Rolling Rates Are Too Likely to Indicate Statistically Significant Trends

By: Daniel Thompson, MPH
Accredited Professional Statistician™

August 21, 2017

Abstract

Objective: Rolling rates are often used to reduce fluctuations in rates over time. Standard statistical tests for significance assume the data are independent; however, rolling rates are not independent from each other and thus may invalidate significance testing. The purpose of this analysis was to assess the validity of the ordinary least squares (OLS) regression test for statistically significant trends when applied to 3-year rolling rates.

Methods: Infant death rates and age-adjusted overall death rates for Florida were arranged in random order for 5 years and 20 years. Annual single year rates and rolling 3-year rates were calculated for the randomly ordered data. Ordinary least squares (OLS) regression was used to test for statistically significant trends in rolling and non-rolling rates. The process was repeated 100,000 times.

Results: The statistical test for significant trends performed as expected when applied to single year rates with statistical significance indicated for close to 5% of 100,000 iterations. With the rolling rates, statistical significance was indicated for 22% to 27% of the 100,000 randomly ordered series. The results were similar for 5 and 20-year time periods and for rates and log rates.

Conclusion: The OLS regression method of testing for statistically significant trends will produce invalid results when applied to rolling rates.

Introduction

Rolling rates are computed by combining data for several consecutive time periods into one measure for each time period. For example, the 3-year rolling rate for 2017 would be computed using the data for 2015, 2016 and 2017. Likewise, the rolling rate for 2018 would be computed using data for 2016, 2017 and 2018. Rolling rates tend to dampen fluctuations in the data so the highs and lows are not extreme. This may be especially useful when the data are characterized by wide swings from one period to the next. When graphed, the rolling rates tend to show an underlying trend more clearly than the non-rolling rates.

However, problems arise when using the rolling rates in statistical tests for significant trends. Standard statistical methods assume the data are independent which means the value of each data point is not influenced by the value of any of the other data points. Rolling rates are not independent because the statistic for each data point is influenced by the values of the other data points used to compute the rolling rate.

Typically, statistical tests estimate the probability that the observed data occurred due to chance alone. For example, in testing for a significant trend in unemployment rates, the statistical test might indicate there is a 25% chance that the relationship between the time line and the unemployment rates occurred due to chance alone. Given this result, it would be reasonable to conclude there is not a trend. But if the independence assumption is not true, the statistical test may provide an inaccurate estimate of the probability that the trend occurred due to chance alone, which could lead to an unsupported conclusion.

The purpose of this analysis was to assess the validity of the OLS regression test for statistically significant trends when applied to 3-year rolling rates.

Methods

In this analysis, the OLS statistical test for trends was applied to rolling rate data and non-rolling rate, single year data. The data used were unadjusted death rates and infant death rates for Florida. These rates were arranged in random order using random sampling functions in R statistical software. Since the rates were arranged in random order, any trends were the result of chance alone. This process was repeated 100,000 times. The unadjusted death rates were used as an example of events that are not rare and the infant death rates represented rare events.

The alpha level used in the test for significant trends was 5%. With this alpha level, the test is expected to detect statistical significance in about 5% of the situations where the data are arranged in random order. In statistical terms: when the null hypothesis of no trend is true, at the 5% alpha level there is a 5% chance of getting a significant result due to chance alone.

Two time periods were used: 5 years and 20 years. For each of the 100,000 randomly arranged series, the statistical test was applied to the single year data and the 3-year rolling data. The statistical test was OLS regression test using the rate as the dependent variable and the sequence number in the series as the independent variable. If the slope coefficient in the regression equation was statistically significant at the 5% alpha level, then the trend was classified as statistically significant. In some cases, trends are not linear but are log linear. To simulate this situation, the trend test was modified to use the log of the rate as the dependent variable.

There were 16 simulations of 100,000 each. The following criteria were used: infant death rates and overall death rates; 5-year and 20–year time periods; and 1-year rates, 1-year log rates, 3-year rolling rates, and 3-year rolling log rates.

Results

The results of the simulations are in the table below. Simulations where the trend was decreasing and the P-value for the trend slope coefficient was less than 5% were classified as significantly low and are presented in the first column. Simulations where the slope coefficient P-value was 5% or greater were classified as not significant and are presented in the second column. And simulations where the trend was increasing and the P-value for the trend slope coefficient was less than 5% were classified as significantly high and are presented in the third column. The P-value in this case is the probability that the difference between the slope coefficient and zero (no trend) occurred due to chance alone.

All of the 1-year rates conform closely to the expected pattern of the statistical test. Close to 95% of the simulations for the 1-year rates were classified as not significant, with close to 2.5% classified as significantly low and close to 2.5% classified as significantly high. This is true for the 5-year trends and the 20-year trends and for the rates and log rates.

In contrast, the table shows that for the 3-year rolling rates the statistical tests indicate no significant trend in about 75% of the simulations, and significant results in about 25% of the simulations. This is true for the 5-year trends and the 20-year trends and for the rates and log rates. This is very different from the expected pattern of results for the statistical test. In statistical terms, when the null hypothesis of no trend is true, the statistical test is expected to yield a significant result about 5% of the time. In these simulations, the data were arranged in random order so the null hypothesis of no trend was true. However, when applied to the rolling rates, the statistical test yielded a significant result in about 25% of the simulations.

Conclusion

The OLS regression method of testing for statistically significant trends works well when the data in the series are independent. When the data in the series are 3-year rolling rates, the data are not independent and the OLS regression method of testing for statistically significant trends produces misleading results. Specifically, when the test is applied to rolling rates, it is much too likely to falsely detect a statistically significant trend.

This analysis compared single year rates to 3-year rolling rates for 5-year and 20-year time periods. The results for rolling rates computed with more than three years would probably be different. However, using more than three years in the rolling rates would also invalidate the independence assumption and would probably lead to more extreme results.

In conclusion, using the OLS regression method of testing for statistically significant trends with 3-year rolling rates will produce invalid results.

Results of Statistical Test for Significant Trends on 100,000 Randomly Ordered Data Sets

Data Set	Trend Number of Years	Measure	Percentage of Simulations Where Trend Significantly Low	Percentage of Simulations Where Trend Not Significant	Percentage of Simulations Where Trend Significantly High	Number of Simulations
all deaths	5	1 year rates	2.3%	95.4%	2.3%	100,000
infant deaths	5	1 year rates	2.2%	95.5%	2.3%	100,000
all deaths	20	1 year rates	2.5%	95.0%	2.5%	100,000
infant deaths	20	1 year rates	2.5%	95.0%	2.5%	100,000
all deaths	5	1 year log rates	2.8%	94.6%	2.6%	100,000
infant deaths	5	1 year log rates	2.2%	95.5%	2.3%	100,000
all deaths	20	1 year log rates	2.6%	94.8%	2.7%	100,000
infant deaths	20	1 year log rates	2.5%	94.9%	2.6%	100,000
all deaths	5	3 year rolling rates	13.5%	73.1%	13.4%	100,000
infant deaths	5	3 year rolling rates	11.9%	76.2%	11.9%	100,000
all deaths	20	3 year rolling rates	13.5%	73.1%	13.4%	100,000
infant deaths	20	3 year rolling rates	13.3%	73.2%	13.5%	100,000
all deaths	5	3 year rolling log rates	11.3%	77.6%	11.1%	100,000
infant deaths	5	3 year rolling log rates	11.8%	76.4%	11.8%	100,000
all deaths	20	3 year rolling log rates	13.4%	73.1%	13.5%	100,000
infant deaths	20	3 year rolling log rates	13.5%	73.0%	13.4%	100,000