



White Noise Time Series with Python

by Jason Brownlee on MARCH 6, 2017 in [Time Series](#)



Last Updated on August 14, 2020

White noise is an important concept in time series forecasting.

If a time series is white noise, it is a sequence of random numbers and cannot be predicted. If the series of forecast errors are not white noise, it suggests improvements could be made to the predictive model.

In this tutorial, you will discover white noise time series with Python.

After completing this tutorial, you will know:

- The definition of a white noise time series and why it matters.
- How to check if your time series is white noise.
- Statistics and diagnostic plots to identify white noise in Python.

Kick-start your project with my new book [Time Series Forecasting With Python](#), including *step-by-step tutorials* and the *Python source code* files for all examples.

Let's get started.

- **Updated Sept/2019:** Updated examples to use latest API.
- **Updated Oct/2019:** Made the check for white noise clearer (thanks Samuel Corradi)



White Noise Time Series with Python
Photo by [Dan Eckert](#), some rights reserved.

What is a White Noise Time Series?

A time series may be white noise.

A time series is white noise if the variables are independent and identically distributed with a mean of zero.

This means that all variables have the same variance (σ^2) and each value has a zero correlation with all other values in the series.

If the variables in the series are drawn from a Gaussian distribution, the series is called Gaussian white noise.

Why Does it Matter?

White noise is an important concept in time series analysis and forecasting.

It is important for two main reasons:

1. **Predictability:** If your time series is white noise, then, by definition, it is random. You cannot reasonably model it and make predictions.
2. **Model Diagnostics:** The series of errors from a time series forecast model should ideally be white noise.

Model Diagnostics is an important area of time series forecasting.

Time series data are expected to contain some white noise component on top of the signal generated by the underlying process.

For example:

$$1 \quad y(t) = \text{signal}(t) + \text{noise}(t)$$

Once predictions have been made by a time series forecast model, they can be collected and analyzed. The series of forecast errors should ideally be white noise.

When forecast errors are white noise, it means that all of the signal information in the time series has been harnessed by the model in order to make predictions. All that is left is the random fluctuations that cannot be modeled.

A sign that model predictions are not white noise is an indication that further improvements to the forecast model may be possible.

Stop learning Time Series Forecasting the *slow way*!

Take my free 7-day email course and discover how to get started (with sample code).

Click to sign-up and also get a free PDF Ebook version of the course.

Is your Time Series White Noise?

Your time series is probably NOT white noise if one or more of the following conditions are true:

- Is the mean/level non-zero?
- Does the mean/level change over time?
- Does the variance change over time?
- Do values correlate with lag values?

Some tools that you can use to check if your time series is white noise are:

- **Create a line plot.** Check for gross features like a changing mean, variance, or obvious relationship between lagged variables.
- **Calculate summary statistics.** Check the mean and variance of the whole series against the mean and variance of meaningful contiguous blocks of values in the series (e.g. days, months, or years).
- **Create an autocorrelation plot.** Check for gross correlation between lagged variables.

Example of White Noise Time Series

In this section, we will create a Gaussian white noise series in Python and perform some checks.

It is helpful to create and review a white noise time series in practice. It will provide the frame of reference and example plots and statistical tests to use and compare on your own time series projects to check if they are white noise.

Firstly, we can create a list of 1,000 [random Gaussian variables](#) using the `gauss()` function from the [random module](#).

We will draw variables from a Gaussian distribution with a mean (μ) of 0.0 and a standard deviation (σ) of 1.0.

Once created, we can wrap the list in a Pandas Series for convenience.

```
1 from random import gauss
2 from random import seed
3 from pandas import Series
4 from pandas.plotting import autocorrelation_plot
5 # seed random number generator
6 seed(1)
7 # create white noise series
8 series = [gauss(0.0, 1.0) for i in range(1000)]
9 series = Series(series)
```

Next, we can calculate and print some summary statistics, including the mean and standard deviation of the series.

```
1 # summary stats
2 print(series.describe())
```

Given that we defined the mean and standard deviation when drawing the random numbers, there should be no surprises.

```
1 count    1000.000000
2 mean      -0.013222
3 std       1.003685
4 min      -2.961214
5 25%      -0.684192
6 50%      -0.010934
7 75%       0.703915
8 max       2.737260
```

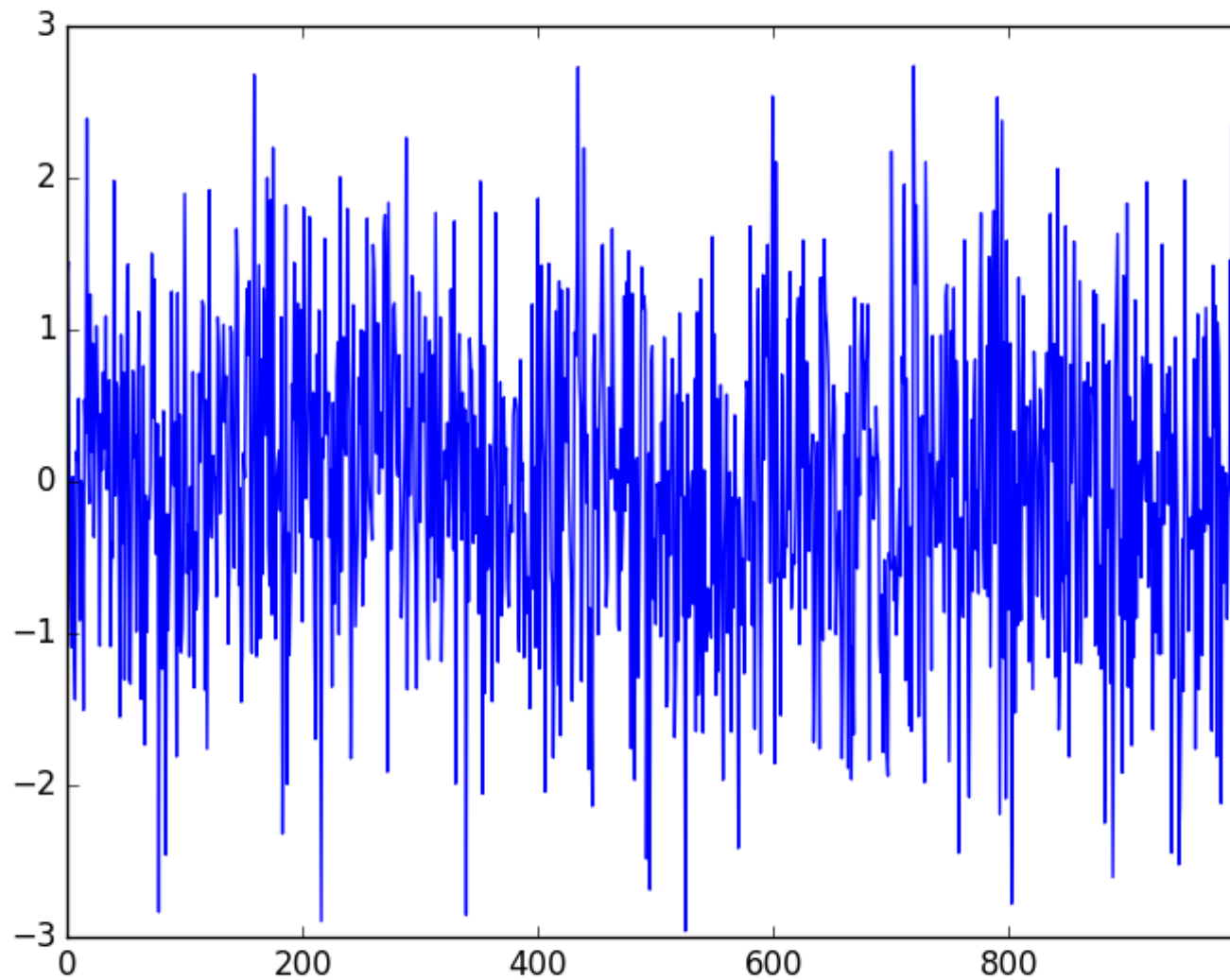
We can see that the mean is nearly 0.0 and the standard deviation is nearly 1.0. Some variance is expected given the small size of the sample.

If we had more data, it might be more interesting to split the series in half and calculate and compare the summary statistics for each half. We would expect to see a similar mean and standard deviation for each sub-series.

Now we can create some plots, starting with a line plot of the series.

```
1 # line plot
2 series.plot()
3 pyplot.show()
```

We can see that it does appear that the series is random.

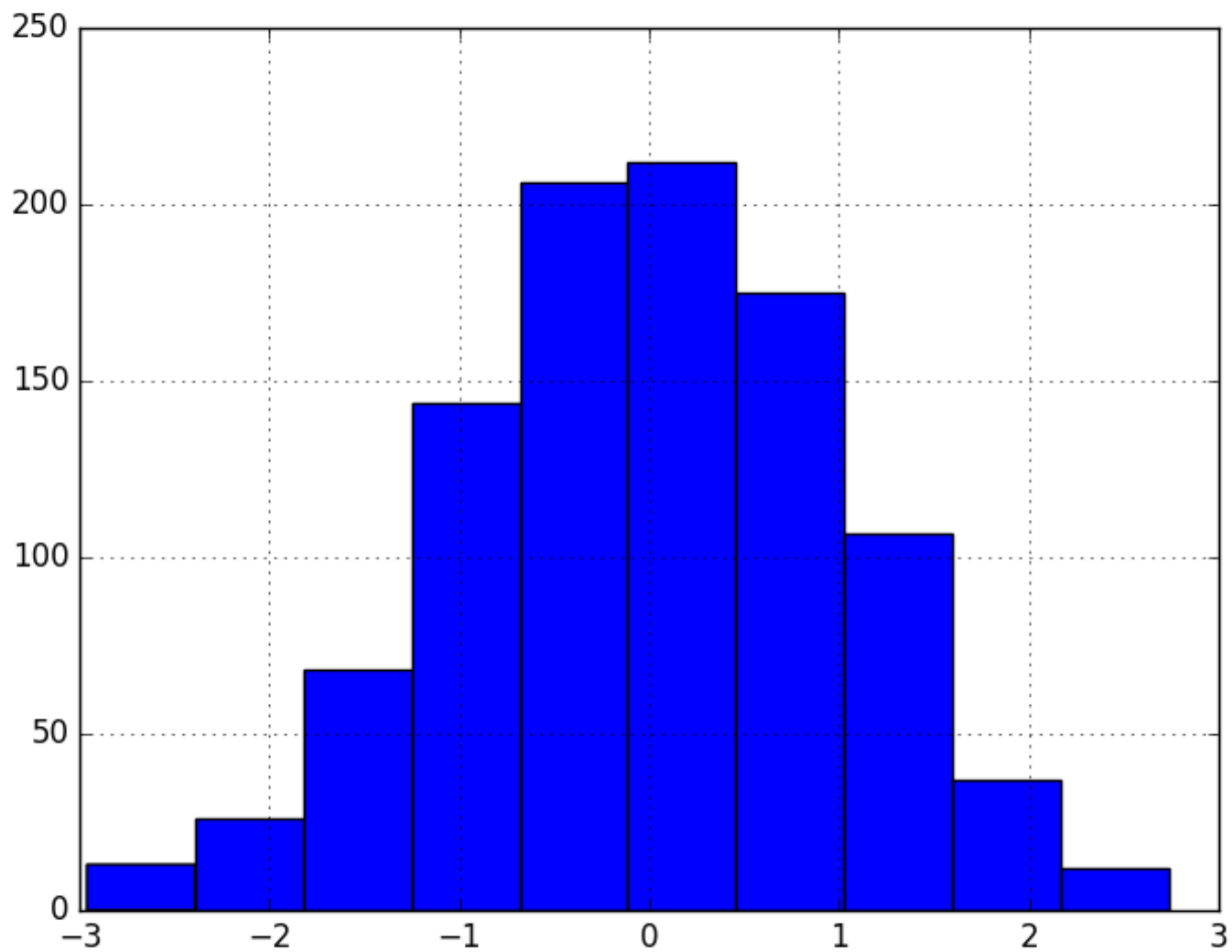


White Noise Series Line Plot

We can also create a histogram and confirm the distribution is Gaussian.

```
1 # histogram plot
2 series.hist()
3 pyplot.show()
```

Indeed, the histogram shows the tell-tale bell-curve shape.



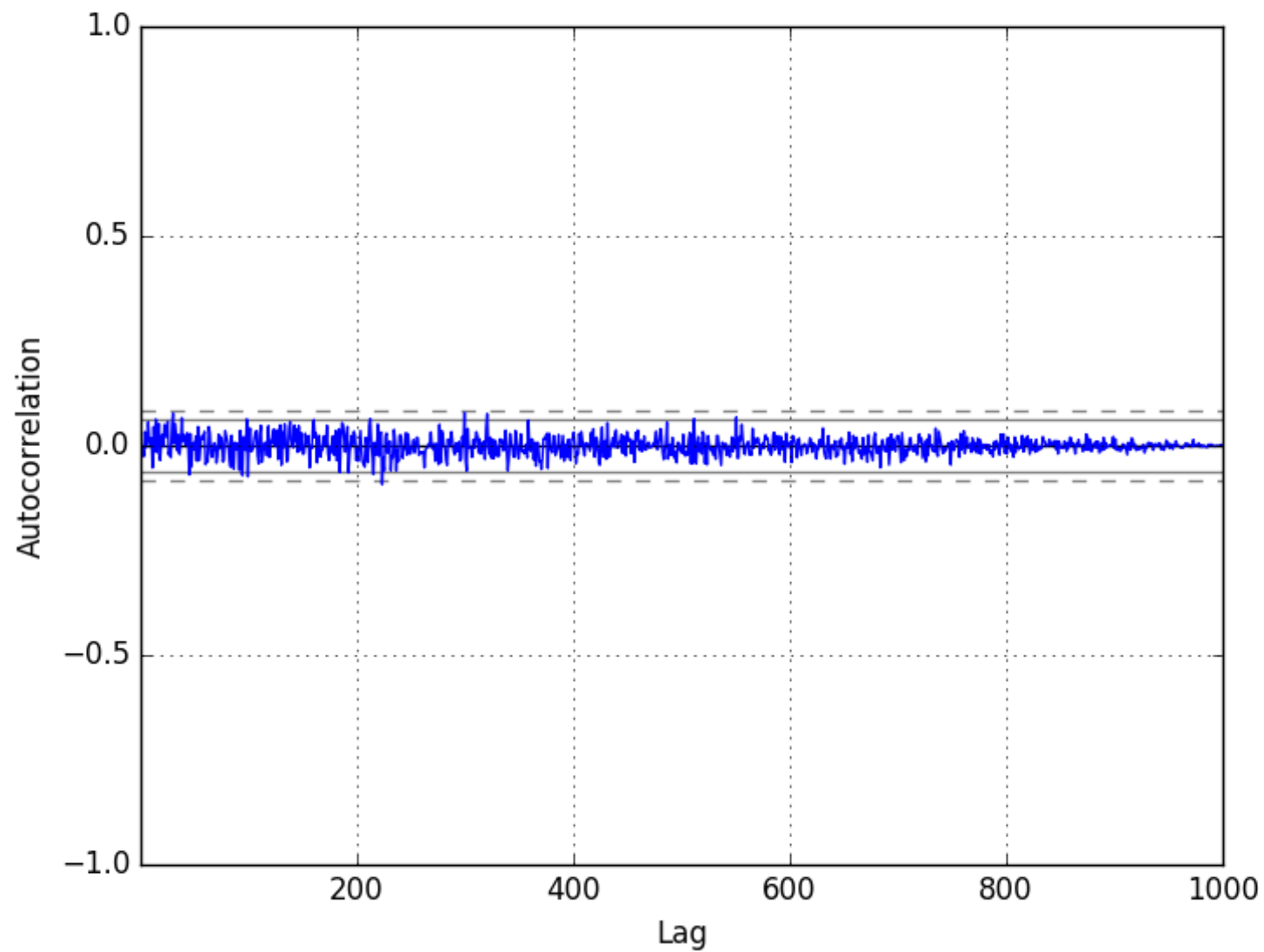
White Noise Series Histogram Plot

Finally, we can create a correlogram and check for any autocorrelation with lag variables.


```
1 # autocorrelation
2 autocorrelation_plot(series)
3 pyplot.show()
```

The correlogram does not show any obvious autocorrelation pattern.

There are some spikes above the 95% and 99% confidence level, but these are a statistical fluke.



White Noise Series Correlogram Plot

For completeness, the complete code listing is provided below.

```
1 from random import gauss
2 from random import seed
3 from pandas import Series
4 from pandas.plotting import autocorrelation_plot
5 from matplotlib import pyplot
6 # seed random number generator
7 seed(1)
8 # create white noise series
9 series = [gauss(0.0, 1.0) for i in range(1000)]
10 series = Series(series)
11 # summary stats
12 print(series.describe())
13 # line plot
14 series.plot()
15 pyplot.show()
16 # histogram plot
17 series.hist()
18 pyplot.show()
19 # autocorrelation
20 autocorrelation_plot(series)
21 pyplot.show()
```

Further Reading

This section lists some resources for further reading on white noise and white noise time series.

- White Noise Process, Page 28, [Time Series Analysis: Forecasting and Control](#).
- Section 4.2 White Noise, [Introductory Time Series with R](#).
- [White Noise](#) on Wikipedia
- [Gaussian noise](#) on Wikipedia

Summary

In this tutorial, you discovered white noise time series in Python.

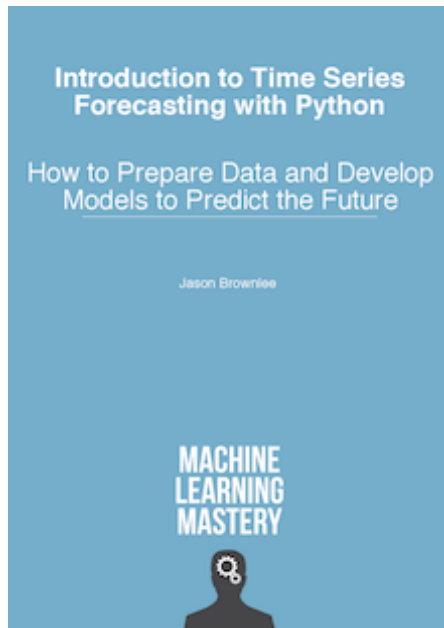
Specifically, you learned:

- White noise time series is defined by a zero mean, constant variance, and zero correlation.

- If your time series is white noise, it cannot be predicted, and if your forecast residuals are not white noise, you may be able to improve your model.
- The statistics and diagnostic plots you can use on your time series to check if it is white noise.

Do you have any questions about this tutorial? Ask your questions in the comments below and I will do my best to answer.

Want to Develop Time Series Forecasts with Python?



Develop Your Own Forecasts in Minutes

...with just a few lines of python code

Discover how in my new Ebook:

[Introduction to Time Series Forecasting With Python](#)

It covers **self-study tutorials** and **end-to-end projects** on topics like: *Loading data, visualization, modeling, algorithm tuning*, and much more...

Finally Bring Time Series Forecasting to Your Own Projects

Skip the Academics. Just Results.

SEE WHAT'S INSIDE



Tweet



Share



Share

About Jason Brownlee

Jason Brownlee, PhD is a machine learning specialist who teaches developers how to get results with modern machine learning methods via hands-on tutorials.

[View all posts by Jason Brownlee →](#)



< Python Environment for Time Series Forecasting

4 Strategies for Multi-Step Time Series Forecasting >

30 Responses to *White Noise Time Series with Python*



Mak July 14, 2017 at 7:57 pm #

REPLY ↩

Do I need to remove the trend of data before we check for white noise ? Thanks



Jason Brownlee July 15, 2017 at 9:42 am #

REPLY ↩

Yes, great question!



Narendra July 18, 2018 at 10:38 pm #

REPLY ↩

zero mean is the requirement of white noise series. But the above article says the opposite. Please clarify.



Jason Brownlee July 19, 2018 at 7:53 am #

REPLY ↩

I mention that white noise has a zero mean in the article.



Jagadish November 17, 2018 at 11:22 pm #

REPLY ↩

Hello,

I have a doubt on how do we calculate the error term in the moving average model,. in most of the sites it was mentioned that it is the difference between the yactual-yhat , however If i am trying to use the error term to find the yhat , how do I have the value of yhat, until I predict it.

Please give reference on how to calculate the error term in Moving average time series.

Thanks in advance,

Jagadish V.

Fellow at Gradvalley.in.



Jason Brownlee November 18, 2018 at 6:42 am #

REPLY ↩

In ARIMA, MA refers to modeling the series of residual errors. Learn more here:

<https://machinelearningmastery.com/model-residual-errors-correct-time-series-forecasts-python/>



rookie December 10, 2018 at 7:05 pm #

REPLY ↩

Hello! In the section of "Is your Time Series White Noise?", you list three questions for us to check if our series is White Noise and the first question is "Does your series have a zero mean?". According to the meaning of this part, if our data mean is 0, then it is not white noise. I think there is some contradiction here because you mention that the mean value of White Noise is 0 as well.

Jason Brownlee December 11, 2018 at 7:41 am #

REPLY ↩

Sorry, I don't follow. Can you please elaborate?



Momo March 31, 2019 at 3:47 am #

REPLY ↩

Hello,

I'm working on a project, I have an Event log and I want to know if it's possible to find out if it's white noise or no. So I want to know if process activities are predictable or no. If yes, how can I do that?

Thanks in advance

Jason Brownlee March 31, 2019 at 9:32 am #

REPLY ↩

Perhaps try some of the methods listed in the above tutorial?



Henry May 10, 2019 at 10:18 pm #

REPLY ↩

Hi Jason,

I'm fairly new to econometrics and time series. I have a few questions about what you wrote.

You said "This means that all variables have the same variance (σ^2) and each value has a zero correlation with all other values in the series."

I thought that not being correlated with other independent variables was a good thing since it avoid multicollinearity.

Also, you said that “Does the variance change over time?” I thought for a time series analysis, you need to have a constant variance in order to have stationarity. Correct me if I’m wrong but I thought that before doing a time series, you must convert a non-stationary data into a stationary form (perhaps by using differencing). But wouldn’t the “variance change over time” violate stationarity?

Thanks for the help! 😊



Jason Brownlee May 11, 2019 at 6:16 am #

REPLY ↩

No, in a time series, we want a correlation for a variable with prior time steps it tells us we have something linear we can learn.

Variance can change and we will have to power transform to make it stationary. But when we are looking at noise, we are checking if there is any pattern at all.



hong May 14, 2019 at 6:39 am #

REPLY ↩

Hi Jason,

Thanks for the post which is very helpful.

I understand that stationarity in data is required for forecasting. That is why we need to take the differences of the data which contains trends or seasonalities to make it stationary (constant mean, constant variance when we look at the graph). Now, the data becomes white noise (constant mean, constant variance)? But then, you said that we could not forecast a white noise series.

So it makes me so confused. Please correct me if I misunderstood something.

Thanks a lot for your help!



Jason Brownlee May 14, 2019 at 7:54 am #

REPLY ↩

The idea is that there is some signal left after the trend and seasonality is removed. If there is not, then your job is done – you can model your series with a linear trend and seasonality.



Aditya Mahajan May 14, 2019 at 9:24 pm #

REPLY ↩

So we can conclude that we need to put effort to improve our model if our error series after modelling is not a white noise .
If its white noise than we have extracted essential information from the data set and our model contains those information.



Jason Brownlee May 15, 2019 at 8:14 am #

REPLY ↩

Correct, if there is some signal remaining, then we can model it.



Tom October 9, 2020 at 3:58 am #

REPLY ↩

Hi Jason,

Let's say that we want to classify a time series data set.

Do we have a similar way (comparing forecast residuals with white noise) to know that our model can be improved ?

Because for classification, we don't have forecast residuals...

Thanks !



Jason Brownlee October 9, 2020 at 6:49 am #

REPLY ↩

For time series classification? No. You would use a confusion matrix:

<https://machinelearningmastery.com/confusion-matrix-machine-learning/>



Tom October 9, 2020 at 6:58 pm #

Thanks for the help !



Jason Brownlee October 10, 2020 at 7:02 am #

You're welcome.



Samuel Corradi October 28, 2019 at 2:28 am #

REPLY ↩

You wrote: "Your time series is NOT white noise if any of the following conditions are TRUE:
– Does your series HAVE a zero mean?"

This is right?

Thanks!



Jason Brownlee October 28, 2019 at 6:07 am #

REPLY ↩

I have made it clearer, thanks for pointing this out!



Nick February 1, 2020 at 12:55 pm #

REPLY ↩

Hi Jason,

Thanks for the post. Very naive question: I wonder if there is a way to evaluate the 'important variables' of ML models, adding white noise and testing variables robustness (if they are noise or not). If yes, any idea of R package (with a Monte Carlo simulation)?



Jason Brownlee February 2, 2020 at 6:21 am #

REPLY ↩

You're welcome.

Great thinking!

Yes, this almost the basis of permutation importance calculation:

https://scikit-learn.org/stable/modules/permutation_importance.html

I think I like your approach better, code it up!



Sarthak March 4, 2020 at 7:06 am #

REPLY ↩

Can you give some insight into increasing the frequency of data, for example, I have GDP which is yearly and I want to increase its frequency to daily. Any idea on how can I do it using noise?



Jason Brownlee March 4, 2020 at 1:31 pm #

REPLY ↩

Perhaps you can use interpolation:

<https://machinelearningmastery.com/resample-interpolate-time-series-data-python/>



Vignesh Aravind June 25, 2020 at 3:50 am #

REPLY ↩

Thank you for the explanation. i have been reading quite some time about Time Series forecasting in which majority of the papers emphasize that the forecast errors have to be normally distributed.

Speaking of forecast errors what about the white noise error term, guess white noise wouldn't be fitting into the normal distribution of errors.

Most importantly the idea for any time series forecast model is to have the error term that should be completely white noise? so that we covered all factors/relationships in predicting a variable and the only error left is white noise.

Forecast function= signal + white noise

Signal contains residuals (reducible errors) and white noise are irreducible errors?

So the normality of errors they are mentioning is only for the residual errors(reducible errors) and does not include White noise?

Can you please confirm on the above



Jason Brownlee June 25, 2020 at 6:29 am #

REPLY ↩

Good question, white noise is not modelled, it is referred to as "irreducible error".



Ashish July 14, 2020 at 4:41 pm #

REPLY ↩

Is it possible to you to give a real life example of white noise? It will be helpful for me.



Jason Brownlee July 15, 2020 at 8:13 am #

REPLY ↩

Static on the tv screen is white noise.

Leave a Reply

Name (required)

Email (will not be published) (required)

Website

SUBMIT COMMENT



Welcome!

I'm *Jason Brownlee* PhD

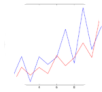
and I **help developers** get results with **machine learning**.

[Read more](#)

Never miss a tutorial:



Picked for you:



[How to Create an ARIMA Model for Time Series Forecasting in Python](#)



[How to Convert a Time Series to a Supervised Learning Problem in Python](#)



[11 Classical Time Series Forecasting Methods in Python \(Cheat Sheet\)](#)



[Time Series Forecasting as Supervised Learning](#)



[How To Backtest Machine Learning Models for Time Series Forecasting](#)

Loving the Tutorials?

The [Time Series with Python](#) EBook is where you'll find the ***Really Good*** stuff.

>> SEE WHAT'S INSIDE

© 2020 Machine Learning Mastery Pty. Ltd. All Rights Reserved.

Address: PO Box 206, Vermont Victoria 3133, Australia. | ACN: 626 223 336.

[LinkedIn](#) | [Twitter](#) | [Facebook](#) | [Newsletter](#) | [RSS](#)

[Privacy](#) | [Disclaimer](#) | [Terms](#) | [Contact](#) | [Sitemap](#) | [Search](#)