

Provisions for claims outstanding, incurred but not reported, with generalized linear models: prediction error formulated according to calendar year

Cálculo de provisiones para prestaciones pendientes de declaración con modelos lineales generalizados: formulación del error de predicción por año de calendario

EVA BOJ DEL VAL¹
TERESA COSTA COR¹
Universidad de Barcelona (España)

Recibido el 27 de febrero de 2015, aceptado el 19 de octubre de 2015

Publicado online el 25 de enero de 2017

Nº de clasificación JEL: C13, C15

DOI: 10.5295/cdg.150526eb

Abstract:

In the current context of Solvency II, insurance companies are required to implement demanding business risk management systems. An important aspect of this risk management is the problem of technical provisions in non-life insurance and, as such, it is in the interest of insurers to calculate the prediction error that has occurred when using methodology to estimate a company's future payments. Furthermore, the predictive distribution of the fitted values, which is descriptive of the risk, allows us to estimate, for example, its Value at Risk at a given confidence level. In this paper we focus on the application of generalized linear models to the amounts of claim losses of a run-off triangle. In order to achieve error distribution, a parameter dependent parametric family is assumed, along with the logarithmic link function. The parametric family has as particular cases the Poisson, the Gamma and the Inverse Gaussian distributions. The particular case which assumes an (over-dispersed) Poisson distribution with the logarithmic link is widely known because it offers the same provision estimation as the deterministic Chain-Ladder method. In this study we develop formulas of the prediction error of future payments by calendar years for the general parametric family. This allows us to perform calculations that consider a financial environment, whether employing analytical formulation or bootstrap estimation. In practice, the presented formulations allow a determination to be made of the present value of the incurred but not reported claim of future payments including a risk margin with statistical significance.

Keywords:

Technical provisions, generalized linear model, calendar year, Solvency II.

¹ Departamento de Matemática Económica, Financiera y Actuarial, Facultad de Economía y Empresa, Avenida Diagonal 690, 08034 Barcelona (España). evaboj@ub.edu; costa@ub.edu

Resumen:

El actual contexto de Solvencia II requiere una exigente gestión empresarial del riesgo de las Entidades Aseguradoras. En el problema de cálculo de provisiones técnicas en seguros de no-vida es de interés calcular el error de predicción cometido con la metodología utilizada para la estimación de los pagos futuros de la Entidad. Además, la distribución predictiva de las estimaciones, que es descriptiva respecto del riesgo, permite obtener, por ejemplo, su valor en riesgo a un nivel de confianza fijado. En este trabajo nos centramos en la aplicación de los modelos lineales generalizados a las cuantías de siniestros de un triángulo de desarrollo. Asumimos para la distribución del error una familia paramétrica dependiente de un parámetro, junto con la función de enlace logarítmica. La familia paramétrica tiene como casos particulares las distribuciones de Poisson, Gamma e Inversa Gaussiana. Es conocido el caso particular en que se asume una distribución de Poisson (sobredispersa) junto con el link logarítmico, que ofrece la misma estimación de provisiones que el método determinista Chain-Ladder. En este estudio desarrollamos las fórmulas del error de predicción de los pagos futuros por años de calendario para la familia paramétrica general, que nos permiten realizar cálculos teniendo en cuenta un ambiente financiero, tanto para el caso de utilizar formulación analítica como para el caso de realizar estimación bootstrap. En la práctica, las formulaciones presentadas nos ponen en disposición de poder calcular el valor actual de los pagos futuros para siniestros pendientes incluyendo márgenes de riesgo con significado estadístico.

Palabras clave:

Provisiones técnicas, modelo lineal generalizado; año de calendario, Solvencia II.

1. INTRODUCTION

This paper takes a similar starting point to that of Boj et al. (2014) in that it considers a portfolio of risks and assumes that each claim is settled either in the accident year or in the following k development years. It then considers a family of random variables $\{c_{ij}\}_{i,j \in \{0,1,\dots,k\}}$, where c_{ij} is the amount of claim losses of accident year i which is paid with a delay of j years and hence in development year j and in calendar year $i+j$. c_{ij} refers to the incremental loss of accident year i and development year j and it is assumed that these incremental losses c_{ij} are observable for calendar years $i+j \leq k$ and that they are collected in a run-off triangle as in Figure 1.

Figure 1

Run-off triangle with incremental losses

Accident year	Development year								
	0	1	...	j	...	$k-i$...	$k-1$	k
0	$c_{0,0}$	$c_{0,1}$...	$c_{0,j}$...	$c_{0,k-i}$...	$c_{0,k-1}$	$c_{0,k}$
1	$c_{1,0}$	$c_{1,1}$...	$c_{1,j}$...	$c_{1,k-i}$...	$c_{1,k-1}$	
...		
i	$c_{i,0}$	$c_{i,1}$...	$c_{i,j}$...	$c_{i,k-i}$			
...				
$k-j$	$c_{k-j,0}$	$c_{k-j,1}$...	$c_{k-j,j}$					
...						
$k-1$	$c_{k-1,0}$	$c_{k-1,1}$							
k	$c_{k,0}$								

Source: Own elaboration.

In the run-off triangle, the numbers for year of origin i are grouped into rows while those for development year j are grouped into columns. The numbers on the antidiagonals with $i+j = t$ denote the payments that were made in the calendar year t . The incremental losses are unobservable for calendar years $i+j \geq k + 1$ and are thus very difficult to predict.

There are various kinds of incurred but not reported (IBNR) claim provisions which are of interest. The provisions for the different accident years $i = 1, \dots, k$ are obtained by adding the future incremental losses predicted in the corresponding row of the square:

$$P_i = \sum_{j=k-i+1}^k \hat{c}_{ij} \tag{1}$$

The total provision is calculated by adding all the future incremental losses predicted in the bottom-right part of the run-off triangle:

$$P = \sum_{t=1}^k \sum_{j=k-t+1}^k \hat{c}_{ij}. \quad (2)$$

And the future payments for the different calendar years $t = k + 1, \dots, 2k$ are obtained by adding the incremental losses that were made in the future calendar years, i.e., the values of the same against-diagonal t :

$$FP_t = \sum_{j=t-k}^k \hat{c}_{t-j,j}. \quad (3)$$

The principal focus of this paper is to calculate the present value of the future payments for the different calendar years, taking into account the Solvency II Directive (see *Parlamento Europeo y Consejo de la Unión Europea* 2009, Albarrán and Alonso 2010, Moreno 2013). In Solvency II, it is clearly stated that the *best estimate* of the technical provisions must be calculated by including the time value of money, which is to say, that they must include the expected present value of future payments. Calculating payments by calendar year allow us to work in a financial environment because each amount is situated in the corresponding future calendar year. In order to consider a conservative scenario Solvency II recommends adding a *risk margin* to the *best estimate* to obtain a correct *Solvency Capital Requirement* (SCR). If possible, the SCR should also derive directly from the probability distribution forecast generated by the statistic applied model, using the Value at Risk (VaR) measure.

Section 3 describes two means of determining an adequate Solvency Capital Requirement. While the first of these consists of adding a percentage of the prediction error to the future payments by calendar years, the second consists of directly using the estimated VaR of the predictive distribution of the future payments by calendar years for a given significance level.

A generalized linear model (GLM) is assumed (see, e.g., McCullagh and Nelder 1989; Boj et al. 2004; Boj and Costa 2014) to model the incremental losses of the run-off triangle. The parametric family of error distributions for the variance function is assumed

$$V(\mu_{ij}) = \mu_{ij}^\theta, \quad (4)$$

named “power variance function”, which depends on parameter θ . This parametric family has as particular cases: the Poisson distribution when $\theta = 1$; the Gamma distribution when $\theta = 2$; and the Inverse Gaussian distribution when $\theta = 3$. Assuming (4) the mean and the variance of the GLM are

$$\mu_{ij} = E[c_{ij}] \quad \text{and} \quad Var[c_{ij}] = (\phi/w_{ij})V(\mu_{ij}) = (\phi/w_{ij})\mu_{ij}^\theta, \quad (5)$$

where ϕ is the dispersion parameter and w_{ij} are *a priori* weights of the data, assumed equal to one, $w_{ij} = 1$, for the incremental claim losses of a run-off triangle. The estimation

process of this kind of GLM, which does not have a distribution in the exponential family, uses extended quasi-likelihood equations. A detailed description of the fitting algorithm may be found in McCullagh and Nelder (1989).

Several actuarial methods which are frequently used to complete a run-off triangle can be described by a GLM. Among these are the Chain-Ladder method, the arithmetic and geometric separation methods, and de Vylder's least squares method. In particular, the classical Chain-Ladder deterministic method can be derived from a GLM assuming (over-dispersed) Poisson errors and logarithmic link function (see, e.g., Haberman and Renshaw 1996; England and Verrall 1999, 2002 and 2006; England 2002; Kaas et al. 2008; Boj et al. 2014).

Assume for the GLM the logarithmic link function

$$\log \mu_{ij} = \eta_{ij}. \tag{6}$$

Next, we can define $\log(\mu_{ij}) = c_0 + \alpha_i + \beta_j$, a GLM in which the responses c_{ij} are modelled as random variables with variance function (4), with a logarithmic link function (6) and with linear predictor

$$\eta_{ij} = c_0 + \alpha_i + \beta_j, \tag{7}$$

where α_i is the factor corresponding to the accident year $i = 1, \dots, k$ and β_j is the factor corresponding to the development year $j = 1, \dots, k$. The c_0 value is the term corresponding to the accident year 0 and development year 0. In the Poisson case, where $\emptyset = 1$ is assumed, over-dispersion is taken into account by estimating the unknown scale parameter \emptyset as a part of the fitting procedure.

The predicted values \hat{c}_{ij} of the IBNR provisions (1), (2) and future payments (3) are estimated from

$$\hat{c}_{ij} = \exp(\hat{c}_0 + \hat{\alpha}_i + \hat{\beta}_j). \tag{8}$$

The prediction error of (8) assuming (4) is studied for the accident year provisions (1) and the total provision (2) in, e.g., England and Verrall (1999, 2002, 2006) and England (2002). This study has been extended to the IBNR future payments by calendar years (3) in Boj et al. (2014) and Espejo et al. (2014) but only for the particular case in which the (over-dispersed) Poisson distribution is assumed (a case that gives the same estimation of provisions as the deterministic Chain-Ladder method). The novelty of the current study is that the formulas of the prediction error for the future payments by calendar years (10) and (14) of Boj et al. (2014) are extended to the general parametric family (4).

The paper is organized as follows. Section 2 develops formulas for the prediction error of a GLM for the future payments by calendar years (3) assuming the parametric family of distributions (4). Section 3 applies these formulas to real data and a practical demonstration of Solvency Directive II is given by working with the future payments by calendar year studied in Section 2 (using an interest rate and including a risk margin to the provisions). In Section 4 the main conclusions of the paper are presented while Sections 5 and 6 contain acknowledgments and references.

2. PREDICTION ERROR FOR FUTURE PAYMENTS BY CALENDAR YEARS

Consider a random variable c_{ij} and a predicted value \hat{c}_{ij} . The mean squared error (MSE) of prediction is:

$$E \left[(c_{ij} - \hat{c}_{ij})^2 \right] = E \left[\left((c_{ij} - E[c_{ij}]) - (\hat{c}_{ij} - E[\hat{c}_{ij}]) \right)^2 \right]. \quad (9)$$

Plugging in \hat{c}_{ij} instead of c_{ij} in the final expectation and expanding gives:

$$E \left[(c_{ij} - \hat{c}_{ij})^2 \right] \approx E \left[(c_{ij} - E[c_{ij}])^2 \right] - 2E \left[(c_{ij} - E[c_{ij}]) (\hat{c}_{ij} - E[\hat{c}_{ij}]) \right] + E \left[(\hat{c}_{ij} - E[\hat{c}_{ij}])^2 \right]. \quad (10)$$

The assumption that future observations are independent of past observations gives:

$$E \left[(c_{ij} - \hat{c}_{ij})^2 \right] \approx E \left[(c_{ij} - E[c_{ij}])^2 \right] + E \left[(\hat{c}_{ij} - E[\hat{c}_{ij}])^2 \right]. \quad (11)$$

Thus, the prediction variance is the sum of process variance and $E \left[(c_{ij} - E[c_{ij}])^2 \right] = \text{Var}[c_{ij}]$ the estimation variance $E \left[(\hat{c}_{ij} - E[\hat{c}_{ij}])^2 \right] = \text{Var}[\hat{c}_{ij}]$. The MSE of predictions is given by:

$$\text{MSE}(c_{ij}) = E \left[(c_{ij} - \hat{c}_{ij})^2 \right] \approx \text{Var}[c_{ij}] + \text{Var}[\hat{c}_{ij}]. \quad (12)$$

The following can be derived by using the delta method:

$$\text{Var}[\hat{c}_{ij}] \cong \left| \frac{\partial \mu_{ij}}{\partial \eta_{ij}} \right|^2 \text{Var}[\eta_{ij}]. \quad (13)$$

In the case where the logarithmic link function $\frac{\partial \mu}{\partial \eta} = \mu$, and then, the MSE of predictions could be approximated as:

$$\text{MSE}(c_{ij}) \cong \phi \mu_{ij}^\theta + \mu_{ij}^2 \text{Var}[\eta_{ij}]. \quad (14)$$

The prediction error (PE) for each prediction with $i, j = 1, \dots, k$ could thus be calculated as the root of the MSE (14). Next, the formulas of the MSE are given for the accident year provisions, for the total provision and for the future payments by calendar years. The PE of those amounts could be calculated as the corresponding root.

The PE obtained when assuming the parameter dependent family of distributions (4) for the accident year provisions and the total provision, has been studied in e.g. England and Verrall (1999). The PE for the future payments by calendar years in the (over-dispersed) Poisson case has been studied in Boj et al. (2014) and Espejo et al. (2014) where it is programmed for practitioners with RExcel. Next, the parametric family of distributions (4) is assumed, thereby extending the formulations for the future payments by calendar year (10) and (14) of Boj et al. (2014).

The squared PE for the accident year provisions are:

$$\begin{aligned}
 MSE(P_i) &= E\left[\left(P_i - \hat{P}_i\right)^2\right] \approx \sum_{\substack{j=1,\dots,k \\ i+j>k}} \phi \mu_{ij}^\theta + \mu_i^T Var[\eta_i] \mu_i = \\
 &= \sum_{\substack{j=1,\dots,k \\ i+j>k}} \phi \mu_{ij}^\theta + \sum_{\substack{j=1,\dots,k \\ i+j>k}} \mu_{ij}^2 Var[\eta_{ij}] + 2 \sum_{\substack{j_1, j_2=1,\dots,k \\ i+j_1>k, i+j_2>k}} \mu_{j_1} \mu_{j_2} Cov[\eta_{j_1}, \eta_{j_2}]. \\
 & \quad i = 1, \dots, k
 \end{aligned}
 \tag{15}$$

The squared PE for the total provision is:

$$\begin{aligned}
 MSE(P) &= E\left[\left(P - \hat{P}\right)^2\right] \approx \sum_{\substack{i,j=1,\dots,k \\ i+j>k}} \phi \mu_{ij}^\theta + \mu^T Var[\eta] \mu = \\
 &= \sum_{\substack{i,j=1,\dots,k \\ i+j>k}} \phi \mu_{ij}^\theta + \sum_{\substack{i,j=1,\dots,k \\ i+j>k}} \mu_{ij}^2 Var[\eta_{ij}] + 2 \sum_{\substack{i_1, j_1, i_2, j_2=1,\dots,k \\ i_1+j_1>k, i_2+j_2>k \\ i_1 j_1 \neq i_2 j_2}} \mu_{i_1 j_1} \mu_{i_2 j_2} Cov[\eta_{i_1 j_1}, \eta_{i_2 j_2}].
 \end{aligned}
 \tag{16}$$

And the squared PE for the future payments by calendar years are:

$$\begin{aligned}
 MSE(FP_t) &= E\left[\left(FP_t - \widehat{FP}_t\right)^2\right] \approx \sum_{\substack{i,j=1,\dots,k \\ i+j=t}} \phi \mu_{ij}^\theta + \mu_t^T Var[\eta_t] \mu_t = \\
 &= \sum_{\substack{i,j=1,\dots,k \\ i+j=t}} \phi \mu_{ij}^\theta + \sum_{\substack{i,j=1,\dots,k \\ i+j=t}} \mu_{ij}^2 Var[\eta_{ij}] + 2 \sum_{\substack{i_1, j_1, i_2, j_2=1,\dots,k \\ i_1+j_1=t, i_2+j_2=t \\ i_1 j_1 \neq i_2 j_2}} \mu_{i_1 j_1} \mu_{i_2 j_2} Cov[\eta_{i_1 j_1}, \eta_{i_2 j_2}]. \\
 & \quad t = k + 1, \dots, 2k
 \end{aligned}
 \tag{17}$$

Next, formulas are given for the case in which the $Var[\hat{c}_{ij}]$ is estimated by means of standard error (SE) of the predictive distribution of c_{ij} . The predictive distribution of c_{ij} is estimated by bootstrap methodology (see Efron and Tibshirani 1998). The treatment of the GLM and associated claims in this paper frequently involve the application of bootstrapping residuals based on Pearson residuals (see, e.g., Boj and Costa 2014). When (4) is assumed and $w_{ij} = 1$, the Pearson residuals take the expression:

$$r_{ij}^P = \frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{c}_{ij}^\theta}}.
 \tag{18}$$

The scale parameter $\hat{\theta}^\rho$ could be estimated by:

$$\hat{\theta}^\rho = \frac{1}{n - 2k - 1} \sum_{\substack{i,j=1,\dots,k \\ i+j \leq k}} \left(r_{ij}^P\right)^2 = \frac{1}{n - 2k - 1} \sum_{\substack{i,j=1,\dots,k \\ i+j \leq k}} \frac{\left(c_{ij} - \hat{c}_{ij}\right)^2}{\hat{c}_{ij}^\theta}.
 \tag{19}$$

In this paper, the residuals in the bootstrap process are adjusted for the degrees of freedom as follows:

$$r_{ij}^{P^*} = \sqrt{\frac{n}{n - 2k - 1}} r_{ij}^P,
 \tag{20}$$

as similarly done, for the sake of compatibility, for the scale parameter $\hat{\theta}^\rho$.

In the bootstrap process estimation we have B resamples of the residuals, $r_{ij}^{P^*}$. From (18) and (20) we isolate $c_{ij}^* = r_{ij}^{P^*} \sqrt{\hat{c}_{ij}^\theta} + \hat{c}_{ij}$ for each sample, and then we estimate the

GLM and the corresponding amounts. The B values of the estimations give the predictive distribution of a value \hat{c}_{ij}^{boot} (for each value), of the accident year provisions \hat{p}_i^{boot} (for each accident year), of the total provision \hat{P}^{boot} , and of the future payments by calendar years \widehat{FP}_t^{boot} (for each calendar year).

From the B values it is possible to estimate the variance of the distribution as the square of the standard error (SE) of the distribution. The SE estimated in this way replaces the part of the estimation variance, $Var[\hat{c}_{ij}]$, in formulas (14), (15), (16) and (17).

Next, a description is provided of the formulations of the estimations of the PE, i.e., the root of the MSE is given, for a value, for the accident year provisions, for the total provision and for the future payments by calendar years in the case of bootstrap estimation of $Var[\hat{c}_{ij}]$. We denote this by PE^{boot} .

The bootstrap estimations of the PE for each value are:

$$PE^{boot}(c_{ij}) \approx \sqrt{\hat{\phi}^P \hat{c}_{ij}^\theta + SE(\hat{c}_{ij}^{boot})^2}, \quad i, j = 1, \dots, k. \quad (21)$$

The bootstrap estimations of the PE for the accident year provisions are:

$$PE^{boot}(P_i) \approx \sqrt{\sum_{\substack{j=1, \dots, k \\ i+j > k}} \hat{\phi}^P \hat{c}_{ij}^\theta + SE(\hat{P}_i^{boot})^2}, \quad i = 1, \dots, k. \quad (22)$$

The bootstrap estimation of the PE for the total provision is:

$$PE^{boot}(P) \approx \sqrt{\sum_{\substack{i, j=1, \dots, k \\ i+j > k}} \hat{\phi}^P \hat{c}_{ij}^\theta + SE(\hat{P}^{boot})^2}. \quad (23)$$

And the bootstrap estimations of the PE for the future payments by calendar years are:

$$PE^{boot}(FP_t) \approx \sqrt{\sum_{\substack{i, j=1, \dots, k \\ i+j=t}} \hat{\phi}^P \hat{c}_{ij}^\theta + SE(\widehat{FP}_t^{boot})^2}, \quad t = k+1, \dots, 2k. \quad (24)$$

Note that, unlike the formulas (12), (13), (14) and (15) of Boj et al. (2014), here SE of the bootstrap distributions is not adjusted for degrees of freedom, because it is assumed that work in the bootstrap process is completed with (20), the adjusted Pearson residuals. If direct use were made of the Pearson residuals (18), as in Boj et al. (2014), it would be necessary to correct the SE multiplying it by $n/n(n-2k-1)$, where $2k+1$ is the number of parameters of model (7).

3. APPLICATION

The proposed methodology is illustrated using the triangle of Taylor and Ashe (1983) of Figure 2 with incremental losses. This dataset is used in many texts that deal with IBNR problems, such as those of Renshaw (1989, 1994), England and Verrall (1999) and England

(2002). In particular, England and Verrall (1999), demonstrate (in Tables 1 and 2) the estimation of provisions and the PE as a percentage of the provision estimate (i.e., the called “coefficient of variation” in many text books and computer programs) for the accident year and total provisions, (1) and (2). This is calculated, in cases when the over-dispersed Poisson and the Gamma distributions are assumed using analytic formulas to calculate PE and when the over-dispersed Poisson distribution is assumed using bootstrap methodology for the estimation of PE.

Figure 2

Run-off triangle with 55 incremental losses

Accident year	Development year									
	0	1	2	3	4	5	6	7	8	9
0	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
1	352118	884021	933894	1183289	445745	320996	527804	266172	425046	
2	290507	1001799	926219	1016654	750816	146923	495992	280405		
3	310608	1108250	776189	1562400	272482	352053	206286			
4	443160	693190	991983	769488	504851	470639				
5	396132	937085	847498	805037	705960					
6	440832	847361	1131398	1063269						
7	359480	1061648	1443370							
8	376686	986608								
9	344014									

Source: Taylor and Ashe (1983)

This Section completes the analysis conducted by England and Verrall (1999) by including the study for the future payments by calendar years (3). Poisson (over-dispersed) and Gamma distributions are assumed and an estimation is made of the PE with analytic formula and with bootstrap (using size $B = 1000$ resamples). Both distributions are particular cases of the parametric family (4) described in this paper: when $\theta = 1$ we have the Poisson distribution and when $\theta = 2$ we have the Gamma distribution. In addition, computations are made with the R software (R Development Core Team 2015) and models are fitted with the function *glm* of the *stats* package for R. Results are given in Tables 1 to 8.

The future payments by calendar years enable to work in a financial environment and to calculate the best estimate of provisions. The present value of the future payments by calendar years can be determined by taking into account the time value of money (and thus, following the Solvency II Directive). The present value constitutes the “today real provision” of an insurance company. This Section provides an example which assumes a 1.5% fixed annual interest rate for the future nine years. Additionally, a risk margin can be added to calculate the SCR.

Three possibilities are presented to show how to calculate the best estimate in the context of Solvency II:

- 1) In the first of these, the present value of the future payments by calendar years is computed without any risk margin:

$$IBNR_{actual} = \sum_{t=k+1}^{2k} FP_t (1 + I_1^t)^{-(t-k)}, \quad (25)$$

where I_1^t for $t = k + 1, \dots, 2k$ are the annual interest rates for each of the next calendar years taken into account in the run-off triangle.

2) In the second, the present value of the future payments by calendar years is computed plus a fixed percentage, δ , of the PE. In this way a risk margin is included for each calendar year equal to the fixed percentage of the corresponding PE. In the application of this Section 25% of the PE is added, i.e., $\delta = 0.25$. Use of the analytic formula (17) to estimate the $MSE(FT_p)$ derives:

$$IBNR_{actual} = \sum_{t=k+1}^{2k} \left(FP_t + \delta \sqrt{MSE(FP_t)} \right) (1 + I_1^t)^{-(t-k)}. \quad (26)$$

And use of bootstrap estimation of the PE with formula (24) to estimate the $PE^{boot}(FT_p)$, derives:

$$IBNR_{actual} = \sum_{t=k+1}^{2k} \left(FP_t + \delta PE^{boot}(FP_t) \right) (1 + I_1^t)^{-(t-k)}. \quad (27)$$

3) And in the third possibility, the present value of the values at risk, VaR, is computed at a fixed confidence level α for each calendar year predictive future payments distribution PE^{boot} :

$$IBNR_{actual} = \sum_{t=k+1}^{2k} VaR_{\alpha} \left(\widehat{FP}_t^{boot} \right) (1 + I_1^t)^{-(t-k)}. \quad (28)$$

A risk margin is thus added because the expected mean of the predictive distributions (which must be similar to the future payments by calendar years) is replaced by the quantile α of its predictive distribution. The example given in this Section uses the 99.5% quantile of predictive distribution of the future payments by calendar years. Note that this way of adding a risk margin is only available when estimates are made by bootstrapping the predictive distribution of future payments.

Table 1 calculates the future payments by calendar years, the prediction errors and the coefficients of variation (i.e., the proportion of the PE over the amount estimate, in percentage) for the over-dispersed Poisson distribution using analytic formula to compute the PE as the root of the MSE (17). In Table 2 we have the same computations for the Gamma distribution.

Note that in Table 1, in the second column, the estimations obtained of the future payments by calendar years are the same as those obtained with the Chain-Ladder deterministic method. The reason for this, as explained in the paper, is because when an over-dispersed Poisson distribution is assumed with the logarithmic link function, the estimation with the GLM coincides with that of the Chain-Ladder method.

In Table 3, the means and the standard deviations of the predictive distributions are calculated for the future payments by calendar years, the prediction errors (24), and the

coefficients of variation for the over-dispersed Poisson distribution using bootstrap methodology with $B = 1000$ resamples. Table 4 shows the same computations for the Gamma distribution.

In Table 5, calculations are shown for the present value of the future payments by calendar years as in (25), and the present value of the future payments by calendar years plus 25% of the prediction error using analytic formula as in (26), assuming in both cases the over-dispersed Poisson distribution. Table 6 provides the same computations for the Gamma distribution. In all cases it is assumed that there is a 1.5% fixed annual interest rate for the future nine years.

In Table 7, calculations are shown for the present value of the future payments by calendar years as in (25); the present value of the future payments by calendar years using bootstrap methodology (with 1000 resamples) plus 25% of the prediction error as in (27); and the present value of the Value at Risk as in (28) with a confidence level of 99.5%, assuming the over-dispersed Poisson distribution. In Table 8, the same computations are shown for the Gamma distribution. In both cases a 1.5% fixed annual interest rate is assumed.

Table 1

Prediction errors for over-dispersed Poisson using analytic formula²

Calendar year	Payment	Prediction error	Coefficient of variation
10	5226535.8	747369.6	14.30 %
11	4179394.4	710144.6	16.99 %
12	3131667.5	644139.5	20.57 %
13	2127271.9	479125.6	22.52 %
14	1561878.9	404967.7	25.93 %
15	1177743.7	364294.9	30.93 %
16	744287.4	294424.6	39.56 %
17	445521.3	250986.8	56.34 %
18	86554.6	108268.8	125.09 %

Source: Own elaboration.

Table 2

Prediction errors for Gamma using analytic formula³

Calendar year	Payment	Prediction error	Coefficient of variation
10	5096855.3	847281.6	16.62 %
11	4050001.5	749549.8	18.51 %
12	3064407.7	628141.0	20.50 %
13	2078010.5	431885.8	20.78 %

² Future payments by calendar years, prediction errors and coefficients of variation for the over-dispersed Poisson distribution using analytic formula.

³ Future payments by calendar years, prediction errors and coefficients of variation for the Gamma distribution using analytic formula.

14	1510392.7	345880.7	22.90 %
15	1095402.7	292255.7	26.68 %
16	692118.4	220057.8	31.79 %
17	416539.9	181226.5	43.51 %
18	82075.9	47918.1	58.38 %

Source: Own elaboration.

Table 3

Prediction errors for over-dispersed Poisson using bootstrap⁴

Calendar year	Mean payment	Standard deviation	Prediction error	Coefficient of variation
10	5262187.6	748660.0	756563.2	14.48 %
11	4206004.9	718618.2	721067.3	17.25 %
12	3153556.8	653134.1	649753.2	20.75 %
13	2139244.8	504395.7	487995.9	22.94 %
14	1562523.4	408073.2	411005.7	26.31 %
15	1178586.1	364102.8	365547.7	31.04 %
16	771451.6	302919.9	292974.4	39.36 %
17	455633.0	250906.3	254458.2	57.11 %
18	91579.0	104329.2	107988.8	124.76 %

Source: Own elaboration.

Table 4

Prediction errors for Gamma using bootstrap⁵

Calendar year	Mean payment	Standard deviation	Prediction error	Coefficient of variation
10	5096897.2	1017.5	652964.8	12.81 %
11	4050047.9	1014.2	545647.9	13.47 %
12	3064465.1	920.9	434825.7	14.19 %
13	2078021.4	694.1	297581.3	14.32 %
14	1510393.3	580.7	233914.7	15.49 %
15	1095419.1	493.1	194252.2	17.73 %
16	692130.4	395.9	142601.5	20.60 %
17	416548.9	342.7	109441.6	26.27 %
18	82080.9	152.7	26649.2	32.47 %

Source: Own elaboration.

⁴ Future payments by calendar years, standard deviations, prediction errors and coefficients of variation for the over-dispersed Poisson distribution using bootstrap methodology with 1000 resamples.

⁵ Future payments by calendar years, standard deviations, prediction errors and coefficients of variation for the Gamma distribution using bootstrap methodology with 1000 resamples.

Table 5

Present values of the future payments for over-dispersed Poisson using analytic formula⁶

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction error
10	1	5226535.8	5413378.2
11	2	4179394.4	4356930.6
12	3	3131667.5	3292702.4
13	4	2127271.9	2247053.3
14	5	1561878.9	1663120.8
15	6	1177743.7	1268817.4
16	7	744287.4	817893.5
17	8	445521.3	508268.0
18	9	86554.6	113621.8
Present value		17873967	18820197

Source: Own elaboration.

Table 6

Present values of the future payments for Gamma using analytic formula⁷

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction error
10	1	5096855.3	5308675.7
11	2	4050001.5	4237389.0
12	3	3064407.7	3221442.9
13	4	2078010.5	2185982.0
14	5	1510392.7	1596862.9
15	6	1095402.7	1168466.7
16	7	692118.4	747132.9
17	8	416539.9	461846.5
18	9	82075.9	94055.4
Present value		17310125	18199962

Source: Own elaboration.

⁶ Present values of the future payments by calendar years and of the future payments by calendar years plus 25% of the prediction error for the over-dispersed Poisson distribution using analytic formula and assuming a 1.5% fixed annual interest rate.

⁷ Present values of the future payments by calendar years and of the future payments by calendar years plus 25% of the prediction error for the Gamma distribution using analytic formula and assuming a 1.5% fixed annual interest rate.

Table 7

Present values of the future payments and of the Value at Risk for over-dispersed Poisson using bootstrap⁸

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction Error	VaR _{99.5}
10	1	5226535.8	5415676.6	7417055.0
11	2	4179394.4	4359661.3	6364764.7
12	3	3131667.5	3294105.8	5207534.8
13	4	2127271.9	2249270.9	3682095.3
14	5	1561878.9	1664630.3	2735533.8
15	6	1177743.7	1269130.6	2209257.2
16	7	744287.4	817531.0	1841310.7
17	8	445521.3	509135.8	1262432.7
18	9	86554.6	113551.8	473412.3
Present value		17873967	18830614	29688278

Source: Own elaboration.

Table 8

Present values of the future payments and of the Value at Risk for Gamma using bootstrap⁹

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction Error	VaR _{99.5}
10	1	5096855.3	5260096.5	5099652.6
11	2	4050001.5	4186413.5	4052656.1
12	3	3064407.7	3173114.1	3067038.9
13	4	2078010.5	2152405.9	2079754.3
14	5	1510392.7	1568871.3	1511867.6
15	6	1095402.7	1143965.8	1096662.3
16	7	692118.4	727768.8	693145.23
17	8	416539.9	443900.3	417478.6
18	9	82075.9	88738.2	82457.2
Present Value		17310125	17938348	17324230

Source: Own elaboration.

It may be observed in Tables 1 to 4 that when an analysis is made of the future payments by calendar years, in general and for this dataset, lower coefficients of variation are

⁸ Present values of the future payments by calendar years, of the future payments by calendar years plus 25% of the prediction error and of the Value at Risk with a confidence level of the 99.5% for the over-dispersed Poisson distribution using bootstrap methodology with 1000 resamples and assuming a 1.5% fixed annual interest rate.

⁹ Present values of the future payments by calendar years, of the future payments by calendar years plus 25% of the prediction error and of the Value at Risk with a confidence level of the 99.5% for the Gamma distribution using bootstrap methodology with 1000 resamples and assuming a 1.5% fixed annual interest rate.

obtained for $\Theta = 2$, the case of the Gamma distribution, if the results are compared with those of the over-dispersed Poisson where $\Theta = 1$. And, specifically, it becomes more evident when we compare Tables 3 and 4, the case in which an estimation is made of the real distribution by bootstrap. The same pattern may be observed in the results for accident year provisions and total provisions in Tables 1 and 2 of England and Verrall (1999).

As may be seen, it is preferable to use the bootstrap estimations (21) to (24) of the PE instead of the analytic estimation formulas (14) to (17), because in the run-off triangle there is usually a small dataset and the data do not always follow the hypothesis assumed in the GLM. It is thus preferable to simulate the real distribution of our portfolio than apply the theoretical formulations of the distribution. This will help to obtain more accurate information on future losses incurred by insurance companies.

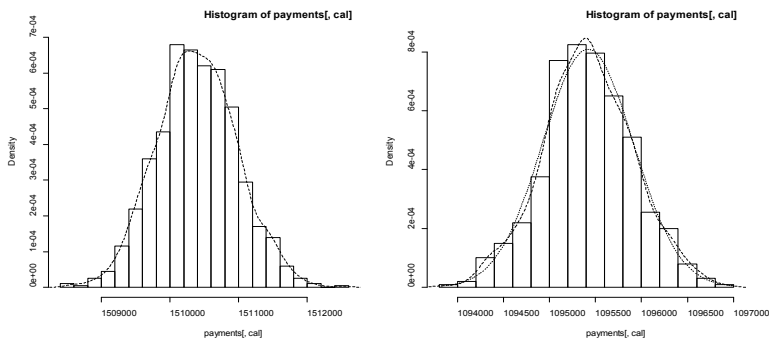
It may be observed in Tables 5 to 8 that the amount of the “today real provision”, i.e. the present value, depends on the added risk margin. While the lower amount always represents the simple present value the simple present value (25), the amount will vary depending on the added percentage δ for formulas (26) and (27), and depending on the confidence level α for (28). As might be expected, higher δ (or α) will obtain higher present values, as will be seen below.

Some additional present values are calculated with formulas (27) and (28) for the Gamma model. First, some percentages δ are used to complete the results of Table 8, with $\delta = \{0.1, 0.4, 0.5, 0.95\}$, and the corresponding present values of 17561414, 18315283, 18566572 and 19697375, respectively, in monetary units. Second, the present values are calculated with formula (28) and taking into account the confidence levels for the VaR: $\alpha = \{75\%, 85\%, 90\%, 95\%\}$. The corresponding present values are 17313858, 17315730, 17317046 and 17319257, respectively, in monetary units.

Finally, Figure 3 shows the predictive distributions of the future payments by calendar years for the calendar years 13 and 14 to illustrate graphically the empirical results of the bootstrap process. Calculations of the percentiles for the VaR or the SE of formula (24) are drawn from these distributions.

Figure 3

Histograms of the predictive distribution of the future payments with Gamma¹⁰



Source: Own elaboration.

¹⁰ Histograms of the predictive distribution of the future payments for the calendar years 14 and 15 for the data of Taylor and Ashe (1983) assuming a GLM with Gamma distribution.

4. CONCLUDING REMARKS

This paper has deduced the formulas (17) and (24) related to the PE for the future payments by calendar years (3) for the GLM in the general case of the parametric distribution family (4) assuming the logarithmic link. Studying the use of calendar years for the IBNR provisions problem has provided results that enable the actuary to take decisions regarding the best estimate of the technical provisions and the risk margin, and, therefore, regarding the financial inversions of the SCR of the insurance companies in the current context of Solvency II. As regards the best estimate, this paper proposes that the present value (25) be calculated directly by taking account of the time value of money. Additionally, it is proposed that a risk margin be included by adding a percentage of the prediction error of the future payments by calendar years, as seen in formulas (26) and (27) or by calculating the VaR of the predictive distribution of the future payments by calendar years, as seen formula (28).

While the analysis in the paper is illustrated with the triangle of Taylor and Ashe (1983), the theoretical and practical results complement the study of the use of GLM in the problem of provisions in other actuarial studies where the analysis is done only with the accident year provisions and total provisions (1) and (2), such as those of Haberman and Renshaw (1996), England and Verrall (1999), England (2002), England and Verrall (2002, 2006) and Kaas et al. (2008). In addition, there has been further analysis of the future payments by calendar years (3) for the Poisson case, conducted in Boj et al. (2014) for the general parametric family (4).

5. ACKNOWLEDGEMENTS

Authors have been supported by the *Spanish Ministerio de Educación y Ciencia* under grant MTM2010-17323 and by the *Generalitat de Catalunya, AGAUR* under grant 2014SGR152.

6. REFERENCES

- Albarrán, I. and Alonso, P., 2010. *Métodos estocásticos de estimaciones de las provisiones técnicas en el marco de Solvencia II*. Cuadernos de la Fundación MAPFRE, 158. Madrid: Fundación MAPFRE Estudios.
- Boj, E., Claramunt, M. M. and Fortiana, J., 2004. *Análisis multivariante aplicado a la selección de factores de riesgo en la tarificación*. Cuadernos de la Fundación MAPFRE, 88. Madrid: Fundación MAPFRE Estudios.
- Boj, E. and Costa, T., 2014. *Modelo lineal generalizado y cálculo de la provisión técnica*. Depósito digital de la Universidad de Barcelona. Colección de objetos y materiales docentes (OMADO). <http://hdl.handle.net/2445/49068>
- Boj, E., Costa, T. and Espejo, J., 2014. Provisiones técnicas por años de calendario mediante modelo lineal generalizado. Una aplicación con RExcel. *Anales del Instituto de Actuarios Españoles*, Tercera Época, 20, 83–116.

- Efron, B. and Tibshirani, J., 1998. *An Introduction to the bootstrap*. New York: Chapman & Hall/CRC.
- England, P. D. and Verrall, R. J., 1999. Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25, 281–293.
- England, P. D., 2002. Addendum to “Analytic and bootstrap estimates of prediction errors in claim reserving”. *Insurance: Mathematics and Economics*, 31, 461–466.
- England, P. D. and Verrall, R. J., 2002. Stochastic claims reserving in general insurance (with discussion). *British Actuarial Journal*, 8, 443–544.
- England, P. D. and Verrall, R. J., 2006. Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1 (II), 221–270.
- Espejo, J., Boj, E. and Costa, T., 2014. *Una aplicación de RExcel para el cálculo de provisiones técnicas con modelo lineal generalizado*. Depósito digital de la Universidad de Barcelona. Colección de Investigación- Software. <http://hdl.handle.net/2445/56230>
- Haberman, S. and Renshaw, A. E., 1996. Generalized linear models and actuarial science. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 45 (4), 407–436.
- Kaas, R., Goovaerts, M., Dhaene, J. and Denuit, M., 2008. *Modern actuarial risk theory: using R. Second edition*. Heidelberg: Springer-Verlag.
- Moreno, F. P., 2013. *Jornada sobre las Directrices de EIOPA de preparación a Solvencia II*. Dirección General de Seguros y Fondos de Pensiones. http://www.dgsfp.mineco.es/sector/documentos/Jornada%20Directrices%20EIOPA%20de%20preparacion%20Solvencia%20II_10-12-2013/Fernando%20Moreno_Jornada%20Directrices%20EIOPA.%20DGSFP-UNESPA.pdf
- McCullagh, P. and Nelder, J., 1989. *Generalized linear models. Second edition*. Londres: Chapman and Hall.
- Parlamento Europeo y Consejo de la Unión Europea, 2009. *Directiva 2009/138/CE del Parlamento Europeo y del Consejo, de 25 de noviembre de 2009*. Diario Oficial de la Unión Europea, L 335, 1–155. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:335:0001:0155:es:PDF>
- R Development Core Team, 2016. *R: a language and environment for statistical computing*. Vienna. Austria. <http://www.R-project.org/>
- Renshaw, A.E., 1989. Chain ladder and interactive modelling (claims reserving and GLIM). *Journal of the Institute of Actuaries*, 116 (III), 559–587.
- Renshaw, A. E., 1994. On the second moment properties and the implementation of certain GLIM based stochastic claims reserving models. *Actuarial Research Paper*, No. 65. Department of Actuarial Science and Statistics, City University, London.
- Taylor, G. and Ashe, F.R., 1983. Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics*, 23, 37–61.

