American Journal of Modeling and Optimization, 2015, Vol. 3, No. 2, 50-55
Available online at http://pubs.sciepub.com/ajmo/3/2/3

DOI:10.12691/ajmo-3-2-3

Truncation Point Determination for Small Series Sizes

Salah H Abid*, Ahmed S Ahmed

Mathematics Department, Education College, Al-Mustansirya University, Baghdad, Iraq
*Corresponding author: abidsalah@gmail.com

Received June 11, 2015; Revised June 16, 2015; Accepted August 12, 2015

Abstract The problem of finding consistent estimator of spectral density function represents one of the important scientific subject in spectrum theory. There is a problem inside the above problem, which is the determination of the truncation point of the stochastic process (or equivalently, the spectral bandwidth determination). In this paper we proposed another procedure for small series sizes, since the problem is solved by Abid's method for moderate and large series sizes. We will support our results by two empirical experiments, the first one for truncation point determination of Poisson process, while the second experiments for empirical data generated from AR(2) Process.

Keywords: *spectral density function, Truncation Point Determination, Bandwidth, Abid's method, cross validation procedure, Poisson process, AR(2) process, Wilcoxon matched-pairs signed-rank test*

Cite This Article: Salah H Abid, and Ahmed S Ahmed, “Truncation Point Determination for Small Series Sizes.” *American Journal of Modeling and Optimization*, vol. 3, no. 2 (2015): 50-55. doi: 10.12691/ajmo-3-2-3.

1. Introduction

It is well known that the consistent estimate of the spectral density function from observations taken from stationary stochastic process $\{X_t\}$ is {Priestley(1983) [10]},

$$\hat{f}_T(\mathrm{w}) = \frac{1}{2\pi} \sum_{\mathrm{v}=-\mathrm{T}}^{\mathrm{T}} \hat{\rho}_{\mathrm{v}} \ \mathrm{k}_{\mathrm{T}}(\mathrm{v})\, \mathrm{e}^{-\mathrm{ivw}} \qquad (1)$$

Where, (T) is called the truncation point of the stochastic process which is also called the weight parameter or smooth parameter or bandwidth parameter, $\hat{\rho}_{\mathrm{v}} = \hat{\mathrm{R}}_{\mathrm{v}} \,/\, \hat{\mathrm{R}}_0$ is the estimated autocorrelation function from the sample where,

$$\hat{\mathrm{R}}_{\mathrm{v}} = \frac{1}{N} \sum_{t=1}^{N-|v|} (x_t - \bar{x})(x_{t-|v|} - \bar{x})$$

is the Auto covariance function estimated from the sample, and $\mathrm{k}_T(v)$ is the weight function called lag window function works as weighting function of autocorrelations in the time domain. The Fourier transform of the lag window function results the spectral window function as,

$$\mathrm{K}_{\mathrm{T}}(\mathrm{u}) = \frac{1}{2\pi} \sum_{\mathrm{v}=-\mathrm{T}}^{\mathrm{T}} \mathrm{k}_{\mathrm{T}}(\mathrm{v})\, \mathrm{e}^{-\mathrm{iv}u}.$$

Let K(b) be a scale parameter spectral window whose corresponding lag window has characteristic exponent r and let

$$k^{(r)} = \lim_{u \to 0} \left\{\frac{1-k(u)}{|u|^r}\right\},$$

Priestley (1983) [10] defined the spectral window bandwidth is $\mathrm{B}_p = \mathrm{C}\left[k^{(r)} / T^r\right]^{1/r}$, Where C is a constant to be determined. Priestley take the regular window on ($-\pi/T, \pi/T$) and determined C is $2\sqrt{6}$.

Based on the relationship between the truncation point of the stochastic process (T) and the spectral window bandwidth which is stated previously, for the purpose of separating the values of spectral density function f(w) at w_1 and w_2 for example, it must be chosen (T) large enough to make spectral window bandwidth K(b) less than the distance between w_1 and w_2, if (T) small so as to make spectral window bandwidth is greater than the distance between w_1 and w_2, the two values at w_1 and w_2 will appear compact together, then Priestly concludes that this means that if we want to show all the peaks and lows in the spectrum density function f(w) it must choose (T), so do not make spectral window bandwidth k(b) is greater than the bandwidth of narrower peak or low" for f(w), i.e. $\mathrm{B}_h \geq \mathrm{K(b)}$.

Priestley (1983) [10] summarizes the importance of our problem follows, "The design relations which we have discussed previously, all require a knowledge of the spectral bandwidth before they can be applied in practice. So, if one ask, why not estimate the spectral bandwidth from the data ? Unfortunately, no completely satisfactory method of estimated the spectral bandwidth has been discovered ".

Wei (1990) [13] stated the following about the importance of the problem under consideration, "This issue is a more crucial difficult problem in time series analysis because for a given window, there is no single criterion for choosing the optimal bandwidth ".

Abid (1994) [1,2] suggested the choice validation technique to get traditional and local bandwidth for spectral density function. The choice validation technique was based on new definition of the truncation point presented by Abid [1,2]. Abid constructed a test to determine the truncation point and derived the power of the test. A lot of simulation experiments were conducted describe the performance of his technique comparing with some other famous techniques. Araveeporn (2011) [2] presents the bandwidth selection methods for local polynomial regression with normal, epanechnikov, and uniform kernel function. Comaniciu (2013) [5] present a mean shift-based approach for local bandwidth selection in the multimodal, multivariate case. Slaoui (2014) [12] propose an automatic selection of the bandwidth of the recursive kernel estimators of probability density function defined by the stochastic approximation algorithm.

2. Methods of Truncation Point Determination

Following some famous methods to solve the problem of Truncation Point Determination according to different bases,

2.1. On the Basis of Preliminary Information about the Spectral Bandwidth

De Jong (1988) [6] stated that the possess of some information prior to estimate f(w) about the spectral bandwidth is rare. Actually, there is some of dissatisfaction here, because of the trial and error beside the lost time of the repetition process. There are two cases of the preliminary information exist about the spectral bandwidth, (a) when the number of observations is infinite. (b) when the number of observations is finite.

2.2. On the Basis of the Absence of Preliminary Information about the Bandwidth Spectrum

Priestley (1983) [10] stated three methods accordingly, (a) by using autocorrelation function. (b) the window closing method. (c) choosing T as a fixed percentage of N. (d) Cross–Validation Procedure(CVP) based on the likelihood function.

The CV Procedure has been suggested by Beltrao and Bloomfield (1987) [4]. They are based on the relative mean square error criterion, MSE = E $[(\hat{f}(w) - f(w))/ f(w)]^2$. For the purpose of obtaining a criterion reflects all the properties of f(w) on the interval (o, π), to get the mean integrated square error (MISE) as,

$$\text{MISE} = \text{E}(2/(-1)) \sum_{0<w_j<\pi} \left[\left(\begin{matrix} \hat{f}(w_j, B_h) \\ -f(w_j) \end{matrix} \right) / f(w_j) \right]^2 \quad (2)$$

Beltrao and Bloomfield proved that, if one had a large size, stationary random sample $\{X_t\}$, then

$$CVLF = \sum_{0<W_j<\pi} \left[Ln\, \hat{f}_T(w_j) + (\hat{h}_T(w_j) / \hat{f}_T(w_j)) \right] + ((N-1)/4) MISE + ((N-1/2) O_P(MISE) \quad (3)$$

Thus, the choice of (B_h), which makes CVLFin a minimum will make MISE in a minimum also. This method is based on the use of CVLF to estimate the differences in the values of MISE for different values of B_h, and then we select the optimum value of B_h for which makes CVLF in a minimum value (or equivalently MISE).

Following some disadvantages to the previous methods,

1- In some methods, the value of T (or equivalently B_h) depends on the spectral window and not on the real behavior of stochastic process or at least on the a available data. The spectral window may be chosen incorrectly.

2- Some methods assume existence of preliminary information on the spectral bandwidth to determine the value of (T) and therefore, it requires a prior estimate of the spectral density function f(w), and then continued estimation of f(w) several times until reaching to the value of the spectral bandwidth. Optimal

3 - Some methods assume using some values of T and estimate f(w) at each value of T repeatedly to get the best estimation of f(w). These methods are called the methods of all possible solutions.

4- Some methods assumed that the real f (w) is known, but it is actually unknown and also essential parameters are unknowns. The appropriate lag window and truncation point are unknown. It is impossible actually to get what is unknown from other unknown.

5-The formulas which are used to calculate the biased and variance of the estimated values of the spectral density function are approximate formulas, so the determination of T and N by using these formulas will be approximately also, and so all the tests which are conducted and confidence intervals which are estimated.

6- Priestley (1972) [9] noticed that to reduce the relative error then the exponent of the exact estimation of the spectral density function will lead us to a condition that the number of observations must be large $(N \to \infty)$ which is not unavailable in many cases.

7- If the bandwidth of the spectral window is large (equivalently the truncation point is small) then the smooth is better, since we move away from the target, so the details will be not clear. If the bandwidth of the spectral window reduced gradually, then the details will appear gradually but we will lost the smoothness gradually too.

Anyway, the choice of T must be conformable with objective of the case study available.

3. The Choice Validation Procedure (CHVP) Abid [1.2]

In the previous methods we noticed that there is no one of them based on mathematical logic to determine the truncation point of the stochastic process. Priestly (1983) [10] stated that the truncation point must be determined from real data of the stochastic process. Abid (1991) [1] suggested a procedure based on some properties of the stochastic process which is controlled on the phenomenon under consideration. The idea of this procedure could be clarified through his definition of the truncation point of the stochastic process, since he defined it as a separating point between keeping the stochastic process a way from its essential properties which are distinguish it from other process and non-away. By using the above principle, this procedure can be explained as follows, Let $\{X_t\,; t = 0, \pm 1, \dots \dots\}$ be a stochastic process, then

$$X_{i,j} = \mu_j + e_{i,j}, \text{ i = j, j+1, ……., n}$$

$$\text{j = 1, 2, ……..,}$$

Where, Var $e_{i,1} = \sigma_1^2$ and $Ee_{i,1} = 0$. Then the steps of Abid's procedure are, a- When j=1 then $x_{i,1}$ (i= 1,2, ….,n) is a set of all observations with mean μ_1 and variance σ_1^2. so, we normalized the values of observations without lost the generality as follows, $\frac{X_{i,j}-\mu_1}{\sigma_1} = \frac{\mu_j-\mu_1+e_{i,j}}{\sigma_1}$, Then, one can write,

$$Z_{i,j} = m_j + V_{i,j}, \text{ i = j, j+1, ……., n}$$

$$\text{j= 1, 2, …….., n,}$$

Where, $\hat{m}_j = (\mu_j - \mu_1)/\sigma_1, Z_{i,j} = \frac{X_{i,j}-\mu_1}{\sigma_1}, V_{i,j} = e_{i,j}/\sigma_1$.

Actually, when j = 1 then the normalized values will formed a Gaussian stochastic process with mean zero and variance one. Also, when j=2, 3,…,n, then $Z_{i,j}$ will formed the lags of process $Z_{i,1}$.

b- Abid used theorem (15) which is stated in Mood, Graybill and Boes (1987) [8] to test whether the stochastic process Z_t and it's lagged process Z_{t-s} will formed a Bivariate Gaussian process or not. if Z_t and Z_{t-s} represent a Bivariate Gaussian process, then Z_t will be a Univariate Gaussian Process and also Z_{t-s} .

The idea behind the mention of this theorem is for as long as Z_t and Z_{t-s} obey to Bivariate Gaussian Process, then we ensure that the process did not go away from it's properties distinguish it accurately.

c- Test whether (Z_t, Z_{t-s}) (s = 1, 2, ……, n-1) are Bivariate Gaussian Process or not

$$H_O : \underline{m} = \underline{m}_0,$$

where, $\underline{m_0} = \begin{bmatrix}0\\0\end{bmatrix}, \underline{m} = \begin{bmatrix}m_1\\m_\tau\end{bmatrix} = \begin{bmatrix}0\\m_\tau\end{bmatrix}, \tau = 2, 3, \ldots, n.$

If H_O is rejected at a certain value of τ, let it be (τ= T) then T is the truncation point for the stochastic process.

d –To test hypothesis in c, Abid [1] used modified Hotelling test with the following test statistic under H_O,

$$\frac{N(N-T)}{2N-T}(\underline{\hat{m}}-\underline{m}_O)' \ \Psi^{-1} (\underline{\hat{m}}-\underline{m}_O) \sim \frac{2N-T-2}{2N-T-3} 2F_{2,2N-T-2} \quad (4)$$

Where, S is the estimated standard deviation, $\mathbf{\Psi}$ = Q / (2N –T -2) and

$$\text{Q} = \begin{bmatrix} 1 & \hat{\rho}\, S\,[Z_{t-r}] \\ & S^2[Z_{t-r}] \end{bmatrix}.$$

Abid(1994) [2] proved in Theorem (3-1) that the Choice Validation Criterion to determine the truncation point for the stochastic process T is, C $[C.V.]^{-2}$ / (1 – $\hat{\rho}^2$) ≥ $F_{2,2N-T-2,\alpha}$, Where, C = N(N-T) (2N-T-3) / (2 (2N-T)) and $\hat{\rho}$ is the autocorrelation function between Z_t and Z_{t-r} and C.V. is the coefficient of variation. He also proved in Theorem (3-2) that The power of the test at a certain value for the noncentrality parameter λ and the significant level α is,

$$\text{Po}(\lambda) = 1 - \theta/\pi - \sum_{j=0}^{\infty} b_j \sin(j\theta) \quad (5)$$

Where, $\theta = Cos^{-1}[\,2(1 - 2Ln\alpha\,/\,((\lambda+1)^{-1}\lambda + 1)(2N-T-3))^{-1} - 1]$

$b_1 = 2\,\pi^{-1}(\text{a} - \text{b})\,(a + b)^{-1}$

$b_2 = \pi^{-1}(a + b)^{-1} \quad (a + b + 1)^{-1}\,[\,2(a-b)^2 - (a+b)(a+b-1)]$

$b_{j+2} = (a+b+j+1)^{-1}$ [2(a- b) (j+1) b_{j+1} + (j + 1 – a – b) j b_j]/(j + 2)

j= 1, 2, …, and, a = (1 / 2) (2N-T -2), b = $(\lambda+1)^2$ / (2 λ +1)

In Corollary (3-1) Abid (1994) [2] proved that The power of the test is monotonic, non-decreasing in λ with fixed α and 2N –T-2 and Monotonic, non-decreasing in α with fixed λ and 2N-T-2.

There is an equivalence between CVP and CHVP, since De Jong (1988) [6] stated that, according to the cross validation procedure (CVP), every time one observation excluded from the sample which is containing n of observations and then build a model based on the (n-1) remaining of observations. Based on that model, we forecast for the observation which is excluded in the beginning. Then repeat this process for all observations. So, if we have n of observations $y_1, y_2, \ldots, y_n$ represent the elements of vector $\underline{y}$ with mean E $\underline{y}$ = $\underline{\mu}$ = X $\underline{b}_o$ and variance covariance matrix Var($\underline{y}$) =P, then $\underline{Z} = P^{-1/2}(\underline{Y} - \underline{\mu})$ will be a vector of normalized observations. Dejong (1988) [6] proved that the cross validation criterion is,

$$\left[\frac{(\underline{Y}-\underline{\mu})'P^{-2}(\underline{Y}-\underline{\mu})}{[\,\text{tr}\,(P^{-1})]}\right]^2 = \underline{Z}'P^{-1}\,\underline{Z}\,/\,[\,\text{tr}\,(P^{-1})]^2 \quad (6)$$

Where tr is the trace of a matrix. Now, let we have n of means of samples, $\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_n$ represent the elements of vector $\underline{\tilde{m}}$ with mean $E\,\underline{\tilde{m}} = \underline{m}_0 = \underline{0}$ and variance covariance matrix Var ($\underline{\tilde{m}}$) = Ψ , then according to (3. 21) we will get the cross validation criterion to be,

$$(\underline{\tilde{m}})'\,\Psi^{-2}(\underline{\tilde{m}})/[tr\,\Psi^{-1}]^2 \quad (7)$$

The formula in (7) is equivalent {Abid (1994) [2]} with the choice validation criterion in (4) if,

1. $\underline{\tilde{m}} = \Psi^{-1/2}\,\underline{\tilde{m}}$.
2. n = 2.
3. $\left[\text{tr}\left(\Psi^{-1}\right)\right]^2 = \left(\frac{2(2N-T)}{(2N-T-2)}\right)\Big/\left(\frac{N(N-T)}{(2N-T-3)}\right)$

Based on the choice validation criterion, Abid (1994) [2] concluded that the distribution of the truncation point of the stochastic process is doubly truncated geometric in zero and N from left and right respectively. This is because of that variable comes from the fail event (reject the hypothesis) after T of success events (accept the hypothesis), with the fact that there is no fail in zero since there is no test Basically and that the maximum value for the variable is N-1 which is the last lag for the stochastic process. The distribution of the truncation point of the stochastic process is,

$$\text{F}(\hat{\text{T}}) = \frac{Pq^{\hat{\text{T}}}}{q(1-q^N)} = \frac{Pq^{\hat{\text{T}}-1}}{1-q^N}, \hat{\text{T}} = 1, 2, \ldots\ldots, (N-1) \quad (8)$$

where P is the fail probability (the probability of the hypothesis rejection), which is equal to the power of the test stated in (5). $\hat{B}_{pr} = \eta\ \hat{B}_h$, $0 < \eta < 1$.Where $\hat{B}_{pr}$ is the spectral bandwidth and $\hat{B}_h$ is the spectral window bandwidth, then

$$\hat{B}_h = (2\sqrt{6}\,/\,\eta\,\hat{\text{T}})\,[\,k^{(u)}]^{\frac{1}{u}} = A^*/\hat{\text{T}},$$

where, $A^* = (2\sqrt{6}\,/\,\eta\,)\,[\,k^{(u)}]^{1/u}$

Then Abid (1994) [2] concluded that the distribution of the spectral bandwidth is,

$$f(\hat{B}_h) = \frac{P q^{(A^*/\hat{B}_h)-1}}{1-q^N}, \hat{B}_h = A^*, A^*/2, \ldots, A^*/(N-1) \quad (9)$$

If one divided the total number of observations N to d of blocks, where (N= d N'), then Abid (1994) [2] derived the distribution of observations number in each block N' for using the Fast Fourier Transform (FFT) is,

$$f(\widehat{N'}) = \frac{P q^{A^*\widehat{N'}-1}}{1-q^N}, \widehat{N'} = \frac{C}{A^*}, \frac{2C}{A^*}, \ldots, \frac{2(N-1)}{A^*} \quad (10)$$

where $B_h = C / N'$ and C is a scalar

By using the test criterion which rewritten below, $L_i = \frac{1}{1-P_i^2}(c.v_i)^{-2} \sim F_{2,2N-T-2}$, i = 1,2,………, (n-1), Abid (1994) [2] estimated T through the data of the random variable L_i by using moments methods to be,

$$\widehat{T} = \frac{(2N-4)\,\hat{\mu}_f - 2N+2}{\hat{\mu}_f - 1} \quad (11)$$

Actually, one can get different estimates of T by substituting $\hat{\mu}_f$ by each value observation entered in the calculation of L_i, since we get from that the local estimation for the truncation point at every test we conducted, as follows

$$\widehat{T}_i = \frac{(2N-4)\,L_i - 2N+2}{L_i - 1} \quad i = 1,2, \ldots\ldots\ldots, (N-1)$$

The local estimation of the truncation point T of the stochastic process will indicate to the effective importance of the information which included in the spectral density function calculation at a certain point, this is actually best. Abid (1994) [2] noted that,

1-we can get the local estimation of the spectral bandwidth as

$$\widehat{B}_{hi} = \frac{A^*(L_i - 1)}{(2N-4)\,L_i - 2N+2}, \; i = 1,2,\ldots\ldots(N-1),$$

where, $A^* = (2\sqrt{6} / \eta)\,[k^{(u)}]^{\frac{1}{u}}, 0 < \eta < 1.$

2-From equation (10) we can get the local estimation of the number of observations for each block to conduct FFT as follows

$$\widehat{N}_i' = \frac{C[(2N-4)\,L_i - 2N+2]}{A^*(L_i - 1)}, \; i = 1,2, \ldots\ldots\ldots, (N-1)$$

Now, if the stochastic process at a point v get away from its essential properties, consequently, this leads to take T=v according to CHVP. The question that imposes itself on us now, is there possibility for v > T that the stochastic process will returned to its essential properties. If the answer is yes, we will necessarily lost amount of information. Through this question, one can establish new class of spectral density functions based on terms from (N-1) to - (N-1) on the basis of effective importance, and then calculate local estimates for the truncation point parameter T. This new class will be discontinuous Abid (1994) [2],

$$f^*(w) = \sum_{v=-(N-1)}^{N-1} \delta_v\, \hat{\rho}_v\, K(v/(N-1))\, e^{-ivw} \quad (12)$$

where, $\delta_v = \begin{cases} o \text{ if the hypothesis is rejected at lag v} \\ 1 \text{ if the hypothesis is accepted at lag v} \end{cases}$

Thus, it is clear that the harmful behavior of the spectral density function is not only from the autocorrelation function at largest values of v ($|v| \to (N-1)$), but the biggest part is due to that the stochastic process at different values of v. Being away from its essential properties. Indeed, here there are harmful behavior, so if one included δ_v in the formula of the consistent estimator of the spectral density function to remove the harmful terms, then he can get another class of spectral density functions. This is mean that, we should not stop test if H_0 is rejected at a certain v, but continue to test the hypothesis at each value of v=1,2, ……, (N-1), and give each rejection the value zero and each accept the value one in terms of the consistent estimator formula of the spectral density function. Now, let x be the number of terms of the equation (12), which is mean that x represents the number of successes (accept the hypothesis) before getting r of failures (reject the hypothesis), So, since we have (N-1) of hypothesis tests, then X + r = N -1. Based on the above argument, Abid (1994) [2] concluded that the distribution of x as a truncated negative Binomial distribution at the upper value N with a mass function,

$$f(x) = \frac{\binom{x+r-1}{r-1} q^x}{\sum_{u=0}^{N} \binom{u+r-1}{r-1} q^u}, \; x = 0,1,\ldots\ldots, (N\text{-}1),$$

Where, q = 1- p = 1- Po (λ) is the probability of hypothesis acceptance.

4. A Practical Suggestion for Small Samples Sizes

The CHVP which is suggested by Abid (1991,1994) [1,2] is very important procedure for large and median sample sizes since this procedure is based on parametric test. If the samples sizes are small or nearly small, then the procedure results becomes questionable. To solve this problem, we will suggest Wilcoxon signed- rank test instead of the parametric test which is exists in Abid's studies. It is well known that the Wilcoxon matched-pairs signed- rank test(WM PS-R)is a non-parametric statistical hypothesis test used to compare between two related samples, so it is very appropriate for our problem since we are comparing every time between the original series and one of its lags. The test procedure [11] is as follows, let N^*=N- i, (i=1,2,……,N-1) be the series size (the number of pairs). Thus there are a total of $2N^*$ points. For k=1,2,…., N^*, let $y_{1,k}$ and $y_{2,k}$ denote the measurements, so the hypothesis can be written as H_0 = the series and its lag represent the same population

H_1 = the series and its lag represent two different population. The steps of (WM-PS-R) test are,

1) For k=1,2,….., N^*, calculate $|y_{2,k} - y_{1,k}|$ and $\text{sgn}(y_{2,k} - y_{1,k})$, where sgn is the sgn function.

2) Exclude pairs with $|y_{2,k} - y_{1,k}|$=0. let N^*_r be the reduced sample size.

3)Order the remaining N^*_r pairs from smallest absolute difference to largest absolute difference.

4) Rank the Paris, starting with the smallest as 1. Ties receive a rank equal to the average of the ranks they span. Let R_k denote the rank.

5) Sum all positive values W^+ and all negative values W^-.

6) The sum with the smaller absolute becomes the test statistic W.

7) If W < W_{α,N^*_r} we reject H_0, where W_{α,N^*_r} is the critical value of W for significant level α and N^*_r.

5. Some Empirical Studies

To support our suggestion, two empirical studies were performed. The first study related with Poisson process while the other study with AR(2) process..

5.1. Truncation Point Determination of Poisson Processes

Let $\{X_t\ ;\ t \geq 0\}$ be a Poisson process which satisfies the following properties,

(i) X(0) = 0. (ii) for $0 \leq t_1 \leq \cdots \ldots . \leq t_n, X_{t_2} - X_{t_1}, \ldots\ldots, X_{t_n} - X_{t_{n-1}}$ are independent of each other's. (iii) For t, h > 0, the distribution of $X_{t+h} - X_t$ is independent of t. (iv)The number of arrival events in time interval of length h, will distributed as Poisson with mean qh is equal to its variance, and pmf,

$$P_n(h) = P\ (X_{t+h} - X_t = n) = \frac{e^{-qh}(qh)^n}{n!}, n = 0,1, \ldots$$

The covariance function of Poisson process for 0 ≤ s ≤ t, can be derived as, $R_x(t,s) = COV(X_s, X_t) = COV(X_s, X_s) + COV(X_s, X_t - X_s) = qs$. In general, one can write (Hoel et al. 1972) [7], $R_x(t,s) = q\ Min\ (s,t), t, s \geq 0$. It is clear that the Poisson process is not stationary from the second order (weakly stationary).

Actually, Poisson process can be transformed to be weakly stationary by taking the first order differences $Y_t = X_{t+1} - X_t$, then the process $\{y_t\ ;\ t \geq 0\}$ is weakly stationary with mean q and covariance function

$$R_y(r) = \begin{cases} q(1 - |r|), & |r| < 1 \\ 0, & |r| \geq 1 \end{cases}$$

Now, let us simulate a stationary Poisson process with q= 0.5, q=1.5 and sample sizes n=7,15,20,25 with run size k=1000, to study the truncation point determination by using the choice validation method and compare it with T=1, which is the value of theoretical truncation point of Poisson process. The power of the test is also studied for the choice validation method under α = 0.01,0.05 significant levels and q =1.2 with T=2 as second order pure moving average Poisson model, $X_t = \alpha_t - \alpha_{t-1} - \Upsilon\ \alpha_{t-2}$. The explanation of the results is as follows,

1) In Table 1 we recorded the results of applying the choice validation method to determine the truncation point,

Table 1.

	n			
q	7	15	20	25
o.5	1.13	1.08	1.02	1
1.5	1.24	1.11	1.04	1

It is clear from Table 1 above that the method is very powerful to determine the truncation point (T).When q decreases, the method powerful will increases, this is due to the homogeneity increases, which is increased clarity of the essential properties of the process.

2) The power of the test is calculated according to the following fact,

$$\text{Power of the test} = 1 - \frac{\#\,accept\ \ H_o\ when\ it\ is\ false}{1000}$$

The results recorded in the following Table 2.

Table 2.

n	7		15		20		25	
γ	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.01$	$\alpha = 0.01$	$\alpha = 0.01$	$\alpha = 0.01$	$\alpha = 0.01$
0.01	0.01	0.05	0.01	0.05	0.01	0.051	0.011	0.052
0.10	0.014	0.062	0.021	0.081	0.033	0.120	0.103	0.255
0.20	0.033	0.111	0.081	0.215	0.191	0.432	0.718	0.896
0.30	0.076	0205	0.224	0.456	0.578	0.846	0.997	1
0.40	0.250	0.520	0.884	0.957	0.998	1	1	1
0.5	0.640	0.815	0.960	0.979	1	1	1	1
0.7	0.906	0.976	0.989	1	1	1	1	1
0.9	0.958	0.993	1	1	1	1	1	1
1.1	0.991	1	1	1	1	1	1	1

It is clear from the above table that

1) If α and γ are fixed, then the power of the test is increased if n is increased.

2) If n and γ are fixed, then the power of the test is increased if α is increased.

3) If n and α are fixed, then the power of the test is increased if γ is increased.

4) If we have two different samples such as n=7,15 then the power of the test for n=7 and α =0.05 is smaller than the power of the test for n=15 and $\alpha = 0.01$ at the small values of γ and vice versa.

5.2 AR(2) Model

21 observations were generated by using simulation according to the following AR(2) model, $Z_t = a_1 Z_{t-1} + a_2 Z_{t-2} + e_1$, Where $a_1 = 0.3\ and\ a_2 = 0.5$ and e_t distributed as standard normal distribution. These Observations are in Table 3 below,

Table 3. 21 Observations generated based on above AR(2) Model

1	0.707945	8	0.334274	15	1.29915
2	0.080243	9	2.1533	16	2.92652
3	-0.12145	10	0.28238	17	1.53618
4	-0.5926	11	1.78571	18	0.744128
5	-0.73577	12	0.953411	19	0.891331
6	0.594315	13	0.882954	20	0.397791
7	1.76534	14	2.10628	21	0.182253

Figure 1 below represents series in Table 3. The essential moments of the series were as follows, the coverage is 0.865413, the standard deviation is 0.9341, the Skewness is 0.356 and the Kurtosis is -0.167. The estimated overall Kolmogorov statistic is 0.13 and the approximate significance level is 0.999997, so under α =0.01, 0.05, 0.1 < 0.999997, the series distributed normally.

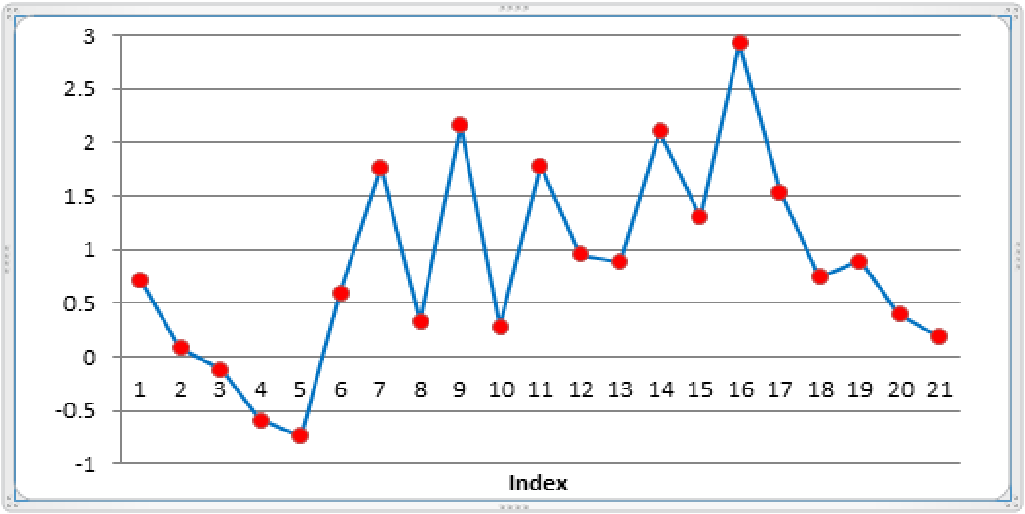

Figure 1. graph of the series in Table 3

By using the nonlinear estimation method, one can estimate the parameters a_1 *and* a_2 as $\hat{a}_1 = 0.32$ *and* $\hat{a}_2 = 0.504$, with estimated white noise standard deviation (SE) = 0.876 and Chi-square test statistic (Portmanteau χ^2) on first 10 residual autocorrelation =5.3084 with probability of a large value given white noise = 0.724, so the model is crossed over the diagnostic checking test since $\alpha = 0.01$, 0.05, 0.1 is less than 0.724. According to the above results, one can use the following formula (Priestley p. 241), of the spectral density function of AR (2) Process

$$f(w) = \frac{(1-a_2)[(1-a_2)^2-a_1{}^2]}{2\pi\,(1+a_2)[(1-a_2)^2+a_1^2+2a_1(1+a_2)\cos(w)+4a_2\cos^2(w)]},$$

$$-\pi \le w \le \pi$$

Where, $a_1 = 0.32$ and $a_2 = 0.504$. Then we used our method to determine the truncation point T under significant levels 0.01 and 0.05 where the truncation points were 2 and 3 respectively. Based on the results above, we estimate the spectral density function according to formula in (3.1) and calculate the mean square error (MSE) twice, the first one between the estimates based on the formula in (3.1) and the estimates based on the formula in (3.36), while the second MSE was between the estimates based on the formula in (3.1) (with local truncation point determination) and estimates based on the formula in (3.36). The results are recorded in the following table

Table 4.

Significant level α	**0.01**	**0.05**
T	3	2
MSE (traditional T)	0.0834	0.0981
MSE (local T)	0.0799	0.0964

From Table 4 one can stated that, (1) MSE decreases when α decreases. This is may be due to that an information will be more since T is increasing. (2) The estimation of the spectral density function based on local truncation points is better than the estimation based on the traditional truncation point. This fact is clear from the MSE values.

References

[1] Abid, S, (1991)" Some Aspects Of Spectral Analysis For The Stochastic Process With Practical Applications", Unpublished Msc Thesis, University Of Baghdad.

[2] Abid, S, (1994) "determination of the truncation point for the stochastic process with practical applications" unpublished PhD thesis, University of Baghdad.

[3] Araveeporn, A. (2011)" The Comparison Of Bandwidth Selection Methods Using Kernel Function", KMTTL Sci. Tech. J. Vol. 11, No.2.

[4] Beltrao, K. and Bloomfield, P. (1987) "Determining the Bandwidth of a kernel spectrum estimate", Time series analysis Journal, Vol.8, No.1.

[5] Comaniciu,D. (2003)"Transctions On Patern Alaysis And Mechine Intelligence ", IEEE Vol 25, No. 2.

[6] De Jong, P. (1988) "A cross-validation filter for time series models", Biometrica, Vol.75, No.3.

[7] Hoel, P. & Port, S. and Stone, C. (1972) "Introduction to stochastic process", Houghton Mifflin company, USA.

[8] Mood, A. & Graybill, F. and Boes, D. (1987) "Introduction to the theory of statistics", McGraw-Hill company, USA.

[9] Priestley, M. (1972) "Non-parametric function fitting", JRSS, Ser.B., Vol.34, No.3.

[10] Priestley, M. (1983) "Spectral analysis of time series", Academic press, London.

[11] Sheskin, D. (2000)" Handbook Of Parametric And Non-Parametric Statistical Procedures Second Edition", CRC Press, USA.

[12] Slaoui, Y. (2014). "Bandwidth Selection For Recursive Kernel Density Estimation Defined By Stochastic Approximation Method", Journal Of Probability And Statistics, Vol. 2014, P.11.

[13] Wei, W. (1990). "Time Series Analysis", Addison Wesley Company, USA.