Semi-Parametric Specification Tests for Discrete Probability Models

Yue Fang

Abstract

Loss functions play an important role in analyzing insurance portfolios. A fundamental issue in the study of loss functions involves the selection of probability models for claim frequencies. In this article, we propose a semi-parametric approach based on the generalized method of moments (GMM) to solve the specification problems concerning claim frequency distributions. The GMM-based testing procedure provides a general framework that encompasses many specification problems of interest in actuarial applications. As an alternative approach to the Pearson χ^2 and other goodness-of-fit tests, it is easy to implement and should be of practical use in applications involving selecting and validating probability models with complex characteristics.

Introduction

Virtually all insurance problems are about the building of a mathematical model that can be used to quantify the loss function and to predict future insurance costs. The usual starting place of such practice is the search of a model for the claim frequency distribution. In many cases, when confronting a large collection of distributions from which to choose, one has to narrow the selection to a single model. The chosen model should provide a balance between simplicity and conformity to the available data.

In this article, we provide a semi-parametric approach based on the generalized method of moments (GMM) to solve the specification problems concerning claim frequency distributions. Initially, we develop the arguments in a general setting. Then, for illustrative purposes, we focus on a few important special cases. As a general framework, the GMM-based testing procedure provides much of the flexibility needed to encompass a variety of specification problems. An appealing feature of this semi-parametric approach is that it does not require complete knowledge of the distribution but only demands the specification of a set of moment conditions that the model should satisfy. Since it depends only upon moment restrictions of the model of interest, it is easy to implement even when the problem involves complex distributional forms.

Yue Fang is in the Decision Sciences department, Lundquist College of Business, University of Oregon. The author is grateful to the editor and two anonymous referees for their helpful comments on an earlier draft of the article.

The structure of the article is as follows. "Current Perspectives" outlines current perspectives of frequency distributions and commonly used specification testing methods. The next section presents the GMM-based testing procedure and briefly summarizes some of the properties of the proposed tests. "Two Illustrative Examples" demonstrates the method. "An Empirical Example" illustrates the empirical relevance of the proposed testing procedure through an application to the United Kingdom comprehensive motor policy data of Johnson and Hey (1971). The conclusion follows.

CURRENT PERSPECTIVES

Frequency Distributions

There are various possible choices of claim frequency distributions. A natural starting place is the Poisson distribution with a constant parameter λ, which measures the expected number of accidents. Early work using the Poisson distribution in the context of insurance includes Bohlmann (1909), Lundberg (1909), and Cramèr (1930). At times when the homogeneous Poisson does not adequately describe the characteristics of data, one can obtain generalized distributions in two common ways. First, one may consider mixed Poisson distributions by treating λ as the outcome of a random variable. With more flexibility in shape than the homogeneous Poisson, mixed Poisson distributions have been extensively used when the heterogeneity of risks arises. For example, the negative binomial, which is of central importance within the family of mixed Poisson distributions because of its convenient mathematical properties, has been viewed as one of the most important alternatives to the homogeneous Poisson in the classical collective risk theory (Dickson et al., 1998). The Poisson-inverse Gaussian distribution is another widely used mixed Poisson distribution. Observing that the Poisson-inverse Gaussian distribution has similar statistical properties to the negative binomial, the Poisson-inverse Gaussian may be used to model similar physical phenomena as the negative binomial. Generalizations of the Poisson-inverse Gaussian have been considered by Kestemont and Paris (1985) and Rubinstein et al. (1987), among others. But the use of generalized Poisson-inverse Gaussian distributions becomes less inconvenient because of the increased complexity of the inference due to the complicated mixing distribution and the compound nature of the problem.

Another possibility to strive for maximum flexibility in the model is to use recursive formulas to obtain families of distributions. For instance, one of the most prominent families, the Katz family (1965), whose successive probabilities satisfy first-order recurrence relations, has been studied in Klugman et al. (1998). The Katz family of distributions, which consists of the Poisson, binomial, and negative binomial distributions, forms a simple class with the property of being equi-, under-, or over-dispersed. One can also find extensions to the Katz family in the literature. They are usually derived by introducing more parameters and/or allowing more general recursive formulas with the Katz family as a submodel (see, for example, the Kemp families [1968], the Ord family [1972], the extended Katz family of Gurland and Tripathi [1975], and the modified Katz family considered by Sundt and Jewell [1981] and Willmot [1988]). Defining distributions by a recursive formula enables recursive computation of aggregate claims distributions. The hierarchical nature of the probability mass function (pmf) also allows one to calculate moments and the moment-generating function in most cases.

Selecting and Validating a Model

With more than one model choice, hypothesis testing is the primary tool to narrow the selection to a single model.[1] While various specific tests have been developed for particular models in the past, the most commonly used generic test in actuarial applications is the Pearson χ^2 goodness-of-fit test. The Pearson χ^2 test is applicable for testing data when no parameters need to be estimated. When parameters must be estimated from data, the theory of the Pearson χ^2 goodness-of-fit test becomes more elaborate. It is common practice to replace the parameters by the estimates from the data based on the maximum likelihood estimation (MLE) or other consistent methods under the hypothesized distribution. In this case, one difficulty with the Pearson statistic is that its asymptotic distribution depends on the estimators employed and may lead to difficulties connected with computation of a nonstandard limit distribution (Stuart et al., 1999). Moreover, the Pearson statistic is sensitive to grouping that is conducted to improve the χ^2 approximation. Nevertheless, we apply the Pearson χ^2 test along with our GMM-based testing procedure to the data of Johnson and Hey (1971) in "An Empirical Example" and compare the testing results.

Another class of specification tests is based primarily on likelihood functions. This class of tests may be considered as *discrimination tests*; examples include the likelihood ratio (LR) and the Wald, and the Lagrange tests. Asymptotically, all three test statistics are equivalent, although they have quite different computational requirements (Buse, 1982). These tests may complement generic testing procedures such as the Pearson χ^2 test in some applications. However, they are designed with different purposes in mind. The Lagrange approach essentially starts at the null hypothesis and examines whether movement toward the alternative would be an improvement, while the Wald test starts at the alternative and moves toward the null. The likelihood ratio method considers the two hypotheses on an equal basis. Consequently, this class of tests does not address the issue of whether a particular distribution or a family of distributions is a good fit to the data; rather, it considers the question of which one between the null and alternative provides a significantly better fit. GMM tests based on LR principles have been studied by Newey (1985) and Ahn (1995), among others. We do not examine this class of tests in this article, but we do this class of tests for reference.

THE GMM-BASED TESTING PROCEDURE

The Test Statistic

To introduce the basic idea behind GMM-based tests, it is useful to first consider the GMM estimation. Let $\mathcal{F}$ be a family of distributions with support on a subset of non-negative integers. Suppose that $\mathcal{F}$ involves a finite p-dimensional vector of unknown parameters θ. Based on a random sample, $\{x_i : i = 1, 2, \ldots, n\}$, we test the hypothesis that the sample was generated by a particular member of $\mathcal{F}$.

---

[1] As an alternative to a formal hypothesis test, a set of selection procedures based on information measures such as the Schwartz Bayesian Criterion (SBC) and the Akaike Information Criterion (AIC) has been also used in selecting and validating a model. See, for example, Klugman et al. (1998) for the use of the SBC, and Frees et al. (2001) for the AIC.

Statistical inference procedures based on GMM examine certain features of the random sample by considering the moment restrictions:

$$E[f(x_i,\theta)] = 0, \tag{1}$$

where $f(x_i,\theta)$ is a continuous q-dimensional vector function of θ and x_i with $q \geq p$.

When the number of moment restrictions in Equation (1) equals the number of parameters (i.e., $q = p$), one can find the method of moments (MM) estimator of θ by solving the analogous sample moment conditions in Equation (1). Note that if $q = p$, then Equation (1) represents a set of p equations with p unknowns. In this case, it has a unique solution for θ under certain conditions. However, when the number of moment restrictions in Equation (1) exceeds the number of parameters (i.e., $q > p$), Equation (1) represents a set of q equations in p unknowns. In this overidentifying case, the system typically does not have a solution and hence, the MM estimation is infeasible. One solution is to find the value of θ that is "closest" to satisfying Equation (1) as the estimator of θ based on a distance measurement. In GMM, 1 can be the distance measure (the objective function) to be the scalar

$$Q_n(\theta) = f_n(\theta)'V_n^{-1}f_n(\theta), \tag{2}$$

where $f_n(\theta) \equiv n^{-1}\sum_{i=1}^{n} f(x_i,\theta)$ and V_n is a consistent estimator of $V \equiv \lim_{n\to\infty} Var[n^{1/2}f_n(\theta)]$. And define the GMM estimator of θ to be

$$\hat{\theta} = argmin_{\theta\in\Theta}Q_n(\theta).$$

The objective function of Equation (2) plays a key role not only in obtaining the GMM estimate of θ but also in developing a statistical procedure of testing the validity of the moment restrictions of Equation (1). Note that when $q > p$, the GMM estimator does not set all sample moments in Equation (1) to zero. This opens up the possibility for testing $E[f(x_i,\theta)] = 0$.[2] In fact, as shown in Sowell (1996) and Hall (1999), the GMM estimation decomposes the population moment restrictions in the null hypothesis

$$H_0 : E[f(x_i,\theta)] = 0 \tag{3}$$

into

$$H_0^I : DV^{-1/2}E[f(x_i,\theta)] = 0$$

and

$$H_0^O : (I_{q\times q} - D)V^{-1/2}E[f(x_i,\theta)] = 0,$$

where $D \equiv M(M'M)^{-1}M'$ and $M \equiv V^{-1/2}E[F_n(\theta)]$ with $F_n(\theta) \equiv \partial f_n(\theta)/\partial\theta$. Under this decomposition, it is clear that $H_0 = H_0^I \cap H_0^O$. In particular, H_0^I has p moment restrictions. They are the identifying restrictions for θ, representing the part of the moment restrictions that actually goes into parameter estimation. However, the $(q - p)$ moment

---

[2] If the primary goal is the statistical testing of hypothetical models with data, one needs to have $q > p$. If $q = p$, as we have seen, $f_n(\hat{\theta}) = 0$. In this case, the magnitude in Equation (2) would simply equal zero in all samples, so it is not possible to examine the validity of the moment restrictions in Equation (1) directly.

restrictions in H_0^O are overidentifying restrictions, which are not imposed, so one can test whether they hold in the sample.

If the hypothesis of the model that led to Equation (1) in the first place is incorrect, some of the overidentifying moment restrictions will be systematically violated, providing a basis for developing a specification test. Hence, one can interpret the objective function of Equation (2) as a measure of how close the sample is to satisfying the overidentifying moment restrictions. This motivated Hansen's (1982) GMM-based test for the null hypothesis of Equation (3),

$$J_n(q) \equiv nQ_n(\hat{\theta}) = n f_n(\hat{\theta})' V_n^{-1} f_n(\hat{\theta}). \tag{4}$$

Hansen (1982) has shown that under certain regularity conditions, $J_n(q)$ has an asymptotic χ^2 distribution with $(q - p)$ degrees of freedom under the null hypothesis. We note that $J_n(q)$ is trivial to calculate because it is simply the sample size times the value of the objective function of Equation (2) evaluated at the GMM estimate of θ.

Some Properties of the Tests

Asymptotic properties of the test $J_n(q)$ have been extensively studied in the literature. One can show that $J_n(q)$ is consistent (i.e., has power one asymptotically against alternative hypotheses) and that the asymptotic power of the tests depends only on p, q, and overidentifying moment restrictions (see, for example, Hall, 1999). Fang (2000) has used two alternative approaches to study the asymptotic properties of $J_n(q)$: the local power analysis and the approximate slope method. The result shows that although additional moment conditions are used in $J_n(q)$ with a higher value q, this does not necessarily imply that large values of q are generally more desirable, even asymptotically. The result dramatically emphasizes the fact that the local power of the test may differ substantially for different values of q and across alternatives.

Results of Monte Carlo studies pertaining to the size and power of $J_n(q)$ in finite samples can be found, for example, in Christiano and den Haan (1995), Smith (1999), and Fang (2000). The Monte Carlo evidences suggest that overall, the distribution of the test is well approximated by the asymptotic theory and that the test has satisfactory performance for moderate-size samples.

Finally, the following remark can be made concerning the power comparison between $J_n(q)$ and the Pearson χ^2 test. In general, one anticipates that if the overidentifying moment restrictions capture virtually all information on the difference between the null and the alternative distributions, $J_n(q)$ will be more powerful than the Pearson χ^2 test because the degrees of freedom of $J_n(q)$ are usually much less than those of the Pearson χ^2 test. However, if the overidentifying moment restrictions fail to capture such information, the Pearson χ^2 may be more powerful. By selecting the moment restrictions of Equation (1) appropriately, $J_n(q)$ should be, in general, comparable to or more powerful than the Pearson χ^2 test.

TWO ILLUSTRATIVE EXAMPLES

This section presents two examples to demonstrate the proposed GMM-based testing procedure: the hypothesis of testing the Poisson homogeneity ($H_O^{(1)}$) and that of testing the mixing distribution in a mixed Poisson model ($H_O^{(2)}$).

Example 1: Testing the Poisson Homogeneity

Under the hypothesis $H_O^{(1)}$ of a homogeneous Poisson distribution, each x_i follows a Poisson distribution governed by the same parameter λ. This specification problem has been the subject of a number of studies. Fisher (1950) and Moran (1973), among others, suggested the use of the Poisson index of dispersion based on the first two moments of the sample:

$$\mathcal{T} = \frac{1}{\sqrt{2(n-1)}} \left[\frac{\sum(x_i - \bar{x})^2}{\bar{x}} - (n-1) \right]. \tag{5}$$

As a consistent test, $\mathcal{T}$ has certain optimal power properties under a suitable class of local alternatives. For example, Potthoff and Whittinghill (1966) have shown that $\mathcal{T}$ is asymptotically locally the most powerful among all locally unbiased tests for the mixed Poisson distribution with a gamma as the mixing distribution. It turns out, not surprisingly, that $\mathcal{T}$ can be derived as a special case of tests based on GMM under $H_O^{(1)}$.

In this simple case, $\theta = \lambda$ and $p = 1$. Choose the q-dimensional vector of moment restrictions in Equation (1) as

$$f(x_i, \lambda) = (x_i - \mu_1', x_i^2 - \mu_2', \dots, x_i^q - \mu_q')',$$

where μ_r' is the jth moment about the origin of the Poisson distribution given by

$$\mu_{r+1}' \equiv E(X^{r+1}) = \lambda \sum_{j=0}^{r} \binom{r}{j} \mu_j',$$

with $\mu_0' = 1$ and $\mu_1' = \lambda$.

In this simple case, the identifying restrictions in H_0^I are $(E(x_i) - \mu_1') = 0$ and hence, the GMM estimate of λ is $\bar{x}$. Furthermore, the overidentifying restrictions in H_0^O are $(E(x_i^j) - \mu_j') = 0$ for $j = 2, 3, \dots, q$. The interpretation of this decomposition result is that in the $J_n(q)$ statistic, one uses the first moment restriction, which is automatically satisfied by the sample, in estimating the population mean, resulting in the MME, which is also the MLE estimator of the parameter λ. The higher moment restrictions are satisfied by the Poisson distribution but not by other distributions such as the binomial distribution ($B(N,P)$) or negative binomial distribution ($NB(k,p)$). For example, the second moment restriction for binomial and negative binomial distributions is

$$0 = E(x_i^2 - \mu_2') = E(x_i^2) - (E(x_i) + [E(x_i)]^2) = \begin{cases} -NP^2, & \textit{binomial} \\ kp^2, & \textit{negative binomial}. \end{cases}$$

It is apparent that $J_n(q)$ with $q \geq 2$ has power against the violation of the overidentifying restriction for any departures from the Poisson distribution for either the binomial or negative binomial distribution.

One can obtain the explicit expressions of $J_n(q)$. For example, if $q = 2$, then $f_n(\lambda) = (\bar{x} - \lambda, m_2' - (\lambda + \lambda^2))'$, where $m_2' = n^{-1} \sum x_i^2$. The variance matrix is

$$V = \begin{bmatrix} \lambda & 2\lambda^2 + \lambda \\ 2\lambda^2 + \lambda & 4\lambda^3 + 6\lambda^2 + \lambda \end{bmatrix}.$$

Hence, the GMM test statistic is

$$J_n(2) = \frac{n}{2}\left(\frac{s^2 - \bar{x}}{\bar{x}}\right)^2 \sim \chi_1^2,$$

where $s^2 = n^{-1}\sum(x_i - \bar{x})^2$. Note that $J_n(2)$, which also depends only on the first two moments of the sample, is asymptotically equivalent to $\mathcal{T}$ in Equation (5). One can also derive tests based on higher values of q analytically along similar lines. See, for example, Fang (2003) for results in the case of $q = 3$.

Example 2: Testing the Mixing Distribution in a Mixed Poisson Model

By allowing the Poisson parameter λ itself to be a random variable with the cumulative distribution function (cdf) $U(\cdot)$, one acquires the mixed Poisson distribution

$$P_k \equiv P(x_i = k) = \int_0^\infty e^{-\lambda}\lambda^k \Gamma^{-1}(k+1)dU(\lambda), \tag{6}$$

where $U(0) = 0$. Given that the data follows a mixed Poisson distribution, one may test $H_O^{(2)}$ knowing that the mixing distribution has a specified cdf $U(\cdot)$.

Note that the mixing distribution is unobservable, because the data are drawn from the mixed distribution. The identifiability result from Douglas (1980) allows us to identify the mixing distribution based on the mixed distribution. This may not be true, however, for other mixed distributions. For example, binomial mixtures are not identifiable (Ord, 1972).

When the moments of the mixing variable exist, a necessary and sufficient condition that a distribution be a mixed Poisson is that its factorial moment[3] generating function be equal to the moment-generating function of the mixing distribution (Haight, 1967), namely

$$\mu_{(k)} = v'_k \tag{7}$$

for nonnegative integer k, where $\mu_{(k)}$ and v'_k denote respectively the kth factorial moment of the mixed Poisson variable and the kth moment about zero of the mixing variable. As will be seen, the moment relationships between the mixed Poisson and the mixing distribution in Equation (7) play a central role in developing specification tests for $H_O^{(2)}$.

Let μ'_k be the kth moment about zero of the mixed Poisson variable. Set $\theta = (\mu'_1, \mu'_2, \ldots, \mu'_p)$. Consider $f(x_i,\theta) = \Omega^{-1}g(x_i,\theta)$, where $g(x_i,\theta) = (x_i - \mu'_1, x_i^2 - \mu'_2, \ldots, x_i^q - \mu'_q)'$ and Ω is the transformation matrix such that $(\mu'_1,\mu'_2,\ldots,\mu'_q)' = \Omega(\mu_{(1)},\mu_{(2)},\ldots,\mu_{(q)})'$. The matrix Ω is lower triangular with 1 as diagonal elements, and the lower triangular elements can be obtained from

---

[3] See Stuart and Ord (1994), Section 3.7, for the definition of factorial moments.

$$\mu'_k = \sum_{j=0}^{k}(\Delta^j 0^k /j!)\mu_{(k)}, \tag{8}$$

where $\Delta^j 0^k /j!$ is the Stirling number of the second kind (Johnson et al., 1992).[4] Therefore, one can rewrite Equation (1) as moment restrictions on v'_ks:

$$E[f(x_i,\theta)] = E[(x_i - v'_1, x_i^{(2)} - v'_2, \ldots, x_i^{(q)} - v'_q)'] = 0,$$

where $x_i^{(k)} = x_i(x_i - 1)\ldots(x_i - k + 1)$. Furthermore,

$$V^{-1} = \Omega' W^{-1}\Omega,$$

where the (i,j)th element of W is $(\mu'_{i+j} - \mu'_i\mu'_j)$.

Note that by having $\theta = (\mu'_1, \mu'_2, \ldots, \mu'_p)$, the p identifying restrictions involve only moments up to the pth order. In fact, the first-order conditions for minimizing $Q_n(\theta)$ in Equation (2) imply that $\hat{\theta}$ is the solution to

$$F_n(\theta)' V_n^{-1} f_n(\theta) = 0,$$

suggesting that $\hat{\theta}$ is determined by the first p moments in Equation (1). Hence, the p identifying moment restrictions are

$$(E[x_i] - \mu_{(1)}, E[x_i^{(2)}] - \mu_{(2)}, \ldots, E[x_i^{(p)}] - \mu_{(p)})' = 0,$$

or equivalently

$$(E[x_i] - v'_1, E[x_i^{(2)}] - v'_2, \ldots, E[x_i^{(p)}] - v'_p)' = 0.$$

By contrast, the $(q-p)$ overidentifying restrictions involve higher-order moments $\mu_{(j)}$s (or v'_js) up to order q. If all p identifying restrictions are satisfied, then the overidentifying restrictions are $(q-p)$ linear combinations of

$$(E[x_i^{(p+1)}] - \mu_{(p+1)}, E[x_i^{(p+2)}] - \mu_{(p+2)}, \ldots, E[x_i^{(q)}] - \mu_{(q)})' = 0,$$

or equivalently

$$(E[x_i^{(p+1)}] - v'_{p+1}, E[x_i^{(p+2)}] - v'_{p+2}, \ldots, E[x_i^{(q)}] - v'_q)' = 0.$$

In the special case that $q = p + 1$, the overidentifying moment restriction is $E[x_i^{(p+1)}] - \mu_{(p+1)} = 0$ or $E[x_i^{(p+1)}] - v'_{p+1} = 0$. We revisit this hypothesis by examining the data of Johnson and Hey (1971) in the following section.

---

[4] For the inverse calculation of Equation (8), Ω^{-1}, which is also lower triangular, has Stirling numbers of the first kind as its elements.

TABLE 1

United Kingdom Comprehensive Motor Policies in 1968

Number of Claims	Observed Frequencies	Fitted Poisson	Fitted Negative Binomial
0	370412	369246	370460
1	46545	48644	46411
2	3935	3204	4045
3	317	141	301
4	28	5	21
5	3	–	1

Adapted from Table 2.10.1 of Beard et al. (1984).

An Empirical Example

As an illustration, we consider the data from Johnson and Hey (1971), which contain 421,240 observations. They represent United Kingdom comprehensive motor policies in 1968 that are classified according to the number of claims ranging from 0 to 5, with the average number of claims per policy being 0.13174 and the variance 0.13852. See Table 1 for the data set.

Beard et al. (1984) have shown that the Poisson distribution is a poor fit because of its short tail. They therefore use the overdispersed negative binomial distribution and anticipate that the negative binomial provides a much better fit in the tail region (see Table 1), an observation confirmed by the Pearson χ^2 test using the parameters based on MLE. According to Beard et al., the value of the Pearson χ^2 test statistic is 6.9, which gives a significance level of 14 percent for four degrees of freedom. They conclude that although there is a slight indication that the negative binomial may be underrepresenting the tail, for most applications the model may be safely used. However, Willmot (1987) has worked on the same data set and used a different grouping scheme. Willmot has reported a χ^2 statistic of 7.94 on two degrees of freedom, which has a significance level of only 1.9 percent for the data. In light of Willmot's result, one could argue that the negative binomial distribution may well not be "safely used."

Here we apply the test statistic $J_n(q)$ to assess the fit of the negative binomial model. Note that $\mathcal{F}$ consists of the mixed Poisson with two-parameter gamma

$$p(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} \quad (\alpha > 0, \beta > 0; x > 0) \tag{9}$$

as the mixing distribution. Since $p = 2$, we calculate $J_n(q)$ for $q = 3$, 4, and 5 as 3.7130, 11.3074, and 10.2359, respectively.[5] Accordingly, the test statistic with $q = 3$ is significant at 10 percent, but not at the 1 or 5 percent level. The test statistic based on $q = 4$

---

[5] Various moments for the mixed distribution (the negative-binomial) and the mixing distribution (two-parameter gamma) can be found in, for example, Stuart and Ord (1994). In particular, the jth order of the factorial moment of the negative binomial $NB(k,p)$, $\mu_{(j)}$, is given by $(k + j - 1)!p^j/(k - 1)!$, and the jth moment about zero of the two-parameter gamma in Equation (9), v'_j, equals $\beta^j \Gamma(j + \alpha)/\Gamma(\alpha)$. See also Example 2 for computational details in obtaining $J_n(q)$.

is significant at 1 percent and that with $q = 5$ is significant at 5 percent but not at 1 percent level.

Overall, the evidence is fairly strong against the negative binomial, especially in view of the result of $J_n(4)$ that the hypothesis is rejected at all conventional significance levels. The result is in complete contrast to that of Beard et al. (1984) that the Pearson χ^2 test cannot reject the negative binomial at significance levels less than 14 percent.

To further elaborate the model, we consider the three-parameter gamma with the density function given in Equation (10) as the mixing distribution:[6]

$$p(x) = \frac{(x - \gamma)^{\alpha - 1} e^{-(x - \gamma)/\beta}}{\beta^\alpha \, \Gamma(\alpha)} \quad (\alpha > 0, \beta > 0; x > \gamma). \tag{10}$$

To evaluate the fit of the use of Equation (10), we report the testing result based on $J_n(q)$ with $q = 5$. For the Johnson and Hey (1971) data, $J_n(5)$ is only 0.021 with a corresponding significance level of 99.41 percent. The fit is an almost perfect one, implying that the inclusion of one more parameter is justified. We obtain the same conclusion with other values of q.

CONCLUSION

This article proposes a semi-parametric approach to specification problems concerning claim frequency distributions. The approach has been shown to provide a useful theoretical and empirical framework for studying claim frequency distributions motivated by a great variety of applications. As an alternative approach to the Pearson χ^2 and other goodness-of-fit tests, it is easy to implement and should be of practical use in applications involving selecting and validating probability models with complex characteristics.

REFERENCES

Ahn, S. C., 1995, Model Specification Testing Based on Root-T Consistent Estimators, Discussion paper, Department of Economics, Arizona State University.

Beard, R., T. Pentikäinen, and E. Pesonen, 1984, Risk Theory (London: Chapman and Hall).

Bohlmann, G., 1909, Die Theorie des mittleren Risikos, *Transactions of the International Congress of Actuaries*.

Buse, A., 1982, The Likelihood-Ratio, Wald, and Lagrange Multipler Tests: An Expository Note, *American Statistician*, 36: 153-157.

Christiano, L., and W. Haan den, 1995, Small Sample Properties of GMM for Business Cycle Analysis, Federal Reserve Bank of Minneapolis Staff Report.

Cramèr, H., 1930, *On the Mathematical Theory of Risk* (Stockholm: Skandia Jubilee Volume).

Dickson, D., L. Tedesco, and B. Zehnwirth, 1998, Predictive Aggregate Claims Distributions, *Journal of Risk and Insurance*, 65: 689-709.

---

[6] This mixed Poisson distribution is a Delaporate distribution, which equals the convolution of a Poisson and a negative binomial (Ruohonen, 1988; Willmot and Sundt, 1989).

Douglas, J., 1980, *Analysis with Standard Contagious Distributions* (Fairland, Md). International Co-operative Publishing House

Fang, Y., 2003, GMM Tests for the Katz Family of Distributions, *Journal of Statistical Planning and Inference*, 110(1-2): 55-73.

Fang, Y., 2000, Specification Tests for Families of Distributions: A Generalized Method of Moments Approach, Discussion paper, Department of Decision Sciences, University of Oregon.

Fisher, R., 1950, The Significance of Deviations from Expectation in a Poisson Series, *Biometrics*, 6: 17-24.

Frees, E., V. Young, and Y. Luo, 2001, Case Studies Using Panel Data Models, *The North American Actuarial Journal*, 5: 24-42.

Gurland, J., and R. Tripathi, 1975, Estimation of Parameters on Some Extensions of the Katz Family of Discrete Distributions Involving Hypergeometric Functions, in: G. P. Patil, S. Kotz, and J. K. Ord, eds., *Statistical Distributions in Scientific Work, Volume 1: Models and Structures* (Dordrecht: Reidel).

Haight, F., (1967), *Handbook of the Poisson Distribution* (New York: John Wiley & Sons).

Hall, A., 1999, Hypothesis Testing in Models Estimated by GMM, in: L. Màtyàs, ed., *Generalized Method of Moments Estimation* (Cambridge: Cambridge University Press).

Hansen, L., 1982, Large Sample Properties of Generalized Method of Moment Estimators, *Econometrica*, 50: 1029-1054.

Johnson, N., S. Kotz, and A. Kemp, 1992, *Univariate Discrete Distributions* (New York: John Wiley and Sons).

Johnson, P. D., and G. B. Hey, 1971, A Review of the Scope for Using Claim Histories of Individual Policies in Risk Assessment, *Bulletin Association Royale des Actuaires Belges*, 66: 159-195.

Katz, L., 1965, Unified Treatment of a Broad Class of Discrete Probability Distributions, *Proceedings of the International Symposium on Discrete Distributions*, Montreal, 175-182.

Kemp, A. W., 1968, A Wide Class of Discrete Distributions and Associated Differential Equations, *Sankhyā*, *Series A*, 30: 401-410.

Kestemont, R., and J. Paris, 1985, Sur l'ajustement du nombre de sinistres, *Bulletin of the Association of Swiss Actuaries*, 85: 157-164.

Klugman, S., H. Panjer, and G. Willmot, 1998, *Loss Models: From Data to Decisions* (New York: John Wiley and Sons).

Lundberg, F., 1909, Zur Theorie der rückversicherung, *Transactions of the International Congress of Actuaries*.

Moran, P., 1973, Asymptotic Properties of Homogeneity Tests, *Biometrika*, 60: 79-85.

Newey, W. K., 1985, Generalized Method of Moments Specification Testing, *Journal of Econometrics*, 29: 229-256.

Ord, K., 1972, *Families of Frequency Distributions*, Statistical Monograph No. 30 (London: Griffin).

Potthoff, R., and M. Whittinghill, 1966, Testing for Homogeneity: II. The Poisson Distribution, *Biometrika*, 53: 183-190.

Rubinstein, G., W. Zucchini, and J. Juritz, 1987, Parameter Estimation for the Sichel Distribution and its Multivariate Extension, *Journal of the American Statistical Association*, 82: 938-944.

Ruohonen, M., 1988, A Model for the Claim Number Process, *ASTIN Bulletin*, 18: 57-68.

Sowell, F., 1996, Optimal Tests for Parameter Instability in the Generalized Method of Moments Framework, *Econometrica*, 64: 1085-1108.

Stuart, A., and J. Ord, 1994, *Kendall's Advanced Theory of Statistics*: *Vol. 1, Distribution Theory* (New York: Oxford University Press).

Stuart, A., J. Ord, and S. Arnold, 1999, *Kendall's Advanced Theory of Statistics*: *Vol. 2, Classical Inference and Relationship* (New York: Oxford University Press).

Sundt, B., and W. S. Jewell, 1981, Further Results on Recursive Evaluation of Compound Distributions, *ASTIN Bulletin*, 11: 27-39.

Smith, D. C., 1999, Finite Sample Properties of Tests of the Epstein-Zin Asset Pricing Model, *Journal of Econometrics*, 93: 113-148.

Willmot, G., and B. Sundt, 1989, On Evaluation of the Delaporte Distribution and Related Distributions, *Scandinavian Actuarial Journal*, 89: 101-113.

Willmot, G., 1987, The Poisson-Inverse Gaussian Distribution as an Alternative to the Negative Binomial, *Scandinavian Actuarial Journal*, 87: 113-127.

Willmot, G., 1988, Sundt and Jewell's Family of Discrete Distributions, *ASTIN Bulletin*, 18: 17-29.