ACTUARIAL RESEARCH CLEARING HOUSE
1993 VOL. 2

A Continuous Estimator of a Mixing Distribution

JACQUES CARRIERE

The University of Manitoba

Abstract

Consider a portfolio of insurance policies where the mean frequency of claims for each policy may vary. This heterogeneity in the portfolio may be modeled as a distribution function $F(\lambda)$ that mixes the mean frequency λ. Using the observed claim frequencies of this portfolio, we present a continuous semiparametric estimator of the mixing distribution $F(\lambda)$ that has some unbiased moments and converges uniformly. The estimator that we investigate is a mixture of gamma distributions whose parameters are calculated by considering the determinants of certain moment matrices.

Keywords

Mixing distribution, moment matrices, semiparametric estimator, uniform consistency.

1. Introduction

Suppose that the number of claims N for a policy can be modeled with the Poisson probability density function (pdf)

$$p(n \mid \lambda) = \frac{e^{-\lambda}\, \lambda^{n}}{n!} \tag{1.1}$$

where $\lambda > 0$ is the mean frequency and n=0,1,2,... In a heterogeneous population the mean frequency is distributed according to some unknown distribution $F(\lambda)$. We will assume throughout the discussion that the mixing distribution $F(\lambda)$ is continuous and that $F(0)=0$. Moreover, we will assume that the mixing distribution is uniquely determined by its moments.

Hossack, Pollard and Zehnwirth (1983) gave an asymptotically consistent estimator of $F(\lambda)$ under the assumption that it belongs to a Gamma class of distributions. Willmot (1987) also gave a consistent estimator when $F(\lambda)$ belongs to an Inverse-Gaussian class of distributions. Obviously, these estimators will be asymptotically biased if the true distribution is not in these parametric classes. Lindsay (1989) constructed a discrete estimator that has some unbiased moments and is consistent when $F(\lambda)$ is uniquely determined by its moments and $F(0)=0$. Expanding on Lindsay's result, we will present a continuous estimator $\hat{F}(\lambda)$ that has some unbiased moments and converges uniformly when $F(\lambda)$ belongs to the class of continuous distributions that are uniquely determined by their moments and $F(0)=0$. In section 2 we will show how to calculate this estimator while in section 3 we will present some of its asymptotic properties.

2. A Semiparametric Estimator

The semiparametric estimator of $F(\lambda)$ that we will investigate will have the following form

$$\hat{F}(\lambda) = \sum_{j=1}^{p} \pi_j \int_0^{\lambda} f_j(y)\, dy \tag{2.1}$$

where

$$f_j(y) = \beta_j^{\alpha}\, y^{\alpha-1} \exp(-\beta_j y) \Big/ \Gamma(\alpha). \tag{2.2}$$

Note that the pdf of the estimator in (2.1) is simply a mixture of Gamma densities but this does not mean that the unknown distribution $F(\lambda)$ has this form. Moment estimates of the parameters β_j and π_j for this mixture are given in Titterington, Smith and Makov (1985), but these estimates do not necessarily satisfy parameter constraints such as $\beta_j > 0$. We now show how to estimate ρ, α, $\beta_1,\ldots,\beta_\rho$ and $\pi_1,\ldots,\pi_\rho$ so that all parameter constraints hold.

Suppose we observe the frequencies N_i for $i=1,\ldots,T$ where T is the number of policies in some insurance portfolio. Also suppose that $N_1,N_2,\ldots$ are independent and identically distributed random variables with a common pdf equal to

$$p(n) = \int_{(0,\infty)} p(n \mid \lambda)\; dF(\lambda) \tag{2.3}$$

where $p(n \mid \lambda)$ is given in (1.1). To calculate the integer $\rho > 0$, we need to estimate the moments $m_k = E(\lambda^k)$ for $k=1,2,\ldots$ A consistent and unbiased estimator of m_k is

$$\hat{m}_k = \sum_{n=0}^{\infty} n(n-1)\cdots(n-k+1)\; \hat{p}(n) \tag{2.4}$$

where

$$\hat{p}(n) = \frac{1}{T} \sum_{i=1}^{T} \mathrm{I}(N_i = n). \tag{2.5}$$

Using the moment sequence $\hat{m}_1,\hat{m}_2,\ldots$ we define $\hat{M}_0=\{1\}$, $\hat{M}_0^*=\{\hat{m}_1\}$ and for $k=1,2,\ldots$ we define the moment matrix

$$\hat{M}_k = \begin{bmatrix} 1 & \hat{m}_1 & \cdots & \hat{m}_k \\ \hat{m}_1 & \hat{m}_2 & \cdots & \hat{m}_{k+1} \\ \vdots & \vdots & & \vdots \\ \hat{m}_k & \hat{m}_{k+1} & \cdots & \hat{m}_{2k} \end{bmatrix} \tag{2.6}$$

and the shifted moment matrix

$$\hat{M}_k^* = \begin{bmatrix} \hat{m}_1 & \hat{m}_2 & \cdots & \hat{m}_{k+1} \\ \hat{m}_2 & \hat{m}_3 & \cdots & \hat{m}_{k+2} \\ \vdots & \vdots & & \vdots \\ \hat{m}_{k+1} & \hat{m}_{k+2} & \cdots & \hat{m}_{2k+1} \end{bmatrix}. \tag{2.7}$$

Using $\tilde{M}_k$ and $\tilde{M}_k^s$ for k=0,1,... we define ρ as follows

$$\rho = 1 + \sup\Big\{k : \det(\tilde{M}_i) > 0 \text{ and } \det(\tilde{M}_i^s) > 0 \ \forall i=0,\ldots,k\Big\}. \tag{2.8}$$

To calculate the rest of the parameters, we define

$$r_k = m_k \Big/ \Big(\alpha(\alpha+1)\cdots(\alpha+k-1)\Big) \tag{2.9}$$

for k=1,2,... Next, we define $R_0=\{1\}$, $R_0^s=\{r_1\}$ and for k=1,2,... we define

$$R_k = \begin{bmatrix} 1 & r_1 & \cdots & r_k \\ r_1 & r_2 & \cdots & r_{k+1} \\ \vdots & \vdots & & \vdots \\ r_k & r_{k+1} & \cdots & r_{2k} \end{bmatrix} \tag{2.10}$$

and

$$R_k^s = \begin{bmatrix} r_1 & r_2 & \cdots & r_{k+1} \\ r_2 & r_3 & \cdots & r_{k+2} \\ \vdots & \vdots & & \vdots \\ r_{k+1} & r_{k+2} & \cdots & r_{2k+1} \end{bmatrix}. \tag{2.11}$$

Using R_k and R_k^s for k=0,1,... we define $\phi(\alpha)$ as follows

$$\phi(\alpha) = 1 + \sup\Big\{k : \det(R_i) > 0 \text{ and } \det(R_i^s) > 0 \ \forall i=0,\ldots,k\Big\}. \tag{2.12}$$

We can now define $\alpha > 0$ as any value that satisfies the inequality

$$\phi(\alpha) \geq \rho\,. \tag{2.13}$$

Let us calculate the parameters $\underset{\sim}{\beta}=(\beta_1,\ldots,\beta_\rho)^T$. Consider the polynomial

$$P(t) = \det \begin{bmatrix} 1 & r_1 & \cdots & r_{\rho-1} & 1 \\ r_1 & r_2 & \cdots & r_\rho & t \\ \vdots & \vdots & & \vdots & \\ r_\rho & r_{\rho+1} & \cdots & r_{2\rho-1} & t^\rho \end{bmatrix}. \tag{2.14}$$

Let $b_j > 0$ for $j=1,\ldots,\rho$ denote the distinct real roots of $P(t)$. We set $\beta_j > 0$ equal to $1/b_j$. Finally, let us calculate the parameters $\underset{\sim}{\pi}=(\pi_1,\ldots,\pi_\rho)^T$. Consider the matrix

$$\Upsilon = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_\rho \\ \vdots & \vdots & & \vdots \\ b_1^{\rho-1} & b_2^{\rho-1} & \cdots & b_\rho^{\rho-1} \end{bmatrix} \tag{2.15}$$

and the vector $\underset{\sim}{r}=(1,r_1,\ldots,r_{\rho-1})^T$. Then $\underset{\sim}{\pi}$ is equal to

$$\underset{\sim}{\pi} = \Upsilon^{-1}\,\underset{\sim}{r}\,. \tag{2.16}$$

Let us apply the result to some motor vehicle data given in Johnson and Hey (1971). In this data we find that T=421,240 and that

$$\begin{aligned} \hat{p}(0) &= .879337 & \hat{p}(1) &= .110495 \\ \hat{p}(2) &= .009341 & \hat{p}(3) &= .000753 \\ \hat{p}(4) &= .000066 & \hat{p}(5) &= .000007. \end{aligned} \tag{2.17}$$

Using (2.4) we find that

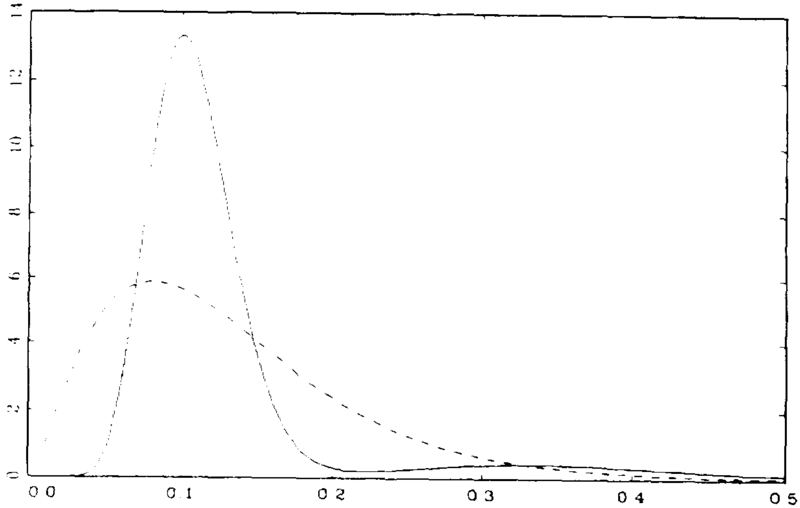
$$\begin{aligned} \hat{m}_1 &= .131735 & \hat{m}_2 &= .024132 \\ \hat{m}_3 &= .006522 & \hat{m}_4 &= .002424 \\ \hat{m}_5 &= .000840 & \hat{m}_6 &= .000000. \end{aligned} \tag{2.18}$$

Using (2.8) we find that ρ=2. Using (2.13) we find that any α greater than 3 is satisfactory. For this example we let α=15. Using the polynomial in (2.14) we found that b_1=.023485, b_2=.007188 and that β_1=42.58, β_2=139.12. Finally, using formula (2.16) we found that π_1=.0978, π_2=.9022. Figure 1 shows a plot of the pdf of the semiparametric estimator in (2.1) and the pdf of the Gamma estimator given in Hossack, Pollard and Zehnwirth (1983). This graph and all the necessary calculations were made with the statistical computing language called GAUSS.

Figure 1

A Plot of the Density of the Semiparametric Estimator

and the Density of the Gamma Estimator



3. Asymptotic Consistency

Let $M_0=\{1\}$ and for $k=1,2,\ldots$ let M_k be equal to the matrix in (2.6) with $\hat{m}_i$ replaced with m_i for $i=1,\ldots,2k$. Also let $M_0^*=\{m_1\}$ and let M_k^* be equal to the matrix in (2.7) with $\hat{m}_i$ replaced with m_i for $i=1,\ldots,2k+1$. Lindsay (1989) called M_k the kth moment matrix of $F(\lambda)$ while M_k^* was called the kth shifted moment matrix of $F(\lambda)$. Using a strong law of large numbers we know that $\hat{m}_k \xrightarrow{a.s.} m_k$ as $T\to\infty$ for $k=1,2,\ldots$ Therefore $\det(\hat{M}_k) \xrightarrow{a.s.} \det(M_k)$ and $\det(\hat{M}_k^*) \xrightarrow{a.s.} \det(M_k^*)$ as $T\to\infty$ because the determinants are continuous functions of the moments. Consulting Shohat and Tamarkin (1943) we find that if $F(0)=0$ and $F(\lambda)$ is continuous then $\det(M_k) > 0$ and $\det(M_k^*) > 0$ for $k=0,1,\ldots$ Using these facts along with the definition of ρ given in (2.8), we get the following result.

Lemma 1. If $T \to \infty$ then $\rho \xrightarrow{a.s.} \infty$. □

Let $\bar{m}_k = \alpha^k r_k$ where r_k is defined in (2.9), then $\bar{m}_k \to \acute{m}_k$ as $\alpha \to \infty$. Let $\bar{M}_k$ be equal to $\acute{M}_k$ with $\acute{m}_k$ replaced by $\bar{m}_k$ and let $\bar{M}_k^*$ be equal to $\acute{M}_k^*$ with $\acute{m}_k$ replaced by $\bar{m}_k$, then $\det(\bar{M}_k) \to \det(\acute{M}_k)$ and $\det(\bar{M}_k^*) \to \det(\acute{M}_k^*)$ as $\alpha \to \infty$. Let $\bar{\rho}$ be equal to (2.8) with $\acute{M}_k$ replaced by $\bar{M}_k$ and $\acute{M}_k^*$ replaced by $\bar{M}_k^*$, then $\exists \alpha_0$ such that $\bar{\rho} \geq \rho$ $\forall \alpha \geq \alpha_0$. According to Lindsay (1989) there exists a discrete distribution with $\bar{\rho}$ distinct atoms of mass τ_j at $a_j > 0$ for $j=1,\ldots,\bar{\rho}$ whose moments are equal to $\bar{m}_k$ for $k=1,\ldots,2\bar{\rho}-1$. Now, consider the discrete distribution with $\bar{\rho}$ atoms of mass τ_j at a_j/α for $j=1,\ldots,\bar{\rho}$. The moments of this discrete distribution are equal to r_k. According to Shohat and Tamarkin (1943), this means that $\det(R_k) > 0$ and $\det(R_k^*) > 0$ $\forall k=0,\ldots,\bar{\rho}-1$ and $\det(R_{\bar{\rho}})=0$ and $\det(R_{\bar{\rho}}^*)=0$ where R_k and R_k^* are defined in (2.10) and (2.11), respectively. Therefore $\phi(\alpha)=\bar{\rho}$ where $\phi(\alpha)$ is defined in (2.12). We summarize the result as follows.

Lemma 2. There exists $\alpha_0 > 0$ such that $\phi(\alpha) \geq \rho$ for all $\alpha \geq \alpha_0$. □

Using our notation we will restate some results given in Lindsay (1989). Note that a version of the first result in the following lemma was used to prove Lemma 2. For the ensuing discussion we will assume that $\alpha \geq \alpha_0$.

Lemma 3. a) If $\det(R_k) > 0$ and $\det(R_k^*) > 0$ $\forall k=0,\ldots,\rho-1$, then there exists a distribution with ρ distinct atoms of mass $\pi_j > 0$ at $b_j > 0$ for $j=1,\ldots,\rho$ whose moments are equal to r_k for $k=1,\ldots,2\rho-1$.

b) Let $P(t)$ be equal to the polynomial given in (2.14), then $P(b_j)=0$ for $j=1,\ldots,\rho$.

c) Let Υ be equal to the matrix given in (2.15) and let $\underline{r}=(1,r_1,\ldots,r_{\rho-1})^T$ and $\underline{\pi}=(\pi_1,\ldots,\pi_\rho)^T$, then $\underline{\pi}=\Upsilon^{-1}\,\underline{r}$. □

We will now give some asymptotic results for the semiparametric estimator $\hat{F}(\lambda)$ given in (2.1). To prove these results we will use certain approximation theorems found in Serfling (1980).

Theorem 4. a) Let $\hat{m}_k=\int_0^{\infty}\lambda^k d\hat{F}(\lambda)$ for $k=1,2,\ldots$, then $\hat{m}_k=\dot{m}_k$ for $k=1,\ldots,2\rho-1$ and for any $k=1,2,\ldots$ we must have $\hat{m}_k \xrightarrow{a.s.} m_k$ as $T\to\infty$.

b) Let $g(\lambda)$ be a bounded and continuous function for all $\lambda > 0$, then $\int_0^{\infty} g(\lambda)\ d\hat{F}(\lambda) \xrightarrow{a.s.} \int_0^{\infty} g(\lambda)\ dF(\lambda)$ as $T\to\infty$.

c) If $T\to\infty$, then $\sup_{\lambda>0} \mid \hat{F}(\lambda)-F(\lambda) \mid \xrightarrow{a.s.} 0$.

d) Let $r > 0$, then $E\mid \hat{F}(\lambda)-F(\lambda)\mid^r\to 0$ and $E(\hat{F}(\lambda))^r\to(F(\lambda))^r$ as $T\to\infty$.

Proof: a) Let $k=1,\ldots,2\rho-1$ and let $b_j=1/\beta_j$. Using Lemma 3 and the definition of $\hat{F}(\lambda)$ in (2.1) and the definition of r_k in (2.9), we find that

$$\hat{m}_k=\sum_{j=1}^{\rho}\pi_j\Big(\alpha(\alpha+1)\cdots(\alpha+k-1)\Big)\Big/\beta_j^k=$$
$$\alpha(\alpha+1)\cdots(\alpha+k-1)\sum_{j=1}^{\rho}\pi_j b_j^k=\alpha(\alpha+1)\cdots(\alpha+k-1)\ r_k=\dot{m}_k.$$

Using Lemma 1 we find that for any $k=1,2,\ldots$ there exists T_k such that for all $T > T_k$ $\hat{m}_k=\dot{m}_k$. Therefore, by a strong law of large numbers $\hat{m}_k \xrightarrow{a.s.} m_k$ as $T\to\infty$.

b) This result follows immediately after applying a theorem by Frechet and Shohat that is given in Serfling (1980), p. 17.

c) This result follows immediately after applying Polya's theorem in Serfling (1980), p. 18.

d) This result follows immediately from standard theorems in Serfling (1980), pp. 11-15. □

References

Hossack, I.B., Pollard, J.H. and Zehnwirth, B. (1983). *Introductory statistics with applications in general insurance*. Cambridge University Press, New York.

Johnson, P.D. and Hey, G.B. (1971). Statistical Studies in Motor Insurance. *Journal of the Institute of Actuaries*. Vol. 97. The Alden Press, Oxford, Great Britain.

Lindsay, B.G. (1989). Moment Matrices: Applications in Mixtures. *The Annals of Statistics*. Vol. 17, No. 2, pp. 722-740.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Shohat, J.A. and Tamarkin, J.D. (1943). *The Problem of Moments. Mathematical Surveys. Number* 1. American Mathematical Society, New York.

Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985). *Statistical Analysis of Finite Mixture Distributions*. Wiley, New York.

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