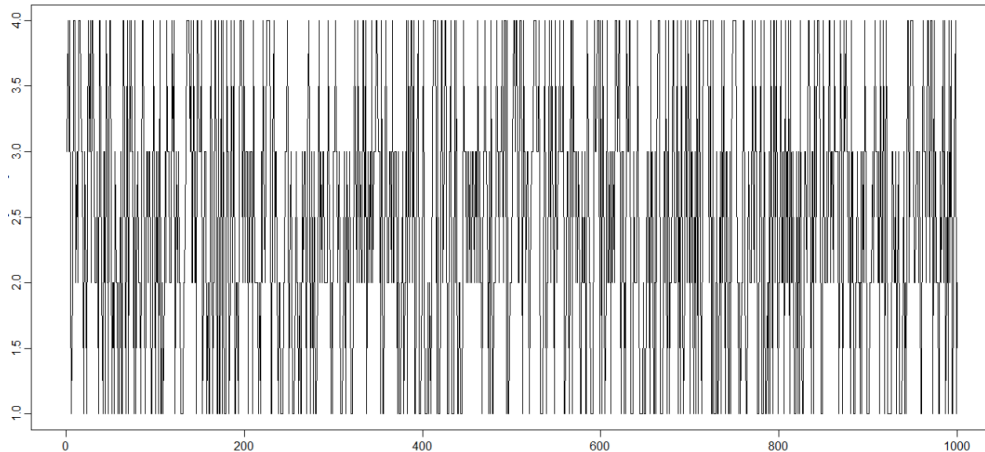


Figure (1) Representing the first 1000 data in the form of a chart where the number 1 represents the nuclide A, the number 2 represents the nuclide C, the number 3 represents the nuclide G, and the number 4 represents the nuclide T

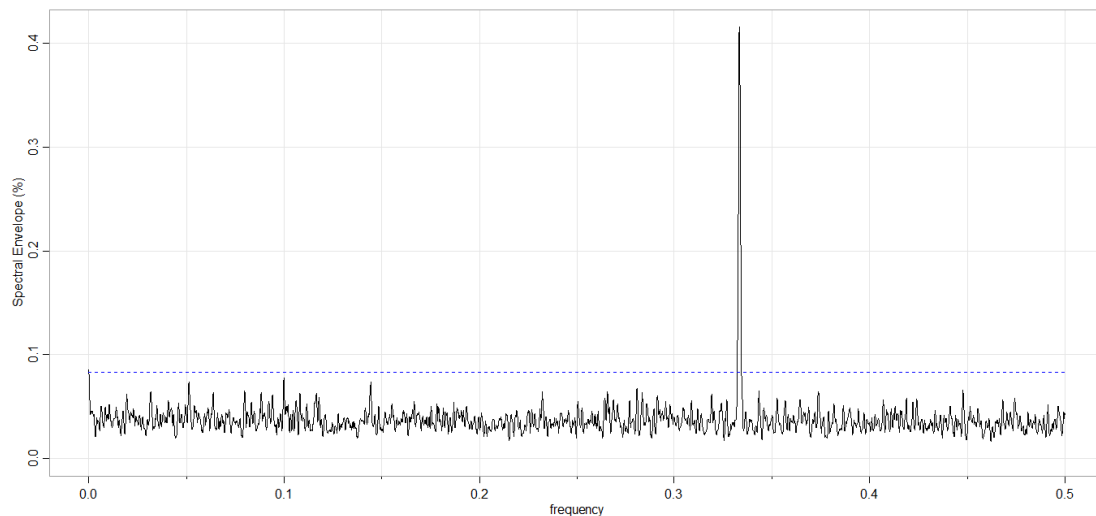


Figure (2) The SNR amount is shown under the modify Daniell method

Here we noticed a peak at frequencies $\omega_{2732} = 0.33349$ which coincide to the period of $P = 18.8407$, this reference that the data shows an approximate nine sequence cycle.

This peak indicates the possibility of a genetic mutation. where during a DNA test, it is stimulated to replicate and reproduce itself. When the DNA is doubled an incorrect association of the nitrogenous bases occurs in the DNA this may lead to the emergence of an unwanted genetic trait and this increases the virulence factors of bacteria or any other microscopic organism, it may lead to the emergence of a certain disease in humans or increase the microorganism's immunity to drugs.

REFERENCES

- [1] Dr. Hassan M. Ahmad, (2017), Investigation of Noise Influence on Speech/Voice Signal, *Al-Baath University Journal*, Volume 93, Issue 52.
- [2] Mortada Mansour Adduallah, (2014), Weather Forecasting in Baghdad City By Using Categorical Time Series, University of Baghdad.
- [3] Maytham Abdul Wahab Saleh, (2022), Using Nonparametric Methods to Study And Analysis Time Series With Long Memory, University of Karbala.
- [4] Abd, S. H. (2008). Propose a new function for estimating the weight Spectrum density function. *Journal of Administration and Economics*, 26(71).
- [5] Dr.Lekaa Ali Muhammad & Amir Ali Abboud, (2016), Comparison of bandwidth estimators (smoothing parameter) using kernel functions in principal component analysis, *Journal Of AL-Turath University College*, (20).
- [6] Chan, K. S., & Cryer, J. D. (2008). *Time series analysis with applications in R*. springer publication.
- [7] Jeong, H. (2012). The spectral analysis of nonstationary categorical time series using local spectral envelope (Doctoral dissertation, University of Pittsburgh).
- [8] Pérez-Cruz, F. (2008, July). Kullback-Leibler divergence estimation of continuous distributions. In *2008 IEEE international symposium on information theory* (pp. 1666-1670). IEEE.
- [9] Stoffer, D. S., Ombao, H. C., & Tyler, D. E. (2002). Local spectral envelope: an approach using dyadic tree-based adaptive segmentation. *Annals of the Institute of Statistical Mathematics*, 54, 201-223.
- [10] Stoffer, D. S. (1999). Detecting common signals in multiple time series using the spectral envelope. *Journal of the American Statistical Association*, 94(448), 1341-1356.
- [11] Stoffer, D. S., Tyler, D. E., & McDougall, A. J. (1993). Spectral analysis for categorical time series: Scaling and the spectral envelope. *Biometrika*, 80(3), 611-622.
- [12] Vandaele, W. (1988). Applied time series and Box-Jenkins models. *時系列入門*, 67-123.
- [13] WEI, W. W. (2006). Univariate and Multivariate Methods. *TIME SERIES ANALYSIS*.