

11-1-2009

Assessing Trends: Monte Carlo Trials with Four Different Regression Methods

Daniel R. Thompson

Florida Department of Health, dan_thompson@doh.state.fl.us

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Thompson, Daniel R. (2009) "Assessing Trends: Monte Carlo Trials with Four Different Regression Methods," *Journal of Modern Applied Statistical Methods*: Vol. 8 : Iss. 2 , Article 4.
DOI: 10.22237/jmasm/1257033780

Assessing Trends: Monte Carlo Trials with Four Different Regression Methods

Daniel R. Thompson
Florida Department of Health

Ordinary Least Squares (OLS), Poisson, Negative Binomial, and Quasi-Poisson Regression methods were assessed for testing the statistical significance of a trend by performing 10,000 simulations. The Poisson method should be used when data follow a Poisson distribution. The other methods should be used when data follow a normal distribution.

Key words: Monte Carlo, simulation, Ordinary least squares regression, Poisson regression, negative binomial regression, Quasi-Poisson regression.

Introduction

In the analysis of trend data, the key question is whether the trend reflects a true change or, alternatively, random variation. Statistical methods can be used to assess the probability that a trend has occurred due to chance. One approach is to use regression techniques to calculate the slope of the line that best fits the trend. If the slope of the line is significantly different from the flat line slope of zero, the trend is assumed to be non-random.

Disease and mortality rates generally change exponentially over time and are therefore linear in terms of the natural logarithm of the rate. Consequently, methods based on the slope of a straight line can be used to examine the natural logarithm of rates over time. The slope of the line that best fits the trend of the logarithm of the rates can also be used to calculate the estimated annual percent change

(EAPC). This is explained in more detail on the National Cancer Institute internet web page under the Surveillance, Epidemiology and End Results program (SEER) (http://seer.cancer.gov/seerstat/WebHelp/Trend_Algorithms.htm).

Several commonly used methods for assessing the statistical significance of trends exist. These methods differ in the assumptions made about the distribution of the data and in the way the slope is calculated. The Poisson regression method assumes the numerator and denominator data for the rates follow a Poisson distribution and the variances are assumed to be equal to the means. The dependent variable is the natural logarithm of the numerators with the natural logarithm of the denominators used as an offset (Dunteman & Ho, 2006). This method has been used by Liu et al to analyze trends in stroke deaths in Japan (Liu, Ikeda & Yamori, 2006); by Botha et al to analyze trends in breast cancer deaths in Europe (Botha, et al., 2001) and by Lieb et al to analyze HIV/AIDS diagnosis trends in Florida (Lieb, et al., 2007).

The Quasi-Poisson and Negative Binomial regression methods are similar to the Poisson regression method but these methods do not assume the variances are equal to the means. For more information on the Quasi-Poisson and Negative Binomial methods see Wolfram Mathworld (<http://mathworld.wolfram.com/NegativeBinomialDistribution.html>) and The R Stats Package (<http://stat.ethz.ch/R-manual/R-patched/library/stats/html/family.html>).

Dan Thompson is in the Division of Family Health Services, Bureau of Family and Community Health. He also serves as adjunct faculty at Florida State University. Note that the R programs used for this analysis are available from: Daniel Thompson, at email: dan_thompson@doh.state.fl.us. Email him at: dan_thompson@doh.state.fl.us.

ASSESSING TRENDS: MONTE CARLO TRIALS WITH FOUR REGRESSION METHODS

The ordinary least squares (OLS) regression method assumes the numerators and denominators follow a Gaussian or Normal distribution and the dependent variable is the natural logarithm of the rates. This method is recommended by the National Cancer Institute and has been used by Olson, et al. to analyze trends in incidence of Primary Central Nervous System Non-Hodgkin Lymphoma in the U.S. (Olson, et al., 2002).

When these methods are applied to randomly generated data, the probability of observing a statistically significant result should be close to the alpha level selected for the test. This is usually 0.05. The performance of these methods can be assessed by repeatedly applying them to randomly generated data and calculating the proportion of trials that result in statistical significance. If the tests are performing well in situations where the null hypothesis is true and there is no trend, this proportion should be close to the alpha level. This is generally known as a Monte Carlo experiment.

Monte Carlo experiments can also be used to assess the performance of these methods when there is a trend and the null hypothesis of no trend is false. Ideally, a method that performs well would detect a trend, when the null hypothesis of no trend is true, in about 5% of the tests; and when the null hypothesis of no trend is false, the ideal method would detect a significant trend in a relatively high proportion of the tests, compared to the other methods. In this analysis, Monte Carlo experiments were used to evaluate and compare the four methods discussed above. The objective is to provide a better understanding regarding the choice of the appropriate method for a given situation.

Methodology

R software (The R Project for Statistical Computing available at: <http://www.r-project.org/>) was used to randomly generate 10 sets of numerators and denominators. These were then used to calculate simulated rates. Random data were generated based on means and standard deviations from four different sets of numerators and denominators taken from actual statistics for the period 1996 through 2005 (Florida Community Health Assessment Resource Tool Kit (CHARTS) at:

<http://www.floridacharts.com/charts/chart.aspx>). The four data sets used were:

- 1) Injury mortality data for Florida;
- 2) Infant mortality (death before age 1) data for Florida;
- 3) Infant low birth weight (birth weight < 2500 grams) data for a Florida county; and
- 4) Infant mortality data for a Florida County.

The means and standard deviations for the numerators and denominators in these 4 data sets are given in table 1.

The data were generated to follow either a Normal (Gaussian) or a Poisson distribution. The 4 methods described in the Introduction were applied to the data sets and the results were compared. These methods were used to derive the equation that best fit the trend. The equation slope coefficient and the standard deviation of the slope coefficient were used to test for a statistically significant trend. The glm (generalized linear model) function in R was used to generate the equations.

This process was repeated 10,000 times and the proportion of trials that indicated statistical significance was compared for the 4 methods. In general, when statistical tests are applied to random data, where the null hypothesis is true, statistical significance will be observed in a proportion close to the alpha level of the test. This follows because the alpha level is defined as the probability of rejecting the null hypothesis when the null hypothesis is true. With trend data, the null hypothesis asserts there is no underlying trend and any observed trend is due to random variation. The four methods were compared in terms of their ability to accept the null hypothesis when the null hypothesis of no trend is true.

The four methods were also assessed for their ability to reject the null hypothesis of no trend when it is false. In this process the random data were generated as described above and then each succeeding simulated year of cases was increased by 1%. The formula for this simulated increase was $(1.01)^{(n-1)}$, where n is the year numbers 1 through 10. These data were generated for 10,000 simulated 10 year periods

Table 1: Means and Standard Deviations from Four Different Sets of Numerators and Denominators Taken From Florida Community Health Statistics 1996-2005

Means and Standard Deviations Used to Generate Simulated Data Sets

Data Set	Numerator Mean	Numerator Stand. Dev.	Denominator Mean	Denominator Stand. Dev.
Florida Injury Mortality	10,293.00	1,311.00	16,275,762.00	1,119,822.00
Florida Infant Mortality	1,483.00	87.80	205,609.00	11,707.00
Florida single county LBW	274.10	27.12	2,983.60	117.04
Florida single county infant mortality	27.50	5.82	2,983.60	117.04

and, as described above, the four methods were used to test for significant trends.

Results

The tables below give the results of the Monte Carlo trials. In the simulations where the null hypothesis of no trend was true, the statistical tests using OLS, Quasipoisson and Negative Binomial regression methods performed well when the data were normally distributed and also when the data followed a Poisson distribution.

As expected, with an alpha level of 0.05, approximately 2.5% of the trials reached statistically high significance and approximately 2.5% reached statistically low significance (Tables 2 through 5). In contrast, the Poisson regression method performed well only when the numerators and denominators followed the Poisson distribution. In simulations where the data followed a Normal distribution, and the null hypothesis of no trend was true, the Poisson regression method indicated statistical significance in far more than 5% of the simulations (Tables 2 through 5). The results for the Poisson method were better for the smaller data sets.

For example, in the test data set with the largest numerators and denominators (Table 2) the Poisson method indicated a significant trend in almost 90% (45.57% significantly low plus 44.24% significantly high) of the simulations where the null hypothesis of no trend was true, while the other 3 methods indicated a significant trend in a proportion close to the alpha level of 5%. The Poisson method performed better as the size of the numerators and denominators in the

test data sets became smaller. For the data set with the smallest numerators and denominators (Table 5) the Poisson method indicated significance in 8.24% of the simulations where the null hypothesis was true, which is much closer to the desired alpha level of 5%.

In the results for the simulations where the null hypothesis of no trend was false (Tables 2 through 5), three of the four methods performed about equally, when the data were normally distributed. In contrast, the Poisson regression method detected the trend in larger proportions of the trials. For example, in Table 2 for the simulations with the normally distributed data, where the null hypothesis was false, the Poisson method detected a significant trend in 68.66% of the simulations. The other 3 methods all detected a significant trend in about 8% of the simulations.

Based on these data, it appears the Poisson method is more likely to detect a trend when the null hypothesis of no trend is false, but, as shown in tables 2 through 5, the Poisson method is also more likely to detect a trend when the null hypothesis of no trend is true. In short, with normally distributed data, the Poisson method is more likely to detect a trend when a trend is present and also when a trend is not present.

When the data followed a Poisson distribution, and the null hypothesis of no trend was false, the Poisson method was more likely to detect a significant trend compared to the other 3 methods. For example, in Table 3, in the simulations where the null hypothesis is false, the Poisson method detected a trend in 94.04% of the Poisson simulations, while the other 3

ASSESSING TRENDS: MONTE CARLO TRIALS WITH FOUR REGRESSION METHODS

methods detected a significant trend in about 86% of the Poisson simulations. In contrast to the simulations of normally distributed data, the Poisson method was not more likely to detect a trend, when the null hypothesis of no trend was true, when the simulated data followed a Poisson distribution. In summary, the Poisson method performed as well as the other 3 methods when the data followed a Poisson distribution, and the null hypothesis of no trend was true. And the Poisson method was more likely to detect a trend when the null hypothesis of no trend was false and the simulated data followed a Poisson distribution.

Conclusion

These results indicate the Poisson regression method, for testing the statistical significance of rate trends, performs well only when the numerator and denominator data follow a Poisson distribution. The Ordinary Least Squares, Quasi-Poisson and Negative Binomial regression methods were more robust and performed well when the data were either Normally distributed or when they followed a Poisson distribution. When the simulation data followed a Poisson distribution and the null hypothesis of no trend was false, the Poisson regression method detected the trend more often

Table 2: Results of 10,000 Simulations of Florida Injury Mortality Rate Trends by Statistical Method and Distribution* Characteristics

Method	Test Data Distribution	Null Hypothesis: No Trend**	Percent Significantly Low	Percent Not Significant	Percent Significantly High
OLS Regression	Normal	TRUE	2.25%	95.20%	2.55%
Poisson Regression	Normal	TRUE	45.57%	10.19%	44.24%
Negative Binomial	Normal	TRUE	2.29%	95.11%	2.60%
Quasipoisson	Normal	TRUE	2.28%	95.17%	2.55%
OLS Regression	Poisson	TRUE	2.25%	95.27%	2.48%
Poisson Regression	Poisson	TRUE	2.33%	95.19%	2.48%
Negative Binomial	Poisson	TRUE	2.26%	95.26%	2.48%
Quasipoisson	Poisson	TRUE	2.27%	95.25%	2.48%
OLS Regression	Normal	FALSE	0.53%	91.60%	7.87%
Poisson Regression	Normal	FALSE	22.89%	8.45%	68.66%
Negative Binomial	Normal	FALSE	0.49%	91.71%	7.80%
Quasipoisson	Normal	FALSE	0.52%	91.73%	7.75%
OLS Regression	Poisson	FALSE	0.00%	0.00%	100.00%
Poisson Regression	Poisson	FALSE	0.00%	0.00%	100.00%
Negative Binomial	Poisson	FALSE	0.00%	0.00%	100.00%
Quasipoisson	Poisson	FALSE	0.00%	0.00%	100.00%

* Simulated 10 years of Florida injury mortality rates with randomly generated numerators at mean 10,293 and denominators at mean 16,275,762. For the random normal data, the standard deviations were 1,311 for the numerators and 1,119,822 for the denominators. For the random Poisson data, the standard deviations were the square roots of the means.

** Where Null Hypothesis of no trend = FALSE, average trend = 0.01 increase per year

THOMPSON

Table 3: Results of 10,000 Simulations of Florida Infant Mortality Rate Trends by Statistical Method and Distribution* Characteristics

Method	Test Data Distribution	Null Hypothesis: No Trend**	Percent Significantly Low	Percent Not Significant	Percent Significantly High
OLS Regression	Normal	TRUE	2.51%	94.88%	2.61%
Poisson Regression	Normal	TRUE	26.23%	46.62%	27.15%
Negative Binomial	Normal	TRUE	2.51%	94.88%	2.61%
Quasipoisson	Normal	TRUE	2.47%	94.89%	2.64%
OLS Regression	Poisson	TRUE	2.44%	94.92%	2.64%
Poisson Regression	Poisson	TRUE	2.26%	95.23%	2.51%
Negative Binomial	Poisson	TRUE	2.42%	94.93%	2.65%
Quasipoisson	Poisson	TRUE	2.43%	94.93%	2.64%
OLS Regression	Normal	FALSE	0.11%	83.87%	16.02%
Poisson Regression	Normal	FALSE	3.91%	26.50%	69.59%
Negative Binomial	Normal	FALSE	0.11%	83.87%	16.02%
Quasipoisson	Normal	FALSE	0.10%	83.98%	15.92%
OLS Regression	Poisson	FALSE	0.00%	14.48%	85.52%
Poisson Regression	Poisson	FALSE	0.00%	5.96%	94.04%
Negative Binomial	Poisson	FALSE	0.00%	14.44%	85.56%
Quasipoisson	Poisson	FALSE	0.00%	14.50%	85.50%

* Simulated 10 years of Florida infant death rates with randomly generated numerators at mean 1,483 and denominators at mean 204,609. For the random normal data, the standard deviations were 87.8 for the numerators and 11,707 for the denominators. For the random Poisson data, the standard deviations were the square roots of the means

** Where Null Hypothesis of no trend = FALSE, average trend = 0.01 increase per year

Table 4: Results of 10,000 Simulations of Low Birth Weight Rate Trends for a Florida County by Statistical Method and Distribution* Characteristics

Method	Test Data Distribution	Null Hypothesis: No Trend**	Percent Significantly Low	Percent Not Significant	Percent Significantly High
OLS Regression	Normal	TRUE	2.76%	95.09%	2.15%
Poisson Regression	Normal	TRUE	13.76%	73.39%	12.85%
Negative Binomial	Normal	TRUE	2.82%	95.01%	2.17%
Quasipoisson	Normal	TRUE	2.85%	94.95%	2.20%
OLS Regression	Poisson	TRUE	2.53%	95.02%	2.45%
Poisson Regression	Poisson	TRUE	2.92%	94.27%	2.81%
Negative Binomial	Poisson	TRUE	2.53%	95.02%	2.45%
Quasipoisson	Poisson	TRUE	2.53%	94.99%	2.48%
OLS Regression	Normal	FALSE	0.38%	88.83%	10.79%
Poisson Regression	Normal	FALSE	2.86%	57.22%	39.92%
Negative Binomial	Normal	FALSE	0.35%	88.68%	10.97%
Quasipoisson	Normal	FALSE	0.35%	88.74%	10.91%
OLS Regression	Poisson	FALSE	0.10%	75.72%	24.18%
Poisson Regression	Poisson	FALSE	0.05%	66.03%	33.92%
Negative Binomial	Poisson	FALSE	0.10%	75.71%	24.19%
Quasipoisson	Poisson	FALSE	0.11%	75.74%	24.15%

* Simulated 10 years of one Florida county low birth weight rates with randomly generated numerators at mean 274.1 and denominators at mean 2983.6. For the random normal data, the standard deviations were 27.12 for the numerators and 117.04 for the denominators. For the random Poisson data, the standard deviations were the square roots of the means.

** Where Null Hypothesis of no trend = FALSE, average trend = 0.01 increase per year

ASSESSING TRENDS: MONTE CARLO TRIALS WITH FOUR REGRESSION METHODS

than the other three methods. When the test data followed a Poisson distribution and the null hypothesis of no trend was true, the Poisson regression method performed as well as the other three methods. However, in the simulations where the null hypothesis of no trend was true and the data followed a normal distribution, the Poisson regression method was far too likely to result in statistical significance, while the other three methods resulted in statistical significance in proportions close to the alpha level of 0.05. In summary, the Poisson method performed as well or better than the other methods when the simulated data followed a Poisson distribution but did not perform as well as the other methods when the simulated data followed a normal distribution.

One of the defining characteristics of the Poisson distribution is the mean is equal to the variance. In situations where the variance exceeds the mean (this is referred to as overdispersion), Poisson regression will tend to underestimate the variance and thereby increase the probability that random results are deemed statistically significant.

Based on the results of this analysis, one recommendation is data should be examined to assess whether it follows a Poisson distribution, and the Poisson regression method should be used only when this condition is met. In practical terms, when using the Poisson regression method, the mean should be approximately equal to the variance. When this is not the case, it would probably be better to use

Table 5: Results of 10,000 Simulations of Infant Mortality Trends for a Florida County by Statistical Method and Distribution* Characteristics

Method	Test Data Distribution	Null Hypothesis: No Trend**	Percent Significantly Low	Percent Not Significant	Percent Significantly High
OLS Regression	Normal	TRUE	2.53%	95.17%	2.30%
Poisson Regression	Normal	TRUE	3.93%	91.76%	4.31%
Negative Binomial	Normal	TRUE	2.63%	95.05%	2.32%
Quasipoisson	Normal	TRUE	2.55%	95.14%	2.31%
OLS Regression	Poisson	TRUE	2.54%	95.31%	2.15%
Poisson Regression	Poisson	TRUE	2.67%	95.16%	2.17%
Negative Binomial	Poisson	TRUE	2.53%	95.34%	2.13%
Quasipoisson	Poisson	TRUE	2.57%	95.36%	2.07%
OLS Regression	Normal	FALSE	1.02%	93.54%	5.44%
Poisson Regression	Normal	FALSE	1.63%	88.25%	10.12%
Negative Binomial	Normal	FALSE	0.97%	93.60%	5.43%
Quasipoisson	Normal	FALSE	0.94%	93.69%	5.37%
OLS Regression	Poisson	FALSE	1.00%	93.39%	5.61%
Poisson Regression	Poisson	FALSE	0.88%	92.19%	6.93%
Negative Binomial	Poisson	FALSE	0.98%	93.45%	5.57%
Quasipoisson	Poisson	FALSE	0.92%	93.57%	5.51%

* Simulated 10 years of one Florida county infant mortality rates with randomly generated numerators at mean 27.5 and denominators at mean 2983.6. For the random normal data, the standard deviations were 5.82 for the numerators and 117.04 for the denominators. For the random Poisson data, the standard deviations were the square roots of the means.

** Where Null Hypothesis of no trend = FALSE, average trend = 0.01 increase per year

the OLS, Quasi-Poisson, or Negative Binomial, regression methods or a nonparametric method

This analysis addressed only trends with 10 discrete points and the test data were generated with characteristics specific to Florida infant death, Low birth weight and injury mortality data. Using more or less points and data with different distribution characteristics could, and probably would, lead to different results and conclusions. The results and conclusions from this analysis apply only to Florida low birth weight, infant death and injury mortality data or data that are very similar. A general conclusion might be that different methods perform differently depending at least in part on the characteristics of the data to which they are applied. Further research is needed to reach a better understanding of the strengths and weaknesses of these methods in various situations.

References

- Dunteman, G. H., & Ho, M-H. R. (2006). *An introduction to generalized linear models*. Thousand Oaks, CA: Sage Publications Inc.
- Liu, L., Ikeda, K., & Yamori, Y. (2001). Changes in stroke mortality rates for 1950 to 1997: A great slowdown of decline trend in Japan. *Stroke* 2001, 32, 1745-1749. Available at: <http://stroke.ahajournals.org/cgi/content/full/32/8/1745>
- Botha, et al. (2003) Breast cancer incidence and mortality trends in 16 European countries. *European Journal of Cancer*, 39, 1718-1729.
- Lieb, et al. (2007). HIV/AIDS Diagnoses among blacks - Florida, 1999-2004. *MMWR*, February 2, 2007, 56(4), 69-73.
- Olsen, J. E., et al. (2002). The continuing increase in the incidence of primary central nervous system non-Hodgkin lymphoma. *Cancer*, October 1, 2002, 95(7), 1504-1510.