A Naïve Autocorrelation Test.

- Jaidev Kutty

Background and Motivation:

Most of the frequentist statistical methodologies developed rely on the assumption of independence and identical nature of distribution. This is the Achilles Heel of Frequentist Statistics. Data from the real world is very unlikely to be independent. Ideally we would like to understand the dependency structure and transform the data so that we can come back to a dataset which has an independence structure. Autocorrelation is one of those dependency structures which is related to time factoring an effect into the data. I came across a statistical hypothesis test which was supposed to check for autocorrelation in random numbers generated from a random number generator.

Methodology:

The test is used to check if we have a random number generator that does not generate random numbers with dependency structure. The assumption of the test is that random numbers internally have a distribution derived from a Uniform (0, 1) independent variable. The whole derivation of the test is mentioned in the book[1].

In order to generalize the test, it had to be transformed to a CDF which can be transformed into a Uniform (0, 1) variable. I initially faced challenges to scale it down for any dataset as I was not sure what the scaling function was supposed to do mathematically.
The R code is not so complicated and easy to follow. I used the "tidyverse" package.

Issues:

The test has some shortcomings and cannot work with small sample sizes after the lag and sub setting argument in the function.

Conclusion:

I see this test as an alternative to the Durbin-Watson test and can be used in that same fashion albeit the code is at present only good for vector values and not generalized to a list which is what an lm or glm statement would give.

Future Work:
1) I wanted to explore if it could show me all the possible lags and not just one lag at a time.
2) Compare the performance of this test with Durbin Watson. (In terms of the testing power).

Reference:

1. Simulation and Analysis of Industrial Systems (J.W.Schmidt & R.E.Taylor. Chapter 6)