The Misuse of The Vuong Test For Non-Nested Models to Test for Zero-Inflation

Paul Wilson[1]

[1] School of Mathematics and Computer Science, University of Wolverhampton, United Kingdom, WV1 1LY

E-mail for correspondence: pauljwilson@wlv.ac.uk

Abstract: The Vuong test for non-nested models is being widely misused as a test of zero-inflation. We show that such use of is erroneous and incorrectly assumes that the distribution of the log-likelihood ratios of zero-inflated models versus their non-zero-inflated counterparts is normal. We see that this stems from a mis-understanding of what is meant by the term "non-nested model", and investigate other approaches for determining zero-inflation

Keywords: Non-Nested Models; Zero-Inflation; Vuong Test

1 The Vuong Test for Strictly Non-Nested Models

Zero-inflated models are those based upon mixtures of a zero and a count distribution $f(y;\Theta)$:

$$f(y;\Theta) = \gamma + (1-\gamma)f(0;\Theta) \quad y=0; \quad (1-\gamma)f(y;\Theta) \quad y=1,2,3,\ldots \tag{1}$$

The "Vuong Test for Non-Nested Models" test was introduced by Vuong (1989), as a test for "strictly non-nested models". In slightly simplified form, it states that under the null hypothesis that two non-nested models F_θ and G_γ fit equally well, i.e. that the expected value of their log-likelihood ratio equals zero, then under H_0 the asymptotic distribution of the log-likelihood ratio statistic, LR, is normal. In particular, (under H_0):

$$LR_n(\hat{\theta}_n, \hat{\gamma}_n)/\hat{\omega}_n\sqrt{n} \longrightarrow N(0,1) \tag{2}$$

where ω denotes the variance of LR_n and n the sample size. Vuong (1989) also presents tests for nested and overlapping models, and shows that, given certain conditions, their log-likelihood ratios are related to χ^2 distributions. Due to the simplicity of its calculation, the test has become popular among statistical practitioners in various disciplines and is implementable in Stata, and the R-package *pscl* (Jackman, 2012).

2 The Misuse of the Vuong Test

The Vuong test for strictly non-nested models is being widely misused as a test of zero-inflation. For example the help page associated with the *vuong* command in *pscl* states: "*The Vuong non-nested test is based on a comparison of the predicted probabilities of two models that do not nest. Examples include comparisons of zero-inflated count models with their non-zero-inflated analogs (e.g., zero-inflated Poisson versus ordinary Poisson, or zero-inflated negative-binomial versus ordinary negative-binomial).*" Desmarais and Hardin (2013) state that:"*researchers commonly use the Vuong test (Vuong 1989) to determine whether the zero-inflated model fits the data statistically significantly better than count regression with a single equation*" and cite *ten* references to publications that have used the Vuong test for this purpose. That this is an incorrect use of the Vuong test for non-nested models is clearly illustrated by Figure 1. The left histogram illustrates the observed distribution of the log-likelihoods obtained when a one-covariate zero-inflated Poisson (ZIP) model and a Poisson model are fitted to $100,000$ samples of size $n = 100$. Clearly the distribution is non-normal. The $97.5th$ percentiles of the observed distribution is 3.45. The right hand histogram is produced using exactly the same software code that produced the left hand histogram, but here the data is simulated from data that is Poisson distributed on variables x_1, x_2, x_3 and x_4 where each x_i is uniformly distributed, and the "competing" models are on $x_1 + x_2$ and $x_3 + x_4$ respectively, and hence are strictly non-nested according to the definition of Vuong (1989). (Here the observed $97.5th$ percentile is 1.969).

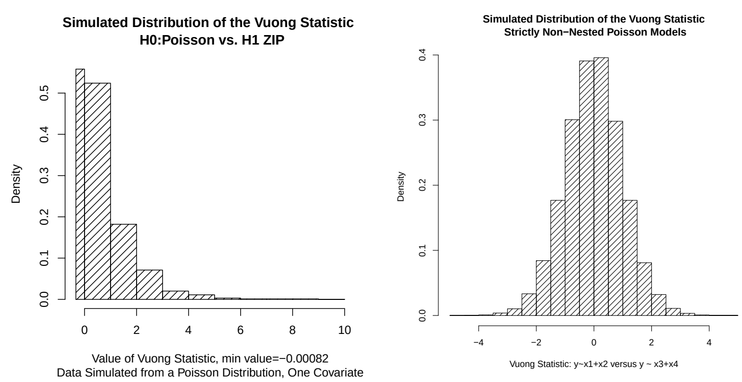

FIGURE 1. **Distributions of the Log-likelihood Ratios of ZIP versus Poisson and Strictly Non-Nested Models**

Desmarais and Hardin (2013) extensively discuss AIC and BIC type adjustments to the distribution of the log-likelihood ratios, and present evidence that this improves the power of the Vuong test; it should be noted that these adjustments are, for any given comparison of models, constants, and hence only effect the mean of the distribution, not its shape.

2.1 The Cause of The Confusion

The misuse of the test stems from misunderstanding of what is meant by the term "non-nested model". This term, along with "nested model" abounds in statistical literature. As is the case with many frequently used terms their meanings are approximately understood by many, but precisely understood by few. Clarke (2001) observes that "*Defining the concept of 'non-nested' precisely is not an easy task. Definitions are often imprecise and uncomplicated or precise and complicated.*" A statement that applies equally well to nested models. (Indeed one category is not the complement of the other, there exists a third, "inbetween" category of what Vuong (1989) refers to as *overlapping models.*) Simple definitions of nested model include that of Davison (2003): "*Two models are said to be nested if one reduces to the other when certain parameters are fixed.*" Vuong (1989) defines a model G_γ to be nested in a model F_θ by: "G_γ *is nested in* F_θ *if and only if* $G_\gamma \subset F_\theta$.
The distribution of the log-likelihood ratios of non-nested models being normal is dependent however on six assumptions presented elsewhere in Vuong (1989), these refer to various topological and measure theoretic properties. In particular whilst the standard formulation of the probability distribution function (see equation (1)) of a zero inflated Poisson distribution with zero-inflation parameter γ, where $0 \leq 0 \leq 1$, reduces to that of a Poisson distribution when $\gamma = 0$, $-2 \times LR$ fails to be χ^2 distributed as $\gamma = 0$ is at the boundary of the parameter space, failing to meet Vuong's prerequisite that it should be interior to the parameter space, this in turn results in non-normality of the sampling distribution of the zero-inflation parameter; as Vuong's subsequent theoretical development of the distributions of log-likelihood ratios (not only of non-nested models, but of nested and overlapping models also), depends upon normality of the sampling distribution of the model parameters, clearly his theory is not applicable.

2.2 Models Fitted Using Link Functions

Zero-inflated models are usually fitted using a logit link to model the expected proportion of perfect zeros. Whilst it is true that the logistic function: $\frac{\exp(t)}{1+\exp(t)} \neq 0$ for all $t \in \mathbb{R}$, and hence in some sense this formulation of the ZIP and Poisson are non-nested, $\lim_{t \to -\infty} \dfrac{\exp(t)}{1+\exp(t)} = 0$, thus this formulation of the zero-inflated model fails to meet Vuong's prerequisite that the parameter space is a compact subset of $\mathbb{R}^p$, and, similar to the scenario presented in the previous section the sampling distribution of the zero-inflation parameter, and hence the distribution of the log-likelihood ratios, is non-normal. Similar statements hold if probit or complementary log-log links are used. It is worth noting that there is confusion in the literature about models being nested if one reduces to the other if certain parameters are fixed, many authors apparently taking this to mean fixed *at zero*. For example, Desmarais and Harden (2013) state: "the count regression f is not nested in the zero-inflated model, because the model does not reduce to f (the count model) when $\gamma = 0$, in which case the probability of a 0 is inflated by 0.50", apparently alluding to the fact that the value of the logistic function $= 0.5$ when $t = 0$.

2.3 The Null Hypothesis of Vuong's Test

Consistent with a test of zero-inflation, the simulated distribution of the Vuong statistic presented in the left hand side of Figure 1 is derived by resampling from non-zero-inflated data. As stated in Section 1 the null hypothesis of Vuong's test for non-nested models is that the expected value of their log-likelihood ratios equals zero, this implies that under the null hypothesis both models are "equally far away" from the data that is being modelled. If we temporarily ignore the issue of whether zero-inflated models and their non-zero-inflated counterparts are non-nested or otherwise, and consider them non-nested, to appropriately simulate the distribution of the log-likelihood ratios it would be necessary to resample from data that was somehow equidistant from zero-inflated and non-zero-inflated data, it is difficult to envisage the nature of such data. More importantly, non-rejection of the null hypothesis of Vuong's test for non-nested models, where the (supposedly) non-nested models are, say, the zero-inflated Poisson and standard Poisson model would mean that there is no evidence to conclude that either model fits the data better than the other, not that there is no evidence to support zero-inflation, and its rejection simply implies that either the zero-inflated Poisson model fits the data better than the Poisson model, or vice-versa, not that zero-inflation is present or absent.

3 Other Approaches

Distributional Methods: Early research by the author indicates that if negative values of the log-likelihood ratio that are very close to zero are considered as zeros, then the distribution of ZIP versus Poisson log-likelihood ratios, where the zero inflation parameter is only modelled by an intercept, is a mixture of a point mass at zero and a χ^2_1 distribution, the weighting of the mixture being dependent on the number of covariates. If the zero-inflation parameter is modelled, a mixture of a zero point mass and some other distribution still occurs; whilst the nature of this other distribution is yet to be determined, a weighted mixture of χ^2 distributions is certainly a candidate, this being consistent with Vuong's theory of overlapping models, but further research is necessary. Note that when the value of γ is allowed to be both positive or negative fitted values of γ do not "pile up" close to zero, and the distribution of zero-modification parameter is normal and hence a non-zero-inflated model is *nested* in its zero-inflated counterpart, and hence a Vuong test for nested models could be used as a test of zero-inflation/deflation. Wilson (2010) presents a method of adapting zero-deflated data so that zero-modified models may be fitted via standard software for fitting zero-inflated models. The lack of software for fitting zero-modified models is due to the fact that the standard link-functions employed to fit zero-inflated models are incompatible with zero-deflation. Dietz and Böhning (2000) proposed a link function that allowed for zero-deflation, and more recently Todem, Hsu and Kim (2012) have proposed a score test that incorporates a link function that allows for both zero-inflation and deflation.

Graphical Methods: Many statistical tests for determining normality (or otherwise) of data exist, practitioners nearly invariably rather use a "normal QQ plot" to assess normality. A parallel here would be to plot the individual fitted probabilities (contributions to the likelihood) of the observed data under the zero-inflated and the non-zero inflated model against each other, if the points lie approximately along the line $x = y$, then zero-inflation is not indicated. Examples are shown in Figure 2, the left-hand diagram where 150 data have been simulated from a Poisson(1) distribution, and the lower diagram where 150 data have been simulated from zero-inflated Poisson data with Poisson mean 1.5 and zero-inflation parameter 0.3. "Jitter" has been applied in both diagrams. We see that in the top diagram the points fall approximately along the line $x = y$, consistent with lack of zero-inflation, whereas in the lower diagram, where the zero-inflated model is appropriate, this is not the case. Another approach would be to plot contributions to the log likelihood under both models against each other.

4 Conclusion

It is beyond doubt that the widespread practice of using Vuong's test for non-nested models as a test of zero-inflation is erroneous. The misuse is rooted in a misunderstanding of what is meant by the term "non-nested model". The derivation of the distribution of the log-likelihood ratios of zero-inflated versus non-zero inflated models is not straightforward, and possible alternative approaches are to develop tests based upon zero-*modified* models where the zero-"inflation" parameter may be negative, or to develop graphical methods.

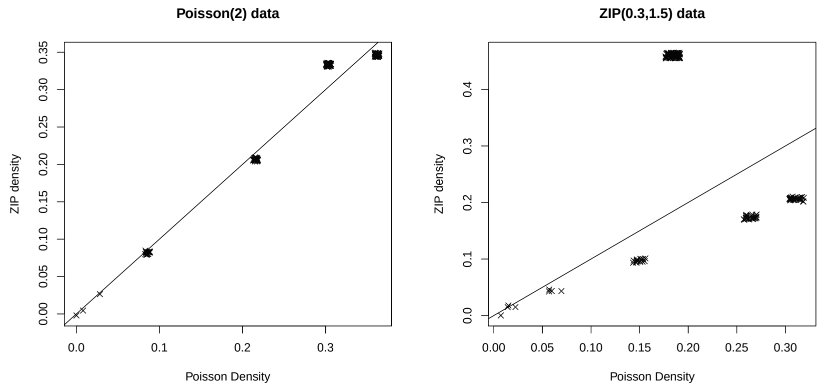

FIGURE 2. **Fitted Probabilities under both models plotted against each other.**

References

Clarke, K. (2001). Testing nonnested models of international relations:re-evaluating realism. *American Journal of Political Science*, **45**, 724 – 744.

Davison, A. (2003). *Statistical Models*. Cambridge University Press.

Desmarais, B.A., and Harden, J.J. (2013). Testing for zero inflation in count models: Bias correction for the Vuong test. *The Stata Journal*, **13**, 810 – 835.

Dietz D. and Böhning D. (2000) On the estimation of the Poisson parameter in zero-modified Poisson models. *Computational Statistics and data Analysis* **34**, 441 – 459.

Jackman, S. (2012). pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University. Stanford University. R package version 1.04.4. URL http://pscl.stanford.edu/

Todem, D., Hsu, W and Kim, K (2012) On the Efficiency of Score Tests for Homogeneity in Two-Component Parametric Models for Discrete Data. *Biometrics* **68** 975 - 982.

Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, **57**, 307 – 333.

Wilson, P. (2010) Zero Augmentation: A method for fitting zero-modified count models that allow both zero-inflation and deflation. In Bowman, A. ed. *Proceedings of the 25th International Workshop on Statistical Modelling*, University of Glasgow, 575 - 580