7 (2): 161-164 (2023)

Journal of Aviation

https://dergipark.org.tr/en/pub/jav

e-ISSN 2587-1676

On Stationarity of Variance Calculation Series

Ongun Yücesan[1*]

[1*]Atilim University, School of Civil Aviation, Avionics Department, 06830, Gölbaşı, Ankara, Türkiye. (ongun.yucesan@atilim.edu.tr)

Article Info

Received: 29 August 2022
Revised: 14 March 2023
Accepted: 30 March 2023
Published Online: 22 June 2023

Keywords:
Stationarity
Heteroscedasticity
Variance
Mean
Mean Time Between Failure (MTBF)
Reliability

Corresponding Author: *Ongun Yücesan*

RESEARCH ARTICLE

https://doi.org/10.30518/jav.1168121

Abstract

While making reliability observations, more samples mean one can make a statistically representative prediction. It is possible to model the failure arrival characteristics statistically using this knowledge. As a natural product of many experiments, a mean and variance figure can be identified for modelling the different occurrences. Even though the different situations can be modelled with such parameters, it may not wholly outline the condition of the product being developed and under test. The variance calculation series derived from the original reliability observation series, which is normally used for simple variance calculation, can be an important consideration. This consideration is rarely encountered. With a mean and a variance figure, a statistical prediction can be made. However, with the very same parameters, another reliability characteristic possessing product or a subcomponent may exist. For this instance, identifying whether the variance calculation series has stationarity and incorporating it in calculations can yield a possible prediction of a more accurate statistical model. In this study, the variance calculation series is considered for their stationary character at hand and is shown to possess such character yielding further modelling possibilities and emphasizing the importance of this consideration.

1. Introduction

Reliability predictions are accurate based on relevant observations and experiments on the final product or software artifact. These observations can target to identify a Mean Time Between Failure (MTBF) figure. Experimentation is lengthy, yet the higher number of occurrences results in better statistical significance. An essential factor in an MTBF observation is how consistent the product is. For example, the calculation of a variance figure can measure consistency. This variance or standard deviation can be used along a Gaussian Random Variable for predictions. With these in mind, one can easily predict the Reliability of the system considered.

The predictions based on sole mean value would lack the variances of the conditions altering or different scenarios. Thus including the variance figure makes up for these factors. Nevertheless, the variance figure that is this important on the deviation of outcomes cannot be just judged by mean squared differences. Then it may be wise to consider the individual samples of this calculation for their variability as well. It requires little intuition to see this need; however, if we had five ordered samples as in the following set A = {3, 4, 5, 6, 7}, our sample variance would be 2.5. If we consider another set B = {3, 3, 5, 6, 6.2657} our sample variance would be 2.5001.

The sets are not the same first of all. The mean of set A is 5, and set B has a mean of 4.6531. These two series are statistically very close. The mean values are within 10% of each other. They will have very similar predictions if they are employed for any forecast employing a Gaussian Distribution. Considering data collected by some manual experiments, one would deem them equal due to human error. However, the results in set A are more uniformly distributed. The results in set B tend to be around 3 and 6 towards the edges of the domain. While considering real-life scenarios with many additional scenarios, two different products can have the same or similar mean and variances. But, one can see that the observations indicate different product characteristics. Set A would feel like it would go off any time, whereas set B would feel like it would go off when 3-unit time has passed.

A third set C = {3.4189, 3.4189, 5.0000, 6.5811, 6.5811} has a mean of 5 and a sample variance of 2.5. This set could feel more resilient once people are told not to continuously work three units of time. Predictions from set C employing the simple set theory would have a zero probability of failure in the domain [0, 3.4189). The same mean and variance results from set A would indicate a 0.2 probability of failure in this domain. Set B should indicate a 0.4 probability of failure within this domain with set theory results. Therefore, when an aviation system of taking responsibility for 500 or more people is combined with thousands of additional operational hours, the difference would relate to safe operational durations.

For all the things mentioned, the stationarity of the time series is calculated from squared deviations off the mean figure can be a measure of consistency. This study will refer to the series form squared deviations as *The Variance Calculation Series*. Such measures addressing the change in variability of results are among the few studies attempting to measure the consistency of variations.

The motivation for the philosophical work is presented in the introduction section. The rest of the document includes details of literature and methods followed in Section.2 Materials and Methods. A general presentation and judgement on the outcomes are presented in section 3. Results and Discussion. The paper concludes with section 4. Conclusions.

2. Materials and Methods

For the scope of this study, a series from an MTBF observation and build-up effort was also employed within the journal paper, which detailed the testbed as well (Yucesan et al., 2021 and 2022). In a brief explanation, this testbed performs data communication between two entities over a Local Area Network (LAN). In addition, the server is on an Embedded PC that supplies data. However, observations indicated that the query activity could no longer be continued without further reset after a while. The gathering of this query repeat amount data over consecutive times, forms the time series. This activity mimics an industrial workshop internet of things (IIoT) environment.

This workshop activity takes place in a very similar fashion over the Ethernet to newly developed inner-plane communications systems Avionics Full-Duplex Switched Ethernet (AFDX) / Aeronautical Radio Incorporated (ARINC) 664. Further, some probe rocket systems with a low budget (Matevska, 2020) incorporate the Open Platform Collaborations – Unified Access (OPC UA). However, the general system under test (SUT) in this study should be seen as a passive cooled embedded data server working under certain environmental conditions.

The stationarity controls in this study was considered by the employment of the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) (Kwiatkowski, 1992), Augmented Dickey-Fuller Test (ADF) (Dickey, 1979), and an autocorrelation result. The homoscedasticity (Kipinski, 2011) of the time series indicates that the standard deviation is statistically independent of the previous time samples or constant. The Engle arch test (Engle, 1988) was employed to check this property. All these tests are performed by MATLAB built-in functions. These considerations will reveal whether the variance calculation series is stationary or Independent Identically Distributed (IID). Such a result will ultimately indicate whether a regular random pattern that can be modelled with a known classical distribution accepting a reasonable error exist or not.

3. Result and Discussion

In this section, the variance calculation series and the test's results will be presented. The discussion will take place along with the presentation of the available data.

3.1. The Variance Calculation Series

As a result of the reliability observations, the series to calculate MTBF has been obtained. The series sample at hand (S_i) was removed from the mean value (μ) and squared to get the samples of the variance calculation series (V_i) as in Equation (1).

$$V_i = (S_i - \mu)^2 \quad (1)$$

The series V_i represented can be seen in Figure 1. As we can see, the series generally has a constant deviation figure. The originating series in this manner could have been told to possess homoscedastic characteristics. However, the concern in this paper is the variance calculation series. Even though the results are grouped mostly nearby, some variations when there are higher reliability occurrences exist. They can be considered as a cost to the producer. However, they would be affecting the mean result and variance result. Still, they would not necessarily validate the IID and stationary considerations. The Gaussian and Exponential are from infinitely long domains. Therefore, any value is possible. Considering the minimum would be zero, worse deviations could have been from an absurdly adverse scenario. Thankfully the reliability observations have a minimum of 0. The mean is around 300 for the originating series visible in Figure 1.

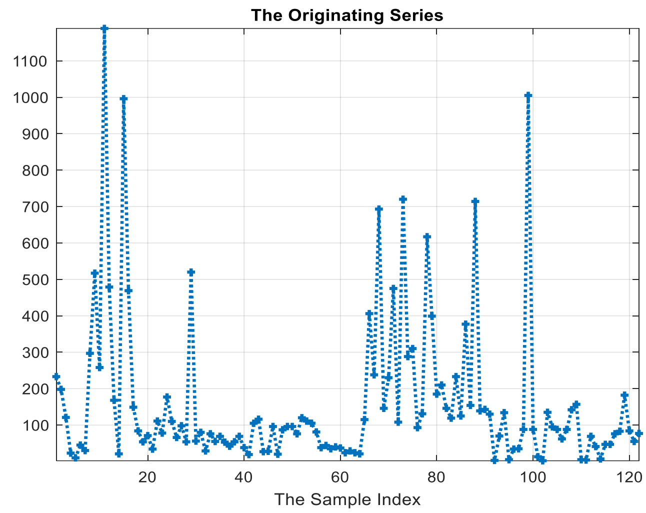

Figure 1. The originating series.

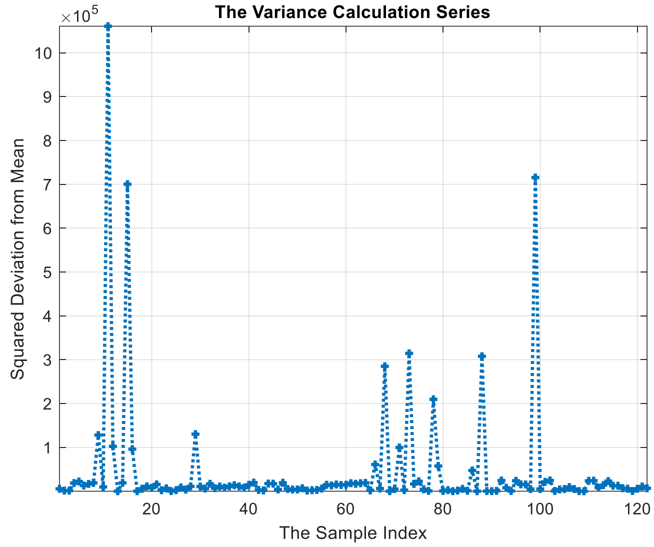

Figure 2. The Variance Calculation Series calculated from the MTBF observations.

The good periods for the duration of operation in originating series are around 70th sample till around 90th sample. However, the period around the 60th sample indicates areas of low reliability. This series was also auxiliary in the originating studies to see how bad the character could be, including all premature and awkward conditions. However, it still possesses some exponential character to findings in the writer's earlier works.

3.2. Auto-correlation Results

The Auto-correlation performed over variance calculation series result is available in the Figure 3.

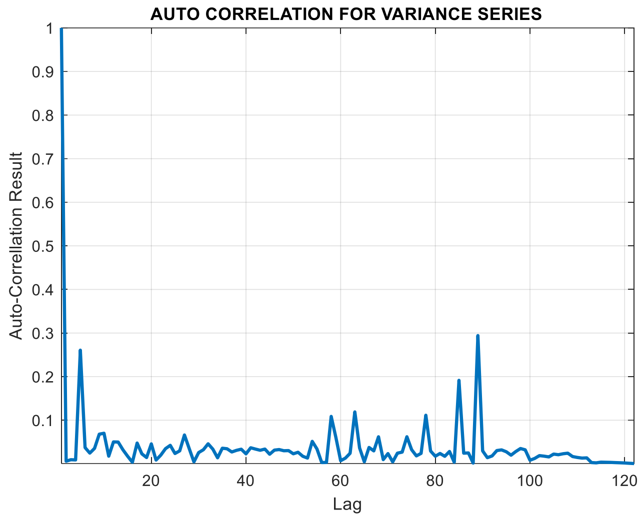

Figure 3. Autocorrelation result with regard to the lags.

As shown in Figure 3, the auto-correlation falls below a threshold and stays there. However, long-range lags around 90th lag are more correlated, possibly due to the slightly seasonal characteristic. This seasonality results are due the actual seasons changing the room temperatures and better cooling the passive cooled embedded PC. Similar seasons like November of the previous year and this year present some correlations. The result indicates the stationarity with an acceptable error.

3.3. Stationarity and Homoscedasticity Test Results

The outline of the KPSS, ADF, and Engle tests results has for the variance calculation series been reported in Table 1.

Table 1. The Stationarity and Homoscedasticity Test Results

Test	Observation
KPSS	Null hypothesis accepted
ADF	Null hypothesis rejected
Engle Test	Null Hypothesis accepted

These tests consider the well-known Auto-Regression (AR) or Auto-Regression Moving Average (ARMA) sort conditions. The KPSS indicates that the null hypothesis is accepted in favor of the existence of an AR equation statistically matching the series at hand. The ADF result rejected the opposing null hypothesis that unit root exists and an AR series statistically matching the series at hand does not exist. Engle's test on the standard deviation constancy is accepted in favor of the null hypothesis that the series does not possess heteroscedastic characteristics against the existence of a complicated formulated alternative. These indicate the existence of a stationary character within the variance calculation series.

3.4. Discussion

The stationary character or IID character can indicate the existence of an underlying stochastic process. However, the character of the distribution becomes critical. As a distribution can exist, evident characteristics and predictability arrive. Deviations, on the contrary, would still become important.

An exponential random variable indicates that any failure at any moment is possible. In this case, the variability should have been limited because it does not consider a variance input. A stationarity test can pave the way to predict a pattern and can bring predictability for avoiding awkward situations by being more cautious, using a model incorporating these factors. However, predictability should not mean that this variance series can highly deviate or that the standard deviation of the random process underlying it can be high. Such would result in short possible usage periods. These would be listed at a cost to the producer. An example can be a car company declaring ten years guarantee or else a car company declaring six months guarantee. Because the cost of the guarantee is due the company and lots of failures would not result in competitive pricing. These numbers can be regarded as safe usage periods. Therefore a plane can make the way to Tokyo from Ankara, while another plane with high variability in the variance calculation series can make a voyage safely from Ankara to Istanbul. However, both can be predicted for their probability of failure. Both missions would carry similar risk factors. If the variance series or the originating series were heteroscedastic or non-stationary, a probability for reliability or failure would not be as known as classical distributions, failing to make predictions. By obtaining stationary results as reported, which was presented according to their traditional methods of reporting, illustrates such prediction is possible for at least an example.

4. Conclusion

The study reveals that underlying stationarity or IID character can be formulated for the variance calculation series at hand. Such character yields an ability to statistically model the underlying conditions by accepting reasonable error. A study considering this variability is rare. A logical way forward is to identify the underlying statistical distribution for the variance calculation series. This information can be employed for identifying a safe period of usage incorporated along the originating series made for MTBF predictions. The consideration for variance series out of such reliability observations is among the few encountered considerations in literature.

A good future work, therefore, can be the identification of a statistical distribution fitting to the variance series and identifying good methods of incorporating these two series by a statistical method. Since variability in results, even with the same statistical parameters, can affect the safety of a variety of scenarios, as explained in the introduction, it is wise to consider the variance calculation series for reliability predictions.

Ethical approval

Not applicable.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgement

Originating Series was recorded under fund by Tubitak KAMAG 110G007 project performed for Turkish Petroleum Anonymous Partnership (TPAO). However, this study had no funds or support from any governmental or non-governmental organization.

References

Dickey, D. A., and W. A. Fuller. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association. Vol. 74, pp. 427-431.

Engle, R. (1988). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica. Vol. 96, pp. 893-920.

Kipiński, L., König, R., Sielużycki, C., & Kordecki, W. (2011). Application of modern tests for stationarity to single-trial MEG data: transferring powerful statistical tools from econometrics to neuroscience. Biological cybernetics, 105(3-4), 183–195.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin. (1992) Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics. Vol. 54, pp. 159-178.

Machiwal, D., & Sharma, A (2008). Testing homogeneity, stationarity and trend in climatic series at Udaipur – a case study. Journal of Agrometeorology, 10(2), 127–135.

Marcus J Chambers, (1996). Fractional integration, trend stationarity and difference stationarity Evidence from some U.K. macroeconomic time series, Economics Letters, Volume 50, Issue 1, Pages 19-24.

Parey, S., Hoang, T.T.H. & Dacunha-Castelle, D. (2019). Future high-temperature extremes and stationarity. Nat Hazards 98, 1115–1134.

Ulrich K. Müller, (2005). Size and power of tests of stationarity in highly autocorrelated time series, Journal of Econometrics, Volume 128, Issue 2, Pages 195-213.

Yucesan, O., Özkil, A. & Özbek, E. (2021). A Reliability Assessment of an Industrial Communication Protocol on a Windows OS Embedded PC for an Oil Rig Control Application, Journal of Science, Technology and Engineering Research, 2 (2), 22-30.

Yucesan, O., Özkil, A. & Özbek, M. E. (2022). Validity of Exponential Distribution for Modelling Inter-Failure Arrival Times of Windows based Industrial Process Control Data Exchange, Journal of Science, Technology and Engineering Research, 3 (1), 1-8.

Zeybekoğlu, U. & Aktürk, G. (2022). Homogeneity and Trend Analysis of Temperature Series in Hirfanli Dam Basin. Türk Doğa ve Fen Dergisi, 11 (1), 49-58.

Cite this article: Yucesan, O. (2023). On Stationarity of Variance Calculation Series. Journal of Aviation, 7(2), 161-164.