ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
ΜΠΣ ΑΝΑΛΟΓΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗΣ ΚΙΝΔΥΝΟΥ

ΣΤΟΧΑΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΑΠΟΘΕΜΑΤΟΠΟΙΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Paid Claims Cumulative (£s)

1990	24,246	728,008	1,090,929	1,371,202	1,569,818	2,474,572	2,890,095	3,023,752	3,026,874	3,033,633
1991	26,318	1,083,408	2,481,757	5,475,379	5,990,662	7,936,989	8,909,402	9,037,144	9,211,469	
1992	445,587	1,890,841	6,158,501	9,295,571	12,448,411	15,416,366	15,925,116	16,383,692		
1993	73,916	295,665	1,431,623	1,903,409	3,782,096	7,410,180	8,315,004			
1994	35,850	274,197	629,127	1,175,169	2,479,717	3,875,138				
1995	12,427	218,881	1,179,929	2,098,748	2,853,578					
1996	31,404	106,259	292,575	598,773						
1997	3,652	191,038	516,081							
1998	8,683	354,917								
1999	85,145									

ΠΕΙΡΑΙΑΣ 12/2010

Υπεύθυνος Καθηγητής: κ. Χατζηκωνσταντινίδης

Ονοματεπώνυμο: Αλεξάνδρα Γεωργίου

ΑΜ: ΜΑΕ07020

Περιεχόμενα

Εισαγωγή

Ασφάλεια είναι μια μορφή διαχείρισης κινδύνου, που χρησιμοποιείται πρώτιστα ως προστασία ενάντια στον κίνδυνο πιθανών οικονομικών ζημιών. Η ασφάλεια ορίζεται ως η δίκαιη μεταφορά του κινδύνου πιθανής ζημιάς από μια οντότητα σε άλλη, σε αντάλλαγμα μιας λογικής αμοιβής. Γενικά, είναι μία σύμβαση, στην οποία ένα συμβαλλόμενο μέρος συμφωνεί να πληρώσει για τις οικονομικές ζημιές ενός άλλου συμβαλλόμενου μέρους, ως αποτέλεσμα ενός διευκρινισμένου γεγονότος.

Αποθεματοποίηση ζημιών είναι ο υπολογισμός του απαιτούμενου αποθέματος που πρέπει να διατηρεί μια ασφαλιστική εταιρεία. Γενικά τα αποθέματα ζημιών αντιπροσωπεύουν τα χρήματα που πρέπει να έχει ένας ασφαλιστής για να μπορεί να καλύψει τις μελλοντικές αποζημιώσεις που απορρέουν από τα ασφαλιστικά συμβόλαια που είχαν δημιουργηθεί στο παρελθόν. Οι μέθοδοι υπολογισμού των αποθεμάτων στις γενικές ασφάλειες είναι διαφορετικές από αυτές που χρησιμοποιούνται στις ασφάλειες ζωής, στις συντάξεις και στις ασφάλειες υγείας, αφού οι γενικές ασφάλειες είναι τυπικά πιο σύντομης διάρκειας. Τα αποθέματα υπολογίζονται προβλέποντας τις μελλοντικές ζημιές από τις ζημιές που συνέβησαν στο παρελθόν.

Είναι απαραίτητος ο υπολογισμός μια αξιόπιστης εκτίμησης του αποθέματος που θα μπορεί να καλύψει τις πιθανές ζημιές, έτσι ώστε να διασφαλιστεί η σταθερότητα της εταιρείας. Υπάρχει ένας μεγάλος αριθμός μεθόδων που έχουν αναπτυχθεί για την αποθεματοποίηση των ζημιών. Μια από τις πιο γνωστές μεθόδους αυτές είναι η ντετερμινιστική «*chain ladder*», η οποία αναπτύσσεται λεπτομερώς στο πρώτο μέρος της εργασίας παρακάτω.

Πέρα από μια εκτίμηση των ζημιών που θα πρέπει να είναι σε θέση να πληρώσει η ασφαλιστική εταιρεία, χρειάζεται και άλλες πληροφορίες σχετικά με την αποθεματοποίηση ζημιών όπως είναι ένα μέτρο που να χαρακτηρίζει το λάθος της εκτίμησης. Για το σκοπό αυτό χρησιμοποιούνται οι στοχαστικές μέθοδοι, όπως η μέθοδος Poisson, το Γενικευμένο Γραμμικό Μοντέλο, η μέθοδος Mack και άλλες που θα περιγράψουμε αναλυτικά στο δεύτερο μέρος της εργασίας.

Θα πρέπει εδώ να τονιστεί ότι η αποθεματοποίηση ζημιών στον κλάδο των ιδιωτικών ασφαλειών διέπεται από ειδική νομοθεσία, η οποία απαιτεί οι εταιρείες να τηρούν συνεχώς αποθέματα για το σύνολο του χαρτοφυλακίου τους και να είναι σε θέση σε κάθε περίπτωση εντός διαστήματος τριάντα ημερών να υπολογίσουν τα αποθέματα. Σε περίπτωση που η Επιτροπή Εποπτείας Ιδιωτικής Ασφάλισης κρίνει ότι με τις μεθόδους που χρησιμοποίησε η εταιρεία δεν πιστοποιείται η επάρκεια των αποθεμάτων, δύναται να απαιτήσει από την εταιρεία αυτή υπολογισμό των αποθεμάτων της και με άλλες μεθόδους. Οι νόμοι αυτοί που αφορούν τις ιδιωτικές ασφαλίσεις περιλαμβάνουν άρθρα με το τι πρέπει να συμβαίνει σε κάθε ειδική περίπτωση.

Για τις διάφορες κατηγορίες αποθεμάτων ο νόμος προβλέπει τα παρακάτω:

- Τα Αποθέματα Μη Δεδουλευμένων Ασφαλίστρων υπολογίζονται από τον Αναλογιστή ξεχωριστά για κάθε ασφαλιστικό συμβόλαιο. Στα σύνθετα Ασφαλιστήρια υπολογίζεται και κάθε κατηγορία κινδύνων ξεχωριστά.

- Τα Αποθέματα Κινδύνων εν Ισχύ υπολογίζονται βάση πρόβλεψης για τις ασφαλιστικές αποζημιώσεις και για τα διοικητικά έξοδα, με εξαίρεση τα έξοδα επενδύσεων που προκύπτουν μετά την ημερομηνία υπολογισμού. Ο υπολογισμός γίνεται ξεχωριστά για κάθε κλάδο ασφάλισης.

- Το Μαθηματικό Απόθεμα Γήρατος υπολογίζεται σύμφωνα με τις Ασφαλιστικές Αρχές, οι οποίες εφαρμόζονται στις Ασφαλίσεις Ζωής. Η εκτίμηση του αποθέματος γίνεται για κάθε κλάδο και είδος ασφάλισης.

- Ο υπολογισμός των Αποθεμάτων Εκκρεμών Ζημιών υπολογίζεται βάσει του κόστους που αναμένεται ότι θα προκύψει από ζημιές από τους ασφαλιστικούς κινδύνους που έχουν επέλθει μέχρι την ημερομηνία υπολογισμού του αποθέματος. Αρχικά υπολογίζεται ξεχωριστά για κάθε αναγγελθείσα ζημιά ξεχωριστά, με τη χρήση στατιστικών μεθόδων.

- Το απόθεμα εξισορρόπησης υπολογίζεται με το κλείσιμο των οικονομικών καταστάσεων σωρευτικά, λαμβάνοντας υπόψη τα Αποθέματα Εξισορρόπησης της προηγούμενης χρήσης και το 75% που προκύπτει από το 75% του τεχνικού πλεονάσματος που προκύπτει στον κλάδο.

Κάθε εταιρεία οφείλει να διατηρεί για 15 έτη τουλάχιστον, ηλεκτρονικά αρχεία των δεδομένων ανά κλάδο ασφάλισης, ανά συμβόλαιο και ανά φάκελο ζημιάς. Επίσης υποχρεούται κάθε εταιρεία να διατηρεί το Βιβλίο Τεχνικών Τιμολογίων και Όρων, το οποίο περιλαμβάνει τις εξής ενότητες: το Κεφάλαιο Τεχνικών Σημειωμάτων, το Κεφάλαιο τιμολογίων, το Κεφάλαιο γενικών και ειδικών όρων και το Κεφάλαιο Πινάκων.

ΜΕΡΟΣ 1^{O} (Ντετερμινιστικοί Μέθοδοι)

1.1 Δεδομένα ανάλυσης Ζημιών

Θεωρούμε ότι υπάρχει ένα σύνολο κινδύνων και υποθέτουμε ότι κάθε αξίωση του χαρτοφυλακίου κανονίζεται είτε στο έτος συμβάντος είτε στα επόμενα (n) έτη ανάπτυξης. Το σύνολο έχει διαμορφωθεί είτε από **αυξητικές** (incremental) ζημιές είτε από **αθροιστικές** (cumulative) ζημιές.

1.1.1 Αυξητικές Ζημιές

Για να διαμορφώσουμε ένα σύνολο από αυξητικές ζημιές, θεωρούμε μια ομάδα τυχαίων μεταβλητών $\{Z_{i,k}\} i,k \in \{0,1 \ldots n\}$ και ερμηνεύουμε την τυχαία μεταβλητή $Z_{i,k}$ σαν την ζημιά του έτους συμβάντος i και ως εκ τούτου στο έτος εξέλιξης k και στο ημερολογιακό έτος k+i. Αναφέρουμε το $Z_{i,k}$ σαν την αυξητική ζημιά του έτους συμβάντος i και του έτους εξέλιξης k.

Θεωρούμε ότι οι αυξητικές ζημιές του $Z_{i,k}$ είναι παρατηρήσιμες για τα ημερολογιακά έτη $i+k \leq n$ και ότι δεν είναι παρατηρήσιμες για τα ημερολογιακά έτη $i+k \geq n+1$.
Οι παρατηρήσιμες αυξητικές ζημιές αναπαριστώνται από το ακόλουθο τριγωνικό διάγραμμα (run-off triangle):

Έτος συμβάντος	Έτος εξέλιξης						
	0	1	…	K	…	n-i	…
0	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,k}$	…	$Z_{0,n-i}$	…
1	$Z_{1.0}$	$Z_{1.1}$	…	$Z_{1.k}$	…	$Z_{1.n-i}$	…
⋮	⋮	⋮	…	⋮	…	⋮	
i	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,k}$	…	$Z_{i,n-i}$	
⋮	⋮	⋮	…	⋮			
n-k	$Z_{n-k,0}$	$Z_{n-k,1}$	…	$Z_{n-k,k}$			
⋮	⋮	⋮					
n-1	$Z_{n-1,0}$	$Z_{n-1,1}$					
n	$Z_{n,0}$						

Το πρόβλημα είναι να προβλέψουμε τις μη παρατηρήσιμες αυξητικές ζημιές.

1.1.2 Αθροιστικές ζημιές

Για να διαμορφώσουμε ένα σύνολο από αθροιστικές ζημιές, παίρνουμε μια ομάδα τυχαίων μεταβλητών $\{Si,k\}i,k \in \{0,1$ ...... n} και ερμηνεύουμε την τυχαία μεταβλητή $S_{i,k}$ σαν την ζημιά του έτους συμβάντος i που εμφανίστηκε με καθυστέρηση k το πολύ έτη και ως εκ τούτου όχι αργότερα από το έτος εξέλιξης k. Αναφερόμαστε στο $S_{i,k}$ σαν τη αθροιστική ζημιά του έτους συμβάντος i και του έτους εξέλιξης k, στο $S_{i,n-i}$ σαν τη αθροιστική ζημιά του παρόντος ημερολογιακού έτους n και στο $S_{i,n}$ σαν την τελική συσσωρευμένη ζημιά.

Θεωρούμε ότι οι αθροιστικές ζημιές του $S_{i,\kappa}$ είναι παρατηρήσιμες για τα ημερολογιακά έτη $i+k \leq n$ και ότι δεν είναι παρατηρήσιμες για τα ημερολογιακά έτη $i+k \geq n+1$. Οι παρατηρήσιμες αθροιστικές ζημιές αναπαριστώνται από το ακόλουθο τριγωνικό διάγραμμα

Έτος συμβάντος	Έτος εξέλιξης						
	0	**1**	**…**	**k**	**…**	**n-i**	**…**
0	$S_{0,0}$	$S_{0,1}$	…	$S_{0,k}$	…	$S_{0,n-i}$	…
1	$S_{1.0}$	$S_{1.1}$	…	$S_{1.k}$	…	$S_{1.n-i}$	…
⋮	⋮	⋮	…	⋮	…	⋮	
i	$S_{i,0}$	$S_{i,1}$	…	$S_{i,k}$	…	$S_{i,n-i}$	
⋮	⋮	⋮	…	⋮			
n-k	$S_{n-k,0}$	$S_{n-k,1}$	…	$S_{n-k,k}$			
⋮	⋮	⋮					
n-1	$S_{n-1,0}$	$S_{n-1,1}$					
n	$S_{n,0}$						

Το πρόβλημα είναι να προβλέψουμε τις μη παρατηρήσιμες αθροιστικές ζημιές.

1.1.3 Παρατηρήσεις

Φυσικά, η μοντελοποίηση ενός χαρτοφυλακίου αυξητικών ζημιών είναι ισοδύναμη με την μοντελοποίηση ενός συνόλου αθροιστικών ζημιών:

Οι αθροιστικές ζημιές προέρχονται από τις αυξητικές ζημιές θέτοντας:

$$S_{i,k} = \sum_{l=0}^{k} Z_{i,l}$$

Οι αυξητικές ζημιές προέρχονται από τις αθροιστικές ζημιές θέτοντας:

$$Z_{i,k} = \begin{cases} S_{i,k} & \text{αν } k=0 \\ S_{i,k} - S_{i,k-1} & \text{διαφορετικά} \end{cases}$$

Στην συνέχεια θα εναλλάσσουμε αυξητικές και αθροιστικές ζημιές όπου είναι απαραίτητο.

Αντίστοιχα, η πρόβλεψη των μη παρατηρήσιμων αυξητικών ζημιών είναι απαραίτητα ισοδύναμη της πρόβλεψης των μη παρατηρήσιμων αθροιστικών ζημιών.

- Αν $\{Zi,k\}i,k \in \{0,1,...n\}$, $i+k \geq n+1$ είναι μια ομάδα εκτιμητών των μη παρατηρήσιμων αυξητικών ζημιών, τότε μια ομάδα εκτιμητών μη παρατηρήσιμων αθροιστικών ζημιών επιτυγχάνεται θέτοντας:

$$\widehat{S}_{i,k} = S_{i,n-1} + \sum_{l=n-i+1}^{k} \widehat{Z}_{i,l}$$

- Αν $\left\{\widehat{S}_{i,k}\right\} i,k \in \{0,1,\ldots,n\}, i+k \geq n+1$ είναι μια ομάδα εκτιμητών μη παρατηρήσιμων αθροιστικών ζημιών, τότε μια ομάδα εκτιμητών μη παρατηρήσιμων αυξητικών ζημιών επιτυγχάνεται θέτοντας:

$$\widehat{Z}_{i,k} = \begin{cases} \widehat{S}_{i,n-i+1} - S_{i,n-1} & \text{αν } k=n-i+1 \\ \widehat{S}_{i,k} - \widehat{S}_{i,k-1} & \text{διαφορετικά} \end{cases}$$

Για ευκολία και για να αποφεύγεται η διάκριση των περιπτώσεων, όπως στον προηγούμενο ορισμό, θα αναφέρουμε επίσης τα $Z_{i,n-i}$ και $S_{i,n-i}$ ως εκτιμητές των $Z_{i,n-i}$ και $S_{i,n-i}$, αν φυσικά αυτές οι τυχαίες μεταβλητές είναι παρατηρήσιμες.

Προσοχή: Όταν η πρόβλεψη υπόκειται σε κριτήριο βελτιστοποίησης, δεν μπορεί να διασφαλίζεται σε γενικές γραμμές ότι οι προηγούμενοι τύποι οδηγούν από τις βέλτιστες προβλέψεις αυξητικών ζημιών σε βέλτιστες προβλέψεις αθροιστικών ζημιών ή το αντίστροφο.

Η απαρίθμηση των ετών συμβάντων των ζημιών και των ετών εξέλιξης αρχίζοντας από το 0 αντί για 1, είναι ευρέως αλλά όχι γενικά αποδεκτή [βλ. Taylor (2000), καθώς και Radtke και Schmidt (2004)]. Αυτό είναι χρήσιμο για διάφορους λόγους.

- Για τις ζημιές που εμφανίστηκαν εντός του έτους συμβάντος, η καθυστέρηση της διευθέτησης είναι 0. Είναι επομένως φυσικό να αρχίσει η καταμέτρηση του έτους εξέλιξης με 0.

- Χρησιμοποιώντας την απαρίθμηση των ετών ανάλυσης και για τα εξεταζόμενα έτη συνεπάγεται ότι η αυξητική ή αθροιστική ζημιά του εξεταζόμενου έτους i και του έτους εξέλιξης k είναι παρατηρήσιμη εάν και μόνο αν $i+k \leq n$. Ειδικότερα, οι αθροιστικές ζημιές $S_{i,n-i}$ είναι αυτές του παρόντος ημερολογιακού έτους n και διαδραματίζουν καθοριστικό ρόλο στις περισσότερες μεθόδους ζημιάς αποθεματικών.

Μετά από όλα αυτά, η σημείωση που χρησιμοποιείται απλοποιεί τους μαθηματικούς τύπους.

1.2 Πρότυπα ανάπτυξης

Η χρήση των τριγωνικών διαγραμμάτων στην αποθεματοποίηση ζημιών μπορεί να δικαιολογηθεί μόνο εάν υποτεθεί ότι η ανάπτυξη των ζημιών σε κάθε έτος συμβάντος ακολουθεί ένα πρότυπο ανάπτυξης το οποίο είναι κοινό για όλα τα έτη συμβάντος. Αυτή η ασαφής ιδέα ενός πρότυπου ανάπτυξης μπορεί να τυποποιηθεί με διάφορους τρόπους.

Στην παρούσα ενότητα εξετάζουμε τρεις τύπους των προτύπων ανάπτυξης που είναι επισήμως διακριτά αλλά μπορούν εύκολα να μετατραπούν το ένα στο άλλο. Αυτά τα πρότυπα ανάπτυξης και η ισοδυναμία τους παρέχουν ένα μέσο για τη σύγκριση των διαφόρων μεθόδων αποθεματοποίησης ζημιών.

Η υπόθεση του υποκείμενου μοντέλου ανάπτυξης μπορεί να θεωρηθεί ως ένα πρωτογενές στοχαστικό μοντέλο αποθεματοποίησης ζημιών.

1.2.1 Αυξητικές Ποσοστώσεις

Το πρότυπο ανάπτυξης για αυξητικές ποσοστώσεις συγκρίνει τις αναμενόμενες αυξητικές ζημιές με τις αναμενόμενες τελικές αθροιστικές ζημιές.

Μοντέλο ανάλυσης για αυξητικές ζημιές: Υπάρχουν παράμετροι $\vartheta_0, \vartheta_1, ..., \vartheta_n$ με $\sum_{l=0}^{n} \vartheta_l = 1$ έτσι ώστε η ταυτότητα

$$\vartheta_k = \frac{E\left[Z_{i,k}\right]}{E\left[S_{i,n}\right]}$$

να ισχύει για $k \in (0,1,...,n)$ και για όλα τα $i \in (0,1,...,n)$.

Η υπόθεση σημαίνει ότι για κάθε έτος εξέλιξης $k \in (0,1,...,n)$ οι αυξητικές ποσοστώσεις $\vartheta_{i,k} = \dfrac{E\left[Z_{i,k}\right]}{E\left[S_{i,n}\right]}$ είναι πανομοιότυπες για όλα τα έτη συμβάντων.

Στην περίπτωση του διαγράμματος για τις πληρωμένες ζημιές ή μετρημένες ζημιές, είναι συνήθως λογικό να υποθέτουμε επιπλέον ότι $\vartheta_k > 0$,ισχύει για όλα τα $k \in (0,1,...,n)$. Στην περίπτωση των πραγματοποιημένων ζημιών, αυτή η πρόσθετη υπόθεση ίσως είναι ακατάλληλη επειδή, εξαιτίας της συντηρητικής αποθεματοποίησης, οι αναμενόμενες αυξητικές ζημιές των ετών εξέλιξης $k \in (0,1,...,n)$ ίσως είναι αρνητικές.

1.2.2 Αθροιστικές Ποσοστώσεις

Το πρότυπο ανάπτυξης για αθροιστικές ποσοστώσεις συγκρίνει τις αναμενόμενες αθροιστικές ζημιές με τις αναμενόμενες τελικές αθροιστικές ζημιές.

Πρότυπο ανάπτυξης για αθροιστικές ποσοστώσεις: Υπάρχουν παράμετροι $\gamma_0, \gamma_1, ..., \gamma_n$ έτσι ώστε η ταυτότητα

$$\gamma_k = \frac{E\left[S_{i,k}\right]}{E\left[S_{i,n}\right]}$$

να ισχύει για $k \in (0,1,...,n)$ και για όλα τα $i \in (0,1,...,n)$.

Η υπόθεση σημαίνει ότι για κάθε έτος ανάπτυξης $k \in (0,1,...,n)$ οι αθροιστικές ποσοστώσεις $\gamma_{i,k} = \dfrac{E\left[S_{i,k}\right]}{E\left[S_{i,n}\right]}$ είναι πανομοιότυπες για όλα τα έτη συμβάντων.

Στην περίπτωση του τριγωνικού διαγράμματος για τις πληρωμένες ζημιές ή μετρημένες ζημιές, είναι συνήθως λογικό να υποθέτουμε επιπλέον ότι $0<\gamma_0<\gamma_1<\ldots<\gamma_n$. Στην περίπτωση των πραγματοποιημένων ζημιών, όμως, αυτή η πρόσθετη υπόθεση ίσως είναι ακατάλληλη επειδή, εξαιτίας της συντηρητικής αποθεματοποίησης, η ακολουθία για τις αναμενόμενες αθροιστικές ζημιές ίσως είναι φθίνουσα.

Τα πρότυπα ανάπτυξης για αυξητικές και αθροιστικές ποσοστώσεις μπορούν να μετατραπούν το ένα στο άλλο:

- Αν $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ είναι ένα μοντέλο ανάπτυξης για αυξητικές ζημιές, τότε το πρότυπο ανάπτυξης για αθροιστικές ζημιές βρίσκεται υπολογίζοντας:

$$\gamma_k = \sum_{l=0}^{k} \vartheta_l$$

- Αν $\gamma_0, \gamma_1, \ldots, \gamma_n$ είναι ένα πρότυπο ανάπτυξης για τις αθροιστικές ζημιές, τότε το πρότυπο ανάπτυξης για τις αυξητικές ζημιές υπολογίζεται:

$$\vartheta_k = \begin{cases} \gamma_0 & \text{αν k=0} \\ \gamma_k - \gamma_{k-1} & \text{διαφορετικά} \end{cases}$$

Επιπλέον η συνθήκη $\theta_\kappa>0$ εκπληρώνεται για όλα τα $k \in (0,1,\ldots,n)$ αν και μόνο αν $0 < \gamma_0 < \gamma_1 < \ldots < \gamma_n$.

1.2.3 Συντελεστές

Το πρότυπο ανάπτυξης για παράγοντες συγκρίνει σημαντικές αναμενόμενες αθροιστικές ζημιές:

Πρότυπο ανάπτυξης για τους παράγοντες : υπάρχουν παράμετροι $\varphi_1, \ldots, \varphi_n$ έτσι ώστε η ταυτότητα

$$\varphi_\kappa = \frac{E\left[S_{i,k}\right]}{E\left[S_{i,k-1}\right]}$$

ισχύει για $k \in (0,1,\ldots,n)$ και για όλα τα $i \in (0,1,\ldots,n)$.

Η υπόθεση σημαίνει ότι για κάθε έτος ανάλυσης $k \in (0,1,...,n)$, οι παράγοντες

$$\varphi_{i,\kappa} = \frac{E\left[S_{i,k}\right]}{E\left[S_{i,k-1}\right]}$$

είναι ίδιοι για όλα τα έτη συμβάντων ζημιών.

Στην περίπτωση του διαγράμματος για τις πληρωμένες ζημιές ή μετρήσιμες ζημιές, είναι συνήθως λογικό να υποθέτουμε επιπλέον ότι $\varphi_k > 1$ που ισχύει για όλα τα $k \in (0,1,...,n)$. Στην περίπτωση των πραγματοποιημένων ζημιών, όμως, αυτή η πρόσθετη υπόθεση ίσως είναι ακατάλληλη επειδή, εξαιτίας της συντηρητικής αποθεματοποίησης, η ακολουθία των αναμενόμενων αθροιστικών ζημιών ίσως είναι φθίνουσα.

Τα πρότυπα ανάπτυξης για αθροιστικές ποσοστώσεις και για παράγοντες μπορούν να μετατραπούν το ένα στο άλλο:

- Αν $\gamma_0, \gamma_1, \ldots, \gamma_n$ είναι ένα μοντέλο ανάλυσης για τις αθροιστικές ζημιές, τότε το πρότυπο ανάπτυξης για παράγοντες υπολογίζεται:

$$\varphi_k = \frac{\gamma_k}{\gamma_{k-1}}$$

- Αν $\varphi_1, \ldots, \varphi_n$ είναι ένα πρότυπο ανάπτυξης για παράγοντες, τότε το πρότυπο ανάπτυξης για αθροιστικές ζημιές υπολογίζεται

$$\gamma_k = \prod_{l=k+1}^{n} \frac{1}{\varphi_l}$$

(τέτοιο ώστε γ_n=1)

Επιπλέον η συνθήκη $0 < \gamma_0 < \gamma_1 < \ldots < \gamma_n$ εκπληρώνεται μόνο αν $\varphi_k > 1$ ισχύει για $k \in (0,1,...,n)$.

Συνδυάζοντας αυτό το αποτέλεσμα με εκείνο της προηγούμενης υποενότητας, φαίνεται ότι τα πρότυπα ανάπτυξης για αυξητικές ποσοστώσεις και για παράγοντες μπορούν να μετατραπούν το ένα στο άλλο. Θα παραλείψουμε τους αντίστοιχους τύπους, δεδομένου ότι δεν θα χρειαστούν στη συνέχεια.

1.2.4 Υπολογισμός

Με την πρώτη ματιά υπάρχει μικρή ελπίδα να υπολογίσουμε τις παραμέτρους των μοντέλων ανάλυσης για στοιχειώδεις ή σωρευμένες ποσοστώσεις επειδή οι μόνοι φανεροί εκτιμητές του θ_κ και γ_κ είναι τα παρατηρημένα πηλίκα $Z_{0,k}/S_{0,n}$ και $S_{0,k}/S_{0,n}$, αντίστοιχα.

Ευτυχώς η κατάσταση είναι διαφορετική για το πρότυπο ανάπτυξης για τους παράγοντες: Για κάθε έτος ανάπτυξης $k \in (0,1,...,n)$, κάθε ένας από τους ατομικούς συντελεστές ανάπτυξης

$$\hat{\phi}_{i,\kappa} = \frac{S_{i,k}}{S_{i,k-1}}$$

με $i \in (0,1,...,n-k)$ είναι ένας λογικός εκτιμητής του φ_k, και αυτό ισχύει επίσης για κάθε σταθμισμένο μέσο όρο

$$\hat{\varphi}_k = \sum_{j=0}^{n-k} W_{j,k}\hat{\varphi}_{j,k}$$

με τυχαίες μεταβλητές ή σταθερές ικανοποιώντας τον τύπο

$$\sum_{j=0}^{n-k} W_{j,k} = 1$$

Ο πιο σημαντικός εκτιμητής αυτής της ομάδας είναι ο chain ladder συντελεστής

$$\hat{\varphi}_k^{CL} = \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

που μπορεί να γραφτεί και ως εξής

$$\hat{\varphi}_k^{CL} = \sum_{j=0}^{n-k} \frac{S_{j,k-1}}{\sum_{h=0}^{n-k} S_{h,k-1}} \hat{\varphi}_{j,k}$$

και χρησιμοποιείται στη μέθοδο chain ladder.

Εξαιτίας της αντιστοιχίας των τριών προτύπων ανάπτυξης, είναι ξεκάθαρο ότι με τον ίδιο τρόπο εκτιμητές των συντελεστών μπορούν να μετατραπούν σε εκτιμητές των αθροιστικών ποσοστώσεων και έτσι σε εκτιμητές των σταδιακών ποσοστώσεων.

1.2.5 Παρατηρήσεις

Στην περίπτωση του διαγράμματος για τις πληρωμένες ζημιές ή μετρημένες ζημιές η διαισθητική συνολική ερμηνεία των προτύπων ανάπτυξης των αυξητικών ή αθροιστικών ποσοστώσεων θα είναι η ερμηνεία τους σαν αυξητικών ή αθροιστικών πιθανοτήτων. Αυτή η ερμηνεία είναι χρήσιμη, αλλά όχι αρκετά σωστή επειδή οι παράμετροι του προτύπου ανάπτυξης ορίζονται σαν πηλίκα προσδοκιών αντί για προσδοκίες των πηλίκων και επειδή αυτές οι ποσότητες είναι γενικά διακριτές.

Κάποιος όμως, μπορεί να υποστηρίξει ότι ο καθορισμός των προτύπων ανάπτυξης είναι ακατάλληλος επειδή δεν ανταποκρίνεται διαισθητικά. Στα επόμενα δύο κομμάτια όμως θα αποδειχθεί ότι οι ορισμοί που δόθηκαν είναι παρά ταύτα λογικοί επειδή παρέχουν ένα δυναμικό και ενιαίο περιεχόμενο της έννοιας για την ερμηνεία και τη σύγκριση πολλών μεθόδων και μοντέλων «αποθεματοποίησης ζημιών».

1.3 Μέθοδοι

Το παρόν κεφάλαιο παρέχει μια ενιαία παρουσίαση των πιο σημαντικών μεθόδων «αποθεματοποίησης ζημιών». Το σημείο αρχής είναι μια γενική εκδοχή της μεθόδου των Bornhuetter-Ferguson η οποία σχετίζεται στενά με τη θεωρία του προτύπου ανάπτυξης για αθροιστικές ποσοστώσεις και αποδεικνύεται ότι είναι η αρχή κάτω από την οποία ενοποιούνται ποικίλες άλλες μέθοδοι αποθεματοποίησης ζημιών.

1.3.1 Μέθοδος Bornhuetter-Ferguson

Η μέθοδος Bornhuetter-Ferguson βασίζεται στην υπόθεση ότι υπάρχουν παράμετροι $\alpha_0, \alpha_1, \ldots, \alpha_n$ και $\gamma_0, \gamma_1, \ldots \gamma_n$ με $\gamma_n = 1$ έτσι ώστε να ισχύει η ταυτότητα

$$E\left[S_{i,k}\right] = \gamma_k \alpha_i$$

για i, $k \in (1, \ldots, n)$. Τότε έχουμε

$$E\left[S_{i,n}\right] = \alpha_i$$

και ως εκ τούτου

$$E\left[S_{i,k}\right] = \gamma_k E\left[S_{i,n}\right]$$

τέτοια ώστε οι παράμετροι $\gamma_0, \gamma_1, \ldots \gamma_n$ σχηματίζουν ένα πρότυπο ανάπτυξης για αθροιστικές ποσοστώσεις.

Η μέθοδος Bornhuetter-Ferguson βασίζεται επίσης στην πρόσθετη υπόθεση ότι οι προγενέστεροι εκτιμητές $\hat{a}_0, \hat{a}_1, ..., \hat{a}_n$ των αναμενόμενων τελικών αθροιστικών ζημιών $E\left[S_{i,n}\right]$ και οι προγενέστεροι εκτιμητές $\hat{\gamma}_0, \hat{\gamma}_1, ..., \hat{\gamma}_n$ του προτύπου ανάπτυξης είναι γνωστοί και ότι $\hat{\gamma}_n = 1$.

Σχόλιο: Οι προγενέστεροι εκτιμητές θα μπορούσαν να αποκτηθούν από πληροφορίες που παρέχονται από διάφορες πηγές:

- Εσωτερική πληροφόρηση: Αυτή είναι κάθε πληροφορία που περιλαμβάνεται στο τριγωνικό διάγραμμα του χαρτοφυλακίου που εξετάζουμε. Η εσωτερική πληροφόρηση θα μπορούσε να χρησιμοποιηθεί υπολογίζοντας το μοντέλο ανάλυσης από το δοθέν διάγραμμα.

- Εξωτερική πληροφόρηση: Αυτή είναι κάθε πληροφορία που δεν περιλαμβάνεται στο διάγραμμα του συνόλου που εξετάζουμε. Η εξωτερική πληροφόρηση θα μπορούσε να αποκτηθεί, π.χ. από στατιστικές της αγοράς, από άλλα σύνολα τα οποία κρίνονται ότι είναι παρόμοια με το δοθέν, ή από ασφάλιστρα ή άλλες μονάδες μέτρησης του εξεταζόμενου χαρτοφυλακίου.

Φυσικά οι πρώτοι εκτιμητές μπορούν να αποκτηθούν συνδυάζοντας εσωτερική και εξωτερική πληροφόρηση. Σε κάθε περίπτωση η επιλογή των προγενέστερων εκτιμητών είναι μια σημαντική απόφαση για τον αναλογιστή.

Οι εκτιμητές της B-F μεθόδου των αθροιστικών ζημιών $S_{i,k}$ με $i + k \geq n$ ορίζονται ως εξής

$$\hat{S}_{i,k}^{BF} = S_{i,n-1} + \left(\hat{\gamma}_k - \hat{\gamma}_{n-t}\right)\hat{\alpha}_i$$

Ο ορισμός των εκτιμητών της B-F μεθόδου υπενθυμίζει την ταυτότητα

$$E\left[S_{i,k}\right] = E\left[S_{i,n-i}\right] + \left(\hat{\gamma}_k - \hat{\gamma}_{n-i}\right)\hat{\alpha}_i$$

που είναι μια συνέπεια της υπόθεσης του μοντέλου.

Ο ορισμός των εκτιμητών Bornhuetter-Ferguson δείχνει ότι οι προγενέστεροι εκτιμητές είναι πολύ σημαντικοί για τα νέα έτη συμβάντων ενώ είναι λιγότερο σημαντικοί για τα παλιότερα έτη εξέλιξης. Επίσης στην ακραία περίπτωση όπου οι προγενέστεροι εκτιμητές είναι επαρκώς καθορισμένοι από εξωτερική πληροφόρηση, το κυριότερο μέρος του τριγωνικού διαγράμματος αγνοείται και μόνο οι συνολικές ζημιές του παρόντος έτους χρησιμοποιούνται. Αυτό είναι λογικό όταν η ποιότητα των δεδομένων παρελθόντων ετών είναι φτωχή.

Παράδειγμα Α: Θεωρούμε το ακόλουθο διάγραμμα για αθροιστικές ζημιές το οποίο περιέχει τις αθροιστικές ζημιές του παρόντος έτους και συμπληρώνεται από τους προγενέστερους εκτιμητές των αναμενόμενων τελικών αθροιστικών ζημιών και το πρότυπο ανάπτυξης:

		Έτος εξέλιξης					
Έτος συμβάντος	$\hat{a}_i$	**0**	**1**	**2**	**3**	**4**	**5**
0	3517						3483
1	3981					3844	
2	4598				3977		
3	5658			3880			
4	6214		3261				
5	6325	1889					
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

Υπολογίζοντας τώρα τους εκτιμητές Bornhuetter-Ferguson, το διάγραμμα συμπληρώνεται ακολούθως:

		Έτος εξέλιξης					
Έτος συμβάντος	$\hat{a}_i$	**0**	**1**	**2**	**3**	**4**	**5**
0	3517						3483
1	3981					3844	4043
2	4598				3977	4391	4621
3	5658			3880	4785	4389	5577
4	6214		3261	4442	5436	5995	6306
5	6325	1889	3344	4546	5558	6127	6443
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

Όταν οι αθροιστικές ζημιές του παρόντος έτους κρίνονται ότι είναι αξιόπιστες, ίσως είναι επιθυμητό να τροποποιήσουμε τους εκτιμητές Bornhuetter-Ferguson με σκοπό να ενδυναμώσουμε το βάρος των αθροιστικών ζημιών του παρόντος ημερολογιακού έτους και να μειώσουμε τους προγενέστερους εκτιμητές των αναμενόμενων τελικών αθροιστικών ζημιών. Αυτός ο στόχος μπορεί να επιτευχθεί με την επανάληψη.

Για παράδειγμα αν από τη μεριά της προηγούμενης φόρμουλας οι προγενέστεροι εκτιμητές $\hat{a}_i$ αντικαθίστανται από τους εκτιμητές Bornhuetter-Ferguson , τότε οι προκύπτοντες εκτιμητές είναι οι εκτιμητές

$$\hat{S}_{i,k}^{BF} = S_{i,n-i} + \left(\hat{\gamma}_k - \hat{\gamma}_{n-i}\right)\hat{S}_{i,n}^{BF}$$

όπου στην περίπτωση $\hat{\gamma}_{n-1} < \hat{\gamma}_k$ αυξάνουν το βάρος των αθροιστικών ζημιών του παρόντος έτους και μειώνουν εκείνο των προηγούμενων εκτιμητών των αναμενόμενων τελικών ζημιών.

Πιο γενικά, οι εκτιμητές Bornhuetter-Ferguson με $m \in \mathrm{N}_0$ ορίζονται υπολογίζοντας

$$\hat{S}^{(m)}_{i,k} = \begin{cases} S_{i,n-i} + \left(\hat{\gamma}_k - \hat{\gamma}_{n-i}\right)\hat{\alpha}_i & \text{αν m=0} \\ S_{i,n-i} + \left(\hat{\gamma}_k - \hat{\gamma}_{n-i}\right)\hat{S}^{(m-1)}_{i,n} & \text{αν m} \geq 1 \end{cases}$$

Τότε έχουμε $\hat{S}^{(0)}_{i,k} = \hat{S}^{BF}_{i,k}$ και $\hat{S}^{(1)}_{i,k} = \hat{S}^{BF}_{i,k}$ και η επαγωγή αποδίδει

$$\begin{aligned} \hat{S}^{(m)}_{i,k} &= \left(1-\left(1-\hat{\gamma}_{n-i}\right)^m\right)\hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} + \left(1-\hat{\gamma}_{n-i}\right)^m \hat{S}^{BF}_{i,k} \\ &= \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} + \left(1-\hat{\gamma}_{n-i}\right)^m \left(\hat{S}^{BF}_{i,k} - \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\right) \\ &= \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} + \left(1-\hat{\gamma}_{n-i}\right)^m \left(\hat{\gamma}_k - \hat{\gamma}_{n-i}\right)\left(\hat{\alpha}_i - \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\right) \end{aligned}$$

για όλα τα $m \in \mathrm{N}_0$.

Στην ειδική περίπτωση που $\hat{a}_i = S_{i,n-1} / \hat{\gamma}_{n-1}$ ή $\hat{\gamma}_{n-1} = 1$, η επανάληψη είναι χωρίς ενδιαφέρον επειδή σε αυτή την περίπτωση η ταυτότητα

$$\hat{S}^{(m)}_{i,k} = \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}$$

ισχύει για τα $m \in \mathrm{N}_0$. Αντίθετα, η επανάληψη είναι αυξημένου ενδιαφέροντος στην περίπτωση όπου $0<\gamma_{n-1}<1$ επειδή σε αυτή των περίπτωση πετυχαίνουμε

$$\lim_{m\to\infty} \hat{S}^{(m)}_{i,k} = \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}$$

Και η σύγκλιση της ακολουθίας των επαναλαμβανόμενων εκτιμητών Bornhuetter-Ferguson είναι μονότονη αλλά βαίνει αύξουσα ή φθίνουσα.

Παράδειγμα Β: Ο ακόλουθος πίνακας περιέχει τους προγενέστερους εκτιμητές των αναμενόμενων τελικών αθροιστικών ζημιών, τους επαναλαμβανόμενους εκτιμητές Bornhuetter-Ferguson

$$\hat{S}_{i,n}^{(m)} = \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\left(1-\hat{\gamma}_{n-i}\right)^{m+1}\left(\hat{\alpha}_i - \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\right)$$

και τα όρια τους

	Προγενέστεροι	Μεταγενέστεροι Εκτιμητές BF									
Έτος συμβάντος	$\hat{a}_i$	$\hat{S}_{i,5}^{(0)}$	$\hat{S}_{i,5}^{(1)}$	$\hat{S}_{i,5}^{(2)}$	$\hat{S}_{i,5}^{(3)}$	$\hat{S}_{i,5}^{(4)}$	$\hat{S}_{i,5}^{(5)}$		$\hat{S}_{i,5}^{(10)}$		**Όρια**
0	3517	3483	3483	3483	3483	3483	3483	…	3483	…	3483
1	3981	4043	4046	4046	4046	4046	4046	…	4046	…	4046
2	4598	4621	4623	4624	4624	4624	4624	…	4624	…	4624
3	5658	5577	5553	5346	5544	5543	5543	…	5543	…	5543
4	6214	6306	6351	6373	6384	6389	6392	…	6394	…	6394
5	6325	6443	6528	6589	6633	6664	6687	…	6730	…	6746

Τα επαναλαμβανόμενα βήματα 0 και 1 ανταποκρίνονται στη μέθοδο Bornhuetter-Ferguson και στην Benktander- Hovinen αντίστοιχα. Ο πίνακας δείχνει ότι η σύγκλιση είναι μονότονη αλλά βαίνει αύξουσα ή φθίνουσα και ότι η σύγκλιση είναι γρήγορη για παλιά έτη ανάπτυξης και αργή για νέα έτη συμβάντων.

1.3.2 Μέθοδος Ανάπτυξης ζημιών Loss-development

Η μέθοδος αυτή βασίζεται στην υπόθεση ότι υπάρχουν παράμετροι $\gamma_0,\gamma_1,\ldots,\gamma_n$ με $\gamma_n=1$ τέτοια ώστε η ταυτότητα

$$E\left[S_{i,k}\right]=\gamma_k E\left[S_{i,n}\right]$$

να ισχύει για όλα τα $i,k\in(0,1,....,n)$. Τότε οι παράμετροι $\gamma_0,\gamma_1,...,\gamma_n$ σχηματίζουν ένα πρότυπο ανάπτυξης για αθροιστικές ποσοστώσεις.

Η μέθοδος ανάπτυξης ζημιών βασίζεται στην υπόθεση ότι οι πρώτοι εκτιμητές $\hat{\gamma}_0,\hat{\gamma}_1,\ldots,\hat{\gamma}_n$ του μοντέλου ανάλυσης δίνονται και ότι $\hat{\gamma}_n=1$.

Οι εκτιμητές ανάλυσης ζημιών για αθροιστικές ζημιές $S_{i,k}$ με $i+k\geq n$ ορίζονται ως

$$\hat{S}_{i,k}^{LD}=\hat{\gamma}_K\frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}$$

Ο ορισμός των εκτιμητών ανάλυσης ζημιών υπενθυμίζει την ταυτότητα

$$E\left[S_{i,k}\right]=\gamma_k\frac{E\left[S_{i,n}\right]}{\gamma_{n-i}}$$

που είναι συνέπεια του μοντέλου της υπόθεσης.

Όταν συγκρίνονται με τους εκτιμητές Bornhuetter-Ferguson, η σημασία των αθροιστικών ζημιών του παρόντος ημερολογιακού έτους και των προηγούμενων εκτιμητών του μοντέλου ανάπτυξης είναι αυξανόμενη στους εκτιμητές ανάπτυξης ζημιών επειδή οι τελευταίοι δεν περιέχουν κανένα προηγούμενο εκτιμητή των αναμενόμενων τελικών αθροιστικών ζημιών.

Παράδειγμα C: Θεωρούμε το ακόλουθο τριγωνικό διάγραμμα για αθροιστικές ζημιές το οποίο περιέχει τις αθροιστικές ζημιές του παρόντος έτους και συμπληρώνεται από τους προηγούμενους εκτιμητές του προτύπου ανάπτυξης.

Έτος συμβάντος	**Έτος εξέλιξης**					
	0	**1**	**2**	**3**	**4**	**5**
0						3483
1					3844	
2				3977		
3			3880			
4		3261				
5	1889					
$\hat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

Υπολογίζοντας τους εκτιμητές της ανάλυσης ζημιών, το διάγραμμα συμπληρώνεται ως εξής:

Έτος συμβάντος	**Έτος εξέλιξης**					
	0	**1**	**2**	**3**	**4**	**5**
0						3483
1					3844	*4046*
2				3977	*4393*	*4624*
3			3880	*4767*	*5266*	*5543*
4		3261	*4476*	*5499*	*6074*	*6394*
5	1889	*3440*	*4722*	*5802*	*6409*	*6746*
$\hat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

Οι εκτιμητές ανάλυσης ζημιών μπορούν να γραφούν ως εξής:

$$\widehat{S}_{i,k}^{LD} = S_{i,n-i} + \left(\widehat{\gamma}_k - \widehat{\gamma}_{n-i}\right)\widehat{S}_{i,n}^{LD}$$

Αυτό δείχνει ότι εκτιμητές ανάπτυξης ζημιών δεν είναι τίποτε άλλο από τους εκτιμητές Bornhuetter-Ferguson σε σχέση με τους προγενέστερους εκτιμητές $\widehat{\alpha}_i^{LD} = \widehat{S}_i^{LD}$ των αναμενόμενων τελικών αθροιστικών ζημιών. Με άλλα λόγια, η μέθοδος ανάλυσης ζημιών είναι μια συγκεκριμένη περίπτωση του μοντέλου Bomhuetter-Ferguson με τους προγενέστερους εκτιμητές των αναμενόμενων αθροιστικών ζημιών οι οποίοι βασίζονται σε εξωτερική και εσωτερική πληροφόρηση.

Επιπλέον, στην περίπτωση όπου $0<\widehat{\gamma}_{n-1}<1$, οι εκτιμητές ανάπτυξης ζημιών είναι ακριβώς τα όρια των ακολουθιών των επαναλαμβανόμενων εκτιμητών B-F σε σχέση με τους τυχαίους προγενέστερους εκτιμητές των αναμενόμενων αθροιστικών τελικών ζημιών, όπως έχει δειχθεί στο τμήμα 4.1.

1.3.3 Μέθοδος Chain Ladder

Η μέθοδος Chain Ladder βασίζεται στην υπόθεση ότι υπάρχουν παράμετροι $\varphi_1,\ldots,\varphi_n$ έτσι ώστε να ισχύει η ταυτότητα

$$E\left[S_{i,k}\right] = \varphi_k E\left[S_{i,k-1}\right]$$

για όλα τα i ϵ{0, 1, …, n}και k ϵ{0, 1, …, n}. Τότε οι παράμετροι $\varphi_1,\ldots,\varphi_n$ σχηματίζουν ένα πρότυπο ανάπτυξης για παράγοντες.

Οι εκτιμητές Chain Ladder των αθροιστικών ζημιών $S_{i,k}$ με $i+k\geq n$ ορίζονται ως

$$\widehat{S}_{i,k}^{CL} = S_{i,n-1} \prod_{l=n-i+1}^{k} \widehat{\varphi}_l^{CL}$$

όπου

$$\hat{\varphi}_k^{CL} = \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

είναι ο chain ladder συντελεστής. Ο ορισμός των εκτιμητών chain ladder υπενθυμίζει την ταυτότητα

$$E\left[S_{i,k}\right] = E\left[S_{i,n-i}\right] \prod_{l=n-i+1}^{k} \varphi_l$$

που είναι μια συνέπεια του μοντέλου υπόθεσης.

Σε σύγκριση με τους εκτιμητές της ανάπτυξης ζημιών, είναι αξιοσημείωτο ότι οι εκτιμητές chain ladder δεν προσδιορίζονται από τις αθροιστικές ζημιές του παροντικού ημερολογιακού έτους, αλλά μέσω των παραγόντων chain ladder, από όλες τις αθροιστικές ζημιές του τριγωνικού σχεδιαγράμματος.

Παράδειγμα D. Θεωρούμε το ακόλουθο τριγωνικό διάγραμμα για τις αθροιστικές ζημιές:

Έτος συμβάντος	**Έτος εξέλιξης**					
	0	**1**	**2**	**3**	**4**	**5**
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	
2	1265	2433	3233	3977		
3	1490	2873	3880			
4	1725	3261				
5	1889					

Υπολογίζοντας αρχικά τους παράγοντες «chain ladder» και έπειτα τους εκτιμητές «chain ladder», το τριγωνικό διάγραμμα συμπληρώνεται όπως ακολούθως:

Έτος συμβάντος	**Έτος εξέλιξης**					
	0	**1**	**2**	**3**	**4**	**5**
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	*4013*
2	1265	2433	3233	3977	*4454*	*4650*
3	1490	2873	3880	*4780*	*5354*	*5590*
4	1725	3261	*4334*	*5339*	*5980*	*6243*
5	1889	*3587*	*4767*	*5873*	*6578*	*6867*

Έχει τονιστεί ότι τα διαφορετικά πρότυπα ανάπτυξης και οι εκτιμητές τους μπορούν να μετατρέπονται ο ένας στον άλλον. Συγκεκριμένα, έστω

$$\gamma_k = \prod_{l=k+1}^{n} \frac{1}{\varphi_l}$$

ότι μετατρέπει ένα πρότυπο ανάπτυξης για παράγοντες σε ένα πρότυπο ανάπτυξης για τις αθροιστικές ποσοστώσεις και έστω

$$\hat{\gamma}_k = \prod_{l=k+1}^{n} \frac{1}{\hat{\varphi}_l}$$

ότι μετατρέπει τους εκτιμητές του προτύπου ανάπτυξης των παραγόντων σε εκτιμητές προτύπου ανάπτυξης για τις αθροιστικές ποσοστώσεις. Έτσι, για

$$\hat{\gamma}_k^{CL} = \prod_{l=k+1}^{n} \frac{1}{\hat{\varphi}_l}$$

οι εκτιμητές «chain ladder» μπορούν να γραφούν ως:

$$\hat{S}_{i,k}^{CL} = \hat{\gamma}_k^{CL} \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}^{CL}}$$

Αυτό δείχνει ότι οι εκτιμητές «chain ladder» δεν είναι τίποτα άλλο από τους εκτιμητές της ανάπτυξης ζημιών σε σχέση με τις αθροιστικές ποσοστώσεις της «chain ladder» $\hat{\gamma}_k^{CL}$ ως προγενέστεροι εκτιμητές των αθροιστικών ποσοστώσεων. Επιπλέον, εμείς έχουμε

$$\hat{S}_{i,k}^{CL} = S_{i,n-1} + (\hat{\gamma}_k^{CL} - \hat{\gamma}_{n-i}^{CL})\hat{S}_{i,n}^{CL}$$

Αυτό δείχνει ότι οι εκτιμητές «chain ladder» είναι ακριβώς οι εκτιμητές «Bornhuetter-Ferguson» σε σχέση με τους προγενέστερους εκτιμητές $\hat{\gamma}_k^{CL}$ των αθροιστικών ποσοστώσεων και των προγενέστερων εκτιμητών

$$\hat{\alpha}_i^{CL} = \hat{S}_{i,n}^{CL}$$

των τελευταίων αναμενόμενων αθροιστικών ζημιών. Με άλλα λόγια η μέθοδος chain ladder είναι μια ειδική περίπτωση της μεθόδου ανάπτυξης ζημιών και ως εκ τούτου της μεθόδου B-F με προγενέστερους εκτιμητές του προτύπου ανάπτυξης και των αναμενόμενων τελευταίων αθροιστικών ζημιών οι οποίοι είναι πλήρως βασισμένοι σε εσωτερικές πληροφορίες.

Αυτή η μέθοδος chain ladder μπορεί να τροποποιηθεί αντικαθιστώντας τους παράγοντες chain ladder $\hat{\varphi}_k^{CL}$ από άλλους εκτιμητές της μορφής

$$\hat{\varphi}_k = \sum_{j=0}^{n-k} W_{j,k}\hat{\varphi}_{j,k}$$

Με τυχαίες μεταβλητές (ή σταθερές) που ικανοποιούν την σχέση

$$\sum_{j=0}^{n-k} W_{j,k} = 1$$

1.3.4 Η μέθοδος «grossing – up»

Η μέθοδος «grossing – up» βασίζεται στην υπόθεση ότι υπάρχουν παράμετροι γ_0, γ_1, …,γ_n με γ_n=1 έτσι ώστε να ισχύει η ταυτότητα

$$E\left[S_{i,k}\right]=\gamma_k E\left[S_{i,n}\right]$$

για όλα τα i, k ϵ{0, 1, …, n} . Τότε οι παράμετροι γ_0, γ_1, …,γ_n αποτελούν ένα πρότυπο ανάπτυξης για τις αθροιστικές ποσοστώσεις.

Οι εκτιμητές «grossing-up» των αθροιστικών ζημιών $S_{i,k}$ με i+k≥n ορίζονται ως

$$\hat{S}_{i,k}^{GU}=\hat{\gamma}_k^{GU}\frac{S_{n-i}}{\hat{\gamma}_{n-i}^{GU}}$$

όπου

$$\hat{\gamma}_k^{GU}=\begin{cases}1, & \text{αν } k=n\\ \dfrac{\sum_{j=0}^{n-k-1} S_{i,k}}{\sum_{j=0}^{n-k-1}\hat{S}_{j,n}^{GU}}, & \text{διαφορετικά}\end{cases}$$

είναι η αθροιστική ποσόστωση «grossing-up» του έτους ανάπτυξης k. Ο ορισμός των εκτιμητών «grossing-up» μας υπενθυμίζει την ταυτότητα

$$E\left[S_{i,k}\right]=\gamma_k\frac{E\left[S_{i,n-i}\right]}{\gamma_{n-i}}$$

η οποία είναι μια συνέπεια της υπόθεσης του μοντέλου.

Ο υπολογισμός των αθροιστικών ποσοστώσεων «grossing-up» και των εκτιμητών «grossing-up» για τις τελικές αθροιστικές ζημιές προέρχονται από αναδρομή κατά τα έτη ατυχήματος, η οποία αποδίδει

$$\hat{\gamma}_n^{GU} = 1 \qquad \text{και} \quad \hat{S}_{0,n}^{GU} = S_{0.n}$$

$$\hat{\gamma}_{n-1}^{GU} = \frac{S_{0,n-1}}{\hat{S}_{0,n}^{GU}} \qquad \text{και} \quad \hat{S}_{1,n}^{GU} = \frac{S_{1,n-1}}{\hat{\gamma}_{n-1}^{GU}}$$

$$\hat{\gamma}_{n-2}^{GU} = \frac{S_{0,n-2} + S_{1,n-2}}{\hat{S}_{0,n}^{GU} + \hat{S}_{1,n}^{GU}} \qquad \text{και} \quad \hat{S}_{2,n}^{GU} = \frac{S_{2,n-2}}{\hat{\gamma}_{n-2}^{GU}}$$

Όπως παρατηρούμε από τον ορισμό, οι εκτιμητές «grossing-up» δεν είναι τίποτα άλλο από τους εκτιμητές της ανάπτυξης ζημιών σε σχέση με τις αθροιστικές ποσοστώσεις «grossing-up» $\hat{\gamma}_k^{GU}$ ως προγενέστερους εκτιμητές των αθροιστικών ποσοστώσεων. Επιπλέον, έχουμε

$$\hat{S}_{i,k}^{GU} = S_{i,n-1} + (\hat{\gamma}_k^{GU} - \hat{\gamma}_{n-i}^{GU})\hat{S}_{i,n}^{GU}$$

η οποία δείχνει ότι οι εκτιμητές «grossing-up» είναι ακριβώς οι εκτιμητές «Bornhuetter-Ferguson» σε σχέση με τους προγενέστερους εκτιμητές $\hat{\gamma}_k^{GU}$ των αθροιστικών ποσοστώσεων και τους προγενέστερες εκτιμητές

$$\hat{\alpha}_i^{GU} = \hat{S}_{i,n}^{GU}$$

των τελευταίων αναμενόμενων αθροιστικών ζημιών. Με άλλα λόγια, η μέθοδος «grossing-up» είναι μια παρόμοια περίπτωση της μεθόδου ανάπτυξης ζημιών και ως εκ τούτου της μεθόδου «Bornhuetter-Ferguson» με προγενέστερους εκτιμητές του προτύπου ανάπτυξης και με αναμενόμενες τελευταίες ζημιές, οι οποίες είναι πλήρως βασισμένες σε εσωτερικές πληροφορίες.

Αφού η τελευταία παρατήρηση εφαρμόζεται και στους εκτιμητές «chain ladder», η ερώτηση που προκύπτει είναι αν υπάρχει καμία διαφορά ανάμεσα στους εκτιμητές «grossing-up» και «chain ladder». Η απάντηση σε αυτή την ερώτηση είναι ότι δεν υπάρχει καμία διαφορά, αφού οι αθροιστικές ποσοστώσεις «grossing-up» και οι αθροιστικές ποσοστώσεις «chain ladder» είναι πανομοιότυπες για όλα τα έτη ανάπτυξης.

Έτσι η «grossing-up» μέθοδος παρέχει μια υπολογιστική εναλλακτική στην μέθοδο «chain ladder», αλλά αυτή η εναλλακτική φαίνεται να εμφανίζει μικρό πρακτικό ενδιαφέρον. Η αναδιατύπωση της μεθόδου «chain ladder» παρεχόμενη με την μέθοδο «grossing-up» είναι όμως, σημαντικού ενδιαφέροντος σε σχέση με τις μεθόδους:

Αρχικά, μεταξύ όλων των μεθόδων για τις συνολικές ζημιές που εξετάζονται εδώ, η μέθοδος «chain ladder» εμφανίζεται μοναδική αφού χρησιμοποιεί εκτιμητές ενός προτύπου ανάπτυξης για παράγοντες αντί των αθροιστικών ποσοστώσεων, αλλά η ισοδυναμία του με τη μέθοδο «grossing-up» δείχνει ότι αυτή η μοναδικότητα είναι μόνο εξαιτίας της πιο έξυπνης διατύπωσης ενός αλγόριθμου που αποφεύγει τις αναδρομές και είναι ως εκ τούτου πιο εύκολα κατανοητή.

Δεύτερον, η μέθοδος «grossing-up» παρέχει μία ουσιαστική σύνδεση μεταξύ της μεθόδου «chain ladder» και της μεθόδου του οριακού ποσού (marginal-sum method).

1.3.5 Η μέθοδος Οριακού-Ποσού (Marginal-Sum Method)

Η μέθοδος οριακού-ποσού βασίζεται στην υπόθεση ότι υπάρχουν παράμετροι α_0, α_1, …, α_n και θ_0, θ_1, …,θ_n με $\sum_{l=0}^{n} \theta_l = 1$ έτσι ώστε η ταυτότητα

$$E\left[Z_{i,k}\right] = \vartheta_k a_i$$

να ισχύει για όλα τα i, k ϵ{0, 1, …, n}.

Η άθροιση αποδίδει

$$E\left[S_{i,n}\right]=a_i$$

και ως εκ τούτου

$$E\left[Z_{i,k}\right]=\vartheta_k E\left[S_{i,n}\right]$$

έτσι ώστε οι παράμετροι θ_0, θ_1, …,θ_n να σχηματίζουν ένα πρότυπο ανάπτυξης για πρόσθετες ποσοστώσεις.

Παρατηρήσιμες τυχαίες μεταβλητές $\hat{\alpha}_0^{MS}, \hat{\alpha}_1^{MS}, ..., \hat{\alpha}_n^{MS}$ και $\hat{\theta}_0^{MS}, \hat{\theta}_1^{MS}, ..., \hat{\theta}_n^{MS}$ λέγονται εκτιμητές marginal-sum method αν αυτές είναι λύσεις των marginal-sum εξισώσεων

$$\sum_{l=0}^{n-i} \hat{\alpha}_i \hat{\vartheta}_l = \sum_{l=0}^{n-i} Z_{i.l}$$

για όλα τα i ϵ{0, 1, …, n}και

$$\sum_{j=0}^{n-k} \hat{\alpha}_j \hat{\vartheta}_k = \sum_{j=0}^{n-k} Z_{j,k}$$

για k ϵ{0, 1, …, n}καθώς και

$$\sum_{l=0}^{n} \hat{\vartheta}_l = 1$$

Οι marginal-sum εξισώσεις μας υπενθυμίζουν τις ταυτότητες

$$\sum_{l=0}^{n-i} \alpha_i \vartheta_l = \sum_{l=0}^{n-i} E[Z_{i.l}]$$

και

$$\sum_{j=0}^{n-k} \alpha_j \vartheta_k = \sum_{j=0}^{n-k} E[Z_{j,k}]$$

καθώς και

$$\sum_{k=0}^{n} \hat{\vartheta}_k = 1$$

που έπονται από τις υποθέσεις του μοντέλου.

Η ερώτηση που ανακύπτει είναι εάν οι marginal-sum εκτιμητές υπάρχουν και είναι μοναδικοί. Η ερώτηση στην απάντηση είναι καταφατική: οι marginal-sum εκτιμητές υπάρχουν και είναι μοναδικοί, και ικανοποιούν την

$$\hat{\alpha}_i^{MS} = \hat{S}_{i,n}^{GU}$$

και

$$\hat{\vartheta}_k^{MS} = \begin{cases} \hat{\gamma}_0^{GU} & \text{αν } k=0 \\ \hat{\gamma}_k^{GU} - \hat{\gamma}_{k-1}^{GU} & \text{αν } k \geq 1 \end{cases}$$

Λαμβάνοντας υπόψη την συζήτηση της μεθόδου «grossing-up», οι προηγούμενες ταυτότητες συνεπάγονται ότι οι marginal-sum εκτιμητές ικανοποιούν την παρακάτω σχέση:

$$\hat{\alpha}_i^{MS} = \hat{S}_{i,n}^{CL}$$

και

$$\hat{\vartheta}_k^{MS} = \begin{cases} \hat{\gamma}_0^{CL} & \text{αν } k=0 \\ \hat{\gamma}_k^{CL} - \hat{\gamma}_{k-1}^{CL} & \text{αν } k \geq 1 \end{cases}$$

Έτσι, έστω

$$\hat{\gamma}_k^{MS} = \sum_{l=0}^{k} \hat{\vartheta}_l^{MS}$$

Παρατηρούμε

$$\hat{\gamma}_k^{MS} = \hat{\gamma}_k^{CL}$$

για όλα τα k ϵ{0, 1, …, n}.

Οι marginal-sum εκτιμητές των συνολικών ζημιών $S_{i,k}$ με i+k≥n ορίζονται ως

$$\hat{S}_k^{MS} = \hat{\gamma}_k^{MS} \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}^{MS}}$$

Τότε έχουμε

$$\hat{S}_{i,k}^{MS} = \hat{S}_{i,k}^{CL}$$

Αυτό δείχνει ότι η μέθοδος marginal-sum είναι ισοδύναμη με την μέθοδο chain ladder.

1.3.6 Cape-Cod Μέθοδος

Η μέθοδος Cape-Cod βασίζεται στην υπόθεση ότι υπάρχουν παράμετροι γ_0, γ_1, …,γ_n με γ_n=1 έτσι ώστε η ταυτότητα

$$E[S_{i,k}] = \gamma_k E[S_{i,n}]$$

να ισχύει για όλα τα i, k ϵ{0, 1, …, n}. Τότε οι παράμετροι $\gamma_0, \gamma_1, \ldots, \gamma_n$ αποτελούν ένα πρότυπο ανάπτυξης για τις αθροιστικές ποσοστώσεις.

Η μέθοδος Cape-Cod βασίζεται επίσης στην υπόθεση ότι τα ασφάλιστρα ή άλλες μονάδες μέτρησης $\pi_o, \pi_1, \ldots, \pi_n \in (0, \infty)$ των χρόνων ατυχήματος είναι γνωστά, ότι οι αναμενόμενοι τελευταίοι συνολικοί συντελεστές ζημιών είναι

$$\kappa_i = E\left[\frac{S_{i,n}}{\pi_i}\right]$$

ίδιες για όλα τα χρόνια ατυχημάτων και ότι οι προγενέστεροι εκτιμητές $\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_n$ του προτύπου ανάπτυξης δίνονται και ικανοποιούν την $\hat{\gamma}_n = 1$.

Οι εκτιμητές Cape-Cod των αθροιστικών ζημιών $S_{i,k}$ με $i+k \geq n$ ορίζονται ως

$$\hat{S}_{i,k}^{CC} = S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\pi_i \hat{\kappa}^{CC}$$

όπου

$$\hat{\kappa}^{CC} = \frac{\sum_{j=0}^{n} S_{j,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j}\pi_j}$$

είναι Cape-Cod συντελεστής ζημιών, ο οποίος είναι ένας εκτιμητής του αναμενόμενου τελευταίου συντελεστή ζημιών (κοινό σε όλα τα χρόνια ατυχήματος).

Οι εκτιμητές Cape-Cod δεν είναι τίποτα άλλο από τους εκτιμητές Bornhuetter-Ferguson εκτιμητές σε σχέση με τους προγενέστερους εκτιμητές

$$\hat{\alpha}_i^{CC} = \pi_i \hat{\kappa}^{CC}$$

των αναμενόμενων τελευταίων αθροιστικών ζημιών οι οποίες βασίζονται και σε εξωτερικές και σε εσωτερικές πληροφορίες.

Παράδειγμα Ε. Θεωρούμε το ακόλουθο μειωμένο τριγωνικό σχεδιάγραμμα για τις αθροιστικές ζημιές οι οποίες περιέχουν τις αθροιστικές ζημιές του παροντικού ημερολογιακού έτους και συμπληρώνεται από τα ασφάλιστρα και τους προγενέστερους εκτιμητές του πρότυπου ανάπτυξης:

		Έτος εξέλιξης					
Έτος συμβάντος	π_i	**0**	**1**	**2**	**3**	**4**	**5**
0	4025						3483
1	4456					3844	
2	5315				3977		
3	5986			3880			
4	6939		4261				
5	8158	1889					
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

Το προηγούμενο τρίγωνο διαφέρει από αυτά που εμφανίστηκαν πριν από αυτό ως προς την αξία του $S_{4,1}$ που είναι 4261 έναντι του 3261, το οποίο υποδεικνύει ότι ίσως υπάρχει μία ακραία τιμή στο έτος ατυχήματος 4. Χρησιμοποιώντας τον πίνακα

i	$S_{i,5-i}$	$\hat{\gamma}_{5-i}$	π_i	$\hat{\gamma}_{5-i}\pi_i$
0	3483	1.000	4025	4025
1	3844	0.950	4456	4233
2	3977	0.860	5315	4571
3	3880	0.700	5986	4190
4	4261	0.510	6939	3539
5	1889	0.280	8158	2284
Σ	**21334**			**22842**

παρατηρούμε ότι $\hat{\kappa}^{CC}$ =0.934.

Υπολογίζοντας τώρα τους προγενέστερους εκτιμητές των αναμενόμενων τελευταίων αθροιστικών ζημιών και τους Cape-Cod εκτιμητές ο πίνακας συμπληρώνεται ως εξής:

Έτος συμβάντος i	$\widehat{\alpha}_i$	**0**	**1**	**2**	**3**	**4**	**5**
				Έτος εξέλιξης k			
0	3758						3483
1	4162					3844	*4052*
2	4964				3977	*4424*	*4672*
3	5591			3880	*4775*	*5278*	*5557*
4	6481		4261	*5492*	*6529*	*7113*	*7437*
5	7619	1889	*3641*	*5089*	*6308*	*6994*	*7375*
$\widehat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

Ο προηγούμενος πίνακας πρέπει να συγκριθεί με τον ακόλουθο, ο οποίος είναι του ίδιου τριγωνικού διαγράμματος συμπληρωμένο με τους εκτιμητές ανάπτυξης ζημιών:

Έτος συμβάντος i	**0**	**1**	**2**	**3**	**4**	**5**
0						3483
1					3844	*4046*
2				3977	*4393*	*4624*
3			3880	*4767*	*5266*	*5543*
4		4261	*5948*	*7185*	*7937*	*8355*
5	1889	*3440*	*4722*	*5802*	*6409*	*6746*
$\widehat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

Το παράδειγμα υποδεικνύει ότι η ανάπτυξη των εκτιμητών Cape-Cod κατά τα χρόνια ατυχημάτων είναι πιο ομαλά από την ανάπτυξη των εκτιμητών ανάπτυξης ζημιών το οποίο σημαίνει ότι η μέθοδος Cape-Cod μειώνει τις επιδράσεις των ακραίων τιμών.

Το αποτέλεσμα της εξομάλυνσης εξαρτάται από τα ασφάλιστρα ή άλλες μονάδες μέτρησης που χρησιμοποιούνται.

Οι παρακάτω θεωρήσεις ίσως βοηθήσουν στο αποτέλεσμα εξομάλυνσης της Cape-Cod μεθόδου: Υποθέτουμε ότι για κάθε χρόνο ατυχήματος i, ο αναμενόμενος τελικός συνολικός συντελεστής ζημιών εκτιμάται από:

$$\hat{\kappa}_i = \frac{\hat{S}_{i,n}^{LD}}{\pi_i} = \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}\pi_i}$$

Τότε ο συντελεστής ζημιών Cape-Cod μπορεί να γραφτεί ως ο σταθμισμένος μέσος :

$$\hat{\kappa}^{CC} = \frac{\sum_{j=0}^{n} S_{j,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j}\pi_j} = \sum_{j=0}^{n} \frac{\hat{\gamma}_{n-j}\pi_j}{\sum_{b=0}^{n} \hat{\gamma}_{n-b}\pi_b} \hat{\kappa}_j$$

και η ταυτότητα

$$S_{i,n-i} = \hat{\gamma}_{n-i}\pi_i \hat{\kappa}_i$$

προτείνει την αποσύνθεση των αθροιστικών ζημιών $S_{i,n-i}$ του παροντικού ημερολογιακού έτους στο κανονικό μέρος

$$T_{i,n-i} = \hat{\gamma}_{n-i}\pi_i \hat{\kappa}^{CC}$$

και στο αποτέλεσμα των ακραίων τιμών

$$X_{i,n-i} = S_{i,n-i} - T_{i,n-i}$$

και τότε να εφαρμόσετε την μέθοδο ανάπτυξης ζημιών στο κανονικό μέρος ενώ διατηρείτε το αποτέλεσμα των ακραίων τιμών σταθερό για όλα τα ακόλουθα χρόνια ανάπτυξης. Αφού:

$$\hat{T}_{i,k}^{LD} + X_{i,n-i} = \hat{\gamma}_k \frac{T_{i,n-i}}{\hat{\gamma}_{n-i}} + (S_{i,n-i} - T_{i,n-i})$$

$$= S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i}) \frac{T_{i,n-i}}{\hat{\gamma}_{n-i}}$$

$$= S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i}) \pi_i \hat{\kappa}^{CC}$$

$$= \hat{S}_{i,k}^{CC}$$

παρατηρούμε ότι οι εκτιμητές που απορρέουν είναι ακριβώς οι Cape-Cod εκτιμητές.

Η Cape-Cod μέθοδος μπορεί να τροποποιηθεί αντικαθιστώντας τον Cape-Cod συντελεστή ζημιών $\hat{\kappa}^{CC}$ από άλλους εκτιμητές με την μορφή:

$$\hat{\kappa} = \sum_{j=0}^{n} W_j \hat{\kappa}_j$$

με τις τυχαίες μεταβλητές (ή σταθερές) να ικανοποιούν τη σχέση

$$\sum_{j=0}^{n} W_j = 1$$

1.3.7 Μέθοδος Additive

Η μέθοδος «additive» βασίζεται στην υπόθεση ότι υπάρχουν οι γνωστές παράμετροι $\pi_0, \pi_1, \ldots, \pi_n \in (0,\infty)$ και οι άγνωστες παράμετροι $\zeta_0, \zeta_1, \ldots, \zeta_n$ έτσι ώστε να ισχύει η ταυτότητα

$$E\left[Z_{i,k}\right] = \zeta_k \pi_i$$

για όλα τα i, k ϵ {0, 1, …, n}.

Αν οι παράμετροι π_0, $\pi_1,\ldots$, π_n ερμηνεύονται ως ασφάλιστρα ή άλλες μονάδες μέτρησης των ετών ατυχήματος, τότε η υπόθεση σημαίνει ότι, για κάθε έτος ανάπτυξης k, οι αναμενόμενες στοιχειώδεις συχνότητες ζημιών

$$\zeta_{i,k} = E\left[\frac{Z_{i,k}}{\pi_i}\right]$$

είναι όμοιες για όλα τα χρόνια ατυχημάτων. Έστω

$$\alpha_i = \pi_i \sum_{k=0}^{n} \zeta_k \quad \text{και} \quad \gamma_k = \frac{\sum_{l=0}^{k} \zeta_l}{\sum_{l=0}^{n} \zeta_l}$$

παρατηρούμε

$$E\left[S_{i,k}\right] = \gamma_k \alpha_i$$

για όλα τα i, k ϵ {0, 1, …, n}έτσι ώστε $\alpha_i = E\left[S_{i,n}\right]$ και οι παράμετροι γ_0, $\gamma_1,\ldots$, γ_n σχηματίζουν ένα πρότυπο ανάπτυξης για αθροιστικές ποσοστώσεις.

Οι «additive» εκτιμητές των πρόσθετων ζημιών $Z_{i,k}$ με i+k≥n να ορίζονται ως

$$\widehat{Z}_{i,k}^{AD} = \widehat{\zeta}_k^{AD} \pi_i$$

και οι «additive» εκτιμητές των αθροιστικών ζημιών $S_{i,k}$ με i+k≥n να ορίζονται ως

$$\widehat{S}_{i,k}^{AD} = S_{i,n-i} + \sum_{l=n-i+1}^{k} \widehat{Z}_{i,l}^{AD}$$

όπου

$$\widehat{\zeta}_k^{AD} = \frac{\sum_{j=0}^{n-k} Z_{i,j}}{\sum_{j=0}^{n-k} \pi_j}$$

είναι ο «additive» στοιχειώδης συντελεστής ζημιών το έτος ανάπτυξης k.

Παράδειγμα F. Θεωρούμε το ακόλουθο τριγωνικό σχεδιάγραμμα για τις αθροιστικές ζημιές ο οποίος συμπληρώνεται από τα ασφάλιστρα:

		Έτος εξέλιξης k					
Έτος συμβάντος i	π_i	**0**	**1**	**2**	**3**	**4**	**5**
0	4025	1001	1855	2423	2988	3335	3483
1	4456	1113	2103	2774	3422	3844	
2	5315	1265	2433	3233	3977		
3	5986	1490	2873	3880			
4	6939	1725	3261				
5	8158	1889					

Έτσι παρατηρούμε το ακόλουθο τριγωνικό σχεδιάγραμμα των πρόσθετων ζημιών που συμπληρώνεται από τους «additive» στοιχειώδεις συντελεστές ζημιών:

		Έτος εξέλιξης k					
Έτος συμβάντος i	π_i	**0**	**1**	**2**	**3**	**4**	**5**
0	4025	1001	854	568	565	347	148
1	4456	1113	990	671	648	422	
2	5315	1265	1168	800	744		
3	5986	1490	1383	1007			
4	6939	1725	1536				
5	8158	1889					
$\hat{\zeta}_k$		0.243	0.222	0.154	0.142	0.091	0.037

Υπολογίζοντας τώρα τους «additive» εκτιμητές των μη παρατηρήσιμων πρόσθετων ζημιών, το τριγωνικό σχεδιάγραμμα των πρόσθετων ζημιών συμπληρώνεται ως εξής:

		Έτος εξέλιξης k					
Έτος συμβάντος i	π_i	**0**	**1**	**2**	**3**	**4**	**5**
0	4025	1001	854	568	565	347	148
1	4456	1113	990	671	648	422	*165*
2	5315	1265	1168	800	744	*484*	*197*
3	5986	1490	1383	1007	*850*	*545*	*221*
4	6939	1725	1536	*1069*	*985*	*631*	*257*
5	8158	1889	*1811*	*1256*	*1158*	*742*	*302*
$\hat{\zeta}_k$		0.243	0.222	0.154	0.142	0.091	0.037

Αναλόγως, το τριγωνικό σχεδιάγραμμα των αθροιστικών ζημιών συμπληρώνεται ακολούθως:

		Έτος εξέλιξης k					
Έτος συμβάντος i	α_i	**0**	**1**	**2**	**3**	**4**	**5**
0	4025	1001	1855	2423	2988	3335	3483
1	4456	1113	2103	2774	3422	3844	*4009*
2	5315	1265	2433	3233	3977	*4461*	*4658*
3	5986	1490	2873	3880	*4730*	*5275*	*5496*
4	6939	1725	3261	*4330*	*5315*	*5946*	*6203*
5	8158	1889	*3700*	*4956*	*6114*	*6856*	*7158*

Έστω

$$\hat{\gamma}_k^{AD} = \frac{\sum_{l=0}^{k} \hat{\zeta}_l^{AD}}{\sum_{l=0}^{n} \hat{\zeta}_l^{AD}}$$

και

$$\hat{\alpha}_i^{AD} = \pi_i \sum_{l=0}^{n} \hat{\zeta}_l^{AD}$$

οι «additive» εκτιμητές των μη παρατηρήσιμων αθροιστικών ζημιών μπορούν να γραφτούν ως:

$$\hat{S}_{i,k}^{AD} = S_{i,n-i} + (\hat{\gamma}_k^{AD} - \hat{\gamma}_{n-i}^{AD})\hat{\alpha}_i^{AD}$$

Αυτό δείχνει ότι οι «additive» εκτιμητές των αθροιστικών ζημιών δεν είναι τίποτα άλλο από τους εκτιμητές «Bornhuetter-Ferguson » σε σχέση με τις αθροιστικές «additive» ποσοστώσεις $\hat{\gamma}_k^{AD}$ και τους προγενέστερους εκτιμητές $\hat{\alpha}_i^{AD}$ των αναμενόμενων αθροιστικών ζημιών. Με άλλα λόγια, η μέθοδος «additive» είναι μια ειδική περίπτωση της «Bornhuetter-Ferguson» μεθόδου με προγενέστερους εκτιμητές των αθροιστικών ποσοστώσεων και των αναμενόμενων τελευταίων αθροιστικών ζημιών, οι οποίοι βασίζονται και σε εσωτερικές και σε εξωτερικές πληροφορίες.

Οι αναμενόμενοι συσσωρευτικοί συντελεστές ζημιών

$$\kappa_i = E\left[\frac{S_{i,n}}{\pi_i}\right]$$

ικανοποιούν

$$\kappa_i = \sum_{l=0}^{n} \zeta_{i.l}$$

Αφού οι αναμενόμενοι στοιχειώδεις συντελεστές ζημιών είναι όμοιοι για όλα τα έτη ατυχημάτων, έπεται επίσης ότι οι αναμενόμενοι συσσωρευτικοί συντελεστές ζημιών είναι όμοιοι για όλα τα έτη ατυχημάτων. Συνεπώς, ο «additive» συντελεστής ζημιών

$$\hat{\kappa}^{AD} = \sum_{l=0}^{n} \hat{\zeta}_l^{AD}$$

μπορεί να ερμηνευτεί ως ένας εκτιμητής του αναμενόμενου τελευταίου συσσωρευτικού συντελεστή ζημιών

$$\kappa = \sum_{l=0}^{n} \zeta_l$$

κοινό για όλα τα έτη ατυχημάτων. Επιπλέον, οι προγενέστεροι εκτιμητές $\hat{\alpha}_i^{AD}$ μπορούν να γραφτούν ως

$$\hat{\alpha}_i^{AD} = \pi_i \hat{\kappa}^{AD}$$

και μπορεί να δειχτεί ότι

$$\hat{\kappa}^{AD} = \frac{\sum_{j=0}^{n} S_{i,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j}^{AD} \pi_j}$$

Αυτό δείχνει ότι οι «additive» εκτιμητές των μη παρατηρήσιμων αθροιστικών ζημιών δεν είναι τίποτα άλλο από τους Cape-Cod εκτιμητές σε σχέση με τις συνολικές «additive» ποσοστώσεις $\hat{\gamma}_k^{AD}$. Με άλλα λόγια, η «additive» μέθοδος είναι μια ειδική περίπτωση της Cape-Cod μεθόδου με προγενέστερους εκτιμητές των αθροιστικών ποσοστώσεων, οι οποίοι βασίζονται και σε εσωτερικές και σε εξωτερικές πληροφορίες.

Η παρατήρηση ότι η «additive» μέθοδος είναι μια ειδική περίπτωση της Cape-Cod μεθόδου οφείλεται στο Zocher (2005).

1.3.8 Παρατηρήσεις

Ο ακόλουθος πίνακας συγκρίνει τις διαφορετικές μεθόδους της αποθεματοποίησης ζημιών σε σχέση με τις επιλογές των προγενέστερων εκτιμητών των αναμενόμενων τελευταίων αθροιστικών ζημιών α_i και των αθροιστικών ποσοστώσεων γ_k:

Αναμενόμενες Τελευταίες Αθροιστικές Ζημιές	**Αθροιστικές Ποσοστώσεις**		
	Arbitrary	$\hat{\gamma}_k^{CL}$	$\hat{\gamma}_k^{AD}$
Arbitrary	Μέθοδος Bornhuetter-Ferguson		
$\hat{S}_{i,n}^{AD}$	Μέθοδος Loss-Development	Μέθοδος Chain Ladder	
$\pi_i \hat{\kappa}^{CC}$	Μέθοδος Cape-Cod		Μέθοδος Additive

Σημειώστε ότι οι προγενέστεροι εκτιμητές $\hat{S}_{i,n}^{LD}$ και $\pi_i \hat{\kappa}^{CC}$ εξαρτώνται από την επιλογή των προγενέστερων εκτιμητών $\gamma_0, \gamma_1, \ldots, \gamma_n$.

Φυσικά, οι τέσσερις άλλοι συνδυασμοί στους οποίους δεν έχει δοθεί ένα όνομα στη λογοτεχνία θα μπορούσαν να χρησιμοποιηθούν επίσης, και ακόμη άλλες επιλογές προγενέστερων εκτιμητών μπορούν των αναμενόμενων τελευταίων ζημιών και των αθροιστικών ποσοστώσεων μπορούν να θεωρηθούν.

Η συζήτηση του παροντικού τμήματος και ειδικά, ο παραπάνω πίνακας δείχνει ότι η μέθοδος Bornhuetter-Ferguson παρέχει ένα γενικό αξίωμα κάτω από το οποίο κάποιες μέθοδοι της αποθεματοποίησης ζημιών μπορούν να υπαχθούν. Η εστίαση

- στους προγενέστερους εκτιμητές των αναμενόμενων τελευταίων αθροιστικών ζημιών

και

- στους προγενέστερους εκτιμητές των αθροιστικών ποσοστώσεων

παρέχει μια μεγάλη μεταβλητότητα στις μεθόδους αποθεματοποίησης. Ο παραπάνω πίνακας περιέχει ενδιαφέρουσες ειδικές περιπτώσεις, αλλά θα μπορούσε σίγουρα να μεγεθυνθεί. Επιπλέον,

- κάποιοι κυρτοί συνδυασμοί των προγενέστερων εκτιμητών των αναμενόμενων τελευταίων αθροιστικών ζημιών αποφέρει νέους προγενέστερους εκτιμητές των αναμενόμενων τελευταίων αθροιστικών ζημιών, και

- κάποιοι κυρτοί συνδυασμοί των προγενέστερων εκτιμητών του προτύπου ανάπτυξης για τις αθροιστικές ποσοστώσεις αποφέρει νέους προγενέστερους εκτιμητές του προτύπου ανάπτυξης.

Αυτό το σημείο γίνεται ακριβές με το ακόλουθο παράδειγμα:

Παράδειγμα G. Έστω είναι οι προγενέστεροι εκτιμητές των α_0, α_1, …, α_n και έστω $\hat{\gamma}_0, \hat{\gamma}_1, ..., \hat{\gamma}_n$ ότι είναι οι προγενέστεροι εκτιμητές των γ_0, γ_1, …, γ_n έτσι ώστε κάθε ένας από αυτούς τους εκτιμητές να βασίζεται πλήρως σε εξωτερικές πληροφορίες. Τότε οι προγενέστεροι εκτιμητές

$$\tilde{\alpha}_i = a_1 \hat{\alpha}_i + a_2 \hat{S}_{i,n}^{LD} + a_3 (\pi_i \hat{\kappa}^{CC})$$

με $\alpha_1+\alpha_2+\alpha_3=1$ και
$$\tilde{\alpha}_i = a_1 \hat{\alpha}_i + a_2 \hat{S}_{i,n}^{LD} + a_3 (\pi_i \hat{\kappa}^{CC})$$
$$\tilde{\gamma}_k = b_1 \hat{\gamma}_k + b_2 \hat{\gamma}_k^{CL} + b_3 \hat{\gamma}_k^{AD}$$

$$\tilde{\gamma}_k = b_1 \hat{\gamma}_k + b_2 \hat{\gamma}_k^{CL} + b_3 \hat{\gamma}_k^{AD}$$

με $b_1+b_2+b_3=1$ να είναι οι προγενέστεροι εκτιμητές των α_0, α_1, …, α_n και γ_0, γ_1, …, γ_n αντίστοιχα, οι οποίοι μέσα από τα βάρη $\alpha_1,\alpha_2,\alpha_3$ και b_1,b_2,b_3 εκφράζουν την αξιοπιστία που έχει αποδοθεί στους προγενέστερους εκτιμητές $\hat{\alpha}_i$, $\hat{S}_{i,n}^{LD}$, $\pi_i \hat{\kappa}^{CC}$ και $\hat{\gamma}_k$, $\hat{\gamma}_k^{CL}$, $\hat{\gamma}_k^{AD}$ αντίστοιχα.

ΜΕΡΟΣ 2^{O}

Στοχαστικές Μέθοδοι Αποθεματοποίησης στις Γενικές Ασφαλίσεις

Τα τελευταία χρόνια έχει παρατηρηθεί αυξημένο ενδιαφέρον για τα στοχαστικά μοντέλα αποθεματοποίησης ζημιών στο χώρο των ασφαλειών, παρόλα αυτά χρησιμοποιούνται από έναν περιορισμένο αριθμό χρηστών. Κάποιοι λόγοι είναι η έλλειψη κατανόησης των μεθόδων αυτών, η έλλειψη κατάλληλου λογισμικού κ.ά. Ο βασικός όμως λόγος είναι ότι τα στοχαστικά μοντέλα είχαν σχεδιαστεί για να παράγουν ακριβώς τις ίδιες εκτιμήσεις για τα αποθέματα με την παραδοσιακή ντετερμινιστική τεχνική της «chain ladder». Αναμφισβήτητα, λοιπόν γεννάται το ερώτημα «γιατί να χρησιμοποιηθεί μια πολύπλοκη στοχαστική διαδικασία για την εκτίμηση των αποθεμάτων, όταν η απλή ντετερμινιστική διαδικασία επαρκεί»; Η απάντηση είναι ότι πέρα από την εκτίμηση του αποθέματος υπάρχουν και άλλα θέματα του μοντέλου που είναι σημαντικά, όπως οι εκτιμήσεις της μεταβλητότητας των εκτιμήσεων των παραμέτρων, οι υποκείμενες κατανεμημένες υποθέσεις του μοντέλου που ταιριάζουν και μία εκτίμηση της πιστότητας του μοντέλου. Είναι επίσης χρήσιμο να γνωρίζουμε που τα δεδομένα αποκλίνουν από το μοντέλο που χρησιμοποιείται, και να έχουμε κάποια πλαίσια μέσα στα οποία άλλα μοντέλα μπορούν να χρησιμοποιηθούν και να συγκριθούν.

Η επιλογή του στοχαστικού μοντέλου ή της μεθόδου πρόβλεψης είναι επιλογή του αναλογιστή και το οποίο ίσως έχει μια σημαντική επίδραση στο αποτέλεσμα. Στο κεφάλαιο αυτό θα παρουσιάσουμε τις πιο σημαντικές στοχαστικές μεθόδους. Οι στοχαστικές μέθοδοι αποθεματοποίησης ζημιών διακρίνονται στους ακόλουθους τύπους:

- Αθροιστικές μέθοδοι (Cumulative methods): Αυτές οι μέθοδοι παράγουν εκτιμήσεις σχετικές με τον ντετερμινιστικό αλγόριθμο της μεθόδου «chain ladder». Επίσης αυτές οι μέθοδοι μετρούν την μεταβλητότητα των αναλογιών σύνδεσης. Γενικά απαιτούν από το χρήστη να επιλέξει τον τύπο της κατανομής (π.χ. κανονική ή λογαριθμοκανονική).
  Π.χ. μέθοδος Mack

- Μέθοδοι Προσομοίωσης (Simulation methods): Στις μεθόδους αυτές η υποκείμενη υπόθεση είναι ότι τα δεδομένα της προσομοίωσης έχουν τις ίδιες στατιστικές ιδιότητες με τα παρατηρούμενα δεδομένα. Γενικά είναι ευκολότερη η χρήση μεθόδων προσομοίωσης από την άντληση ενός μαθηματικού τύπου.

  Π.χ. Bootstrapping μέθοδος

- Μπεϋζιανές μέθοδοι (Bayesian methods): Οι άλλες μέθοδοι αντλούν παραμέτρους άμεσα από τα ιστορικά δεδομένα, ενώ οι Μπεϋζιανές μέθοδοι χρησιμοποιούν τα δεδομένα ως ένα δεύτερο βήμα για να βελτιώσουν τις αρχικές προσδοκίες για τις παραμέτρους του χρήστη. Ο χρήστης κάνει υποθέσεις σύμφωνα με τις αρχικές κατανομές των παραμέτρων.

  Π.χ. το μοντέλο του Verall (2000)

- Βαθμιαίες προσεγγίσεις (Incremental approaches): Οι μέθοδοι αυτές βασίζονται στις πληρωμές πρόσθετων ζημιών. Επίσης σε αυτές τις μεθόδους προσαρμόζουν την καμπύλη σε διαστάσεις ετών ανάπτυξης ή ατυχημάτων ή ημερολογιακών.

  Π.χ. Το γενικό γραμμικό μοντέλο

Παρακάτω παρουσιάζονται τα σημαντικότερα από τα στοχαστικά μοντέλα που έχουν προταθεί, η μεταξύ τους σχέση, πως αυτές οι μέθοδοι μπορούν να εφαρμοστούν στην πράξη και ποια είναι τα χαρακτηριστικά τους.

2.1 Μέθοδος Mack

Ο Mack (1993) ανέπτυξε μία μέθοδος χωρίς τη χρήση κατανομής για να εκτιμήσει την πρόβλεψη του λάθους των εκτιμητών Chain Ladder αποθεμάτων. Για την αποθεματοποίηση ζημιών μια ασφαλιστική εταιρεία συνήθως διαιρεί το χαρτοφυλάκιό της σε υποχαρτοφυλάκια έτσι ώστε η ανάπτυξη συμπεριφοράς του κάθε υποχαρτοφυλακίου να θεωρείται ομογενής. Τότε, για κάθε υποχαρτοφυλάκιο μπορεί να εφαρμοστεί η μέθοδος Chain Ladder για τον υπολογισμό των αποθεμάτων και του λάθους πρόβλεψης.

Αλλά τελικά χρειάζεται να υπολογιστούν τα αποθέματα όλου του χαρτοφυλακίου και το λάθος πρόβλεψής τους. Οι εκτιμήσεις των αποθεμάτων ζημιών κάθε υποχαρτοφυλακίου μπορεί να προστεθούν και να δώσουν το τελικό απόθεμα ζημιών όλου του χαρτοφυλακίου, αυτό μπορεί να γίνει και στην περίπτωση των διασπορών των προβλέψεων υπό την προϋπόθεση ότι τα υποχαρτοφυλάκια είναι ανεξάρτητα. Στην πραγματικότητα όμως η ανεξαρτησία αυτή δεν επαληθεύεται, οπότε δεν μπορεί να εφαρμοστεί η μέθοδος Chain Ladder για την πρόβλεψη του λάθους πρόβλεψης του αποθέματος.

Η πρόβλεψη του λάθους της πρόβλεψης

<u>Αρχικά εκτιμούμε το λάθος της πρόβλεψης για το ποσό των τελευταίων ζημιών ενός έτους ατυχήματος.</u>

Έστω $C_{ik}>0$ είναι το σύνολο των ζημιών του έτους ατυχήματος i, $1\le i \le n$, μετά από k χρόνια ανάπτυξης, $1\le k \le n$,για ένα συγκεκριμένο υποχαρτοφυλάκιο. Τα ποσά C_{ik} με $i+k\le n+1$ είναι παρατηρήσιμα και εμείς ενδιαφερόμαστε να προβλέψουμε τα ποσά C_{in} για i=2, 3, …, n. Η μέθοδος Chain Ladder το κάνει αυτό με τον επαναληπτικό τύπο:

$$\widehat{C}_{ik}=\widehat{C}_{i,k-1} * \widehat{f}_k$$

Με αρχική τιμή $\widehat{C}_{i,n+1-i}=C_{i,n+1-i}$ και τον παράγοντα:

$$\widehat{f}_k = \frac{\sum_{i=1}^{n+1-k} C_{ik}}{C_{<,k-1}} = \sum_{i=1}^{n+1-k} \frac{C_{i,k-1}}{C_{<,k-1}} * F_{ik}$$

ο οποίος είναι ο σταθμισμένος μέσος των ατομικών παραγόντων ανάπτυξης

$$F_{ik} := \frac{C_{i,k}}{C_{i,k-1}} \text{ με } C_{<,k-1} := \sum_{i=1}^{n+1-k} C_{i,k-1}$$

Θεωρούμε : "T_k":όλες τις μεταβλητές $\{C_{ij}| 1\leq i \leq n, 1\leq j \leq k, i+j\leq n+1\}$ με το χρόνο ανάπτυξης k δεδομένο.

"T_n":όταν θεωρείται όλο το τρίγωνο δεδομένο

"T_{ik}":όταν θεωρούνται οι μεταβλητές $\{C_{ij}| 1\leq j \leq k\}$ δεδομένες

Σύμφωνα με τις στοχαστικές υποθέσεις ισχύει ότι :

(1) $E[F_{ik}| T_{i,k-1}) = f_k$

Και

(2) $Var[F_{ik}| T_{i,k-1}) = \dfrac{\sigma_k^2}{C_{i,k-1}}$

Για όλα τα $1\leq i\leq n$ και $2\leq k\leq n$ όπου f_k και σ_k^2 είναι άγνωστοι παράμετροι, τα $\widehat{f}_k$ και $\widehat{C}_{i,n+1-i}$ είναι αμερόληπτα αν τα χρόνια ατυχημάτων είναι ανεξάρτητα. Οι υποθέσεις (1) και (2) μαζί με την υπόθεση της ανεξαρτησίας των ετών ατυχημάτων αποτελούν τη βάση για όσα θα αναφέρουμε παρακάτω.

Το λάθος πρόβλεψης mse($\widehat{C}_{i,n}$) για το τελευταίο ποσό ζημιών ενός έτους ατυχήματος ορίζεται ως:

$$mse(\widehat{C}_{i,n}) := E((C_{in} - \widehat{C}_{i,n})^2| T_n)$$

επειδή για λόγους αποθεματοποίησης μόνο η μελλοντική μεταβλητότητα δεδομένων των παρατηρήσιμων δεδομένων μας ενδιαφέρει, το παραπάνω το γράφουμε ως:

$$mse(\widehat{C}_{i,n}) := Var(C_{in}| T_{n+1-i}) + (E(\widehat{C}_{i,n} | T_{n+1-i}) - \widehat{C}_{i,n})^2$$

όπου για λόγους εκτίμησης προσεγγίζεται από:

$$mse(\widehat{C}_{i,n}) \approx Var(C_{in}| T_{n+1-i}) + Var(\widehat{C}_{i,n} | T_{n+1-i})$$

όπου Var(C_{in}| T_{n+1-i}) ονομάζεται το τυχαίο λάθος και Var($\widehat{C}_{i,n}$| T_{n+1-i}) το λάθος εκτίμησης . Για λόγους απλοποίησης αν i+k>n+1 οι προβλέψεις E(C_{in}), E($\widehat{C}_{i,n}$) και οι διασπορές Var(C_{in}), Var ($\widehat{C}_{i,n}$) είναι το ίδιο με E($C_{i,n} | T_{n+1-i}$), E($\widehat{C}_{i,n} | T_{n+1-i}$) και τις διασπορές Var($C_{i,n} | T_{n+1-i}$), Var ($\widehat{C}_{i,n} | T_{n+1-i}$) .

Στη συνέχεια συνάγουμε επαναλήψεις για το τυχαίο λάθος και το συστηματικό λάθος.

Οπότε έχουμε τις σχέσεις:

E(C_{ik}| $T_{i,k-1}$)= $C_{i,k-1}$ f_k

Var(C_{ik}| $T_{i,k-1}$)= $C_{i,k-1}$ σ_k^2

Τότε για i+k>n+1

Var(C_{ik})= E(Var(C_{ik}| $T_{i,k-1}$))+ Var(E(C_{ik}| $T_{i,k-1}$))= E($C_{i,k-1}$) σ_k^2 + Var($C_{i,k-1}$) f_k^2

Αυτό παράγει για τον εκτιμητή $\widehat{Var}(C_{in})$ του τυχαίου λάθους Var(C_{in}) του ποσού των τελευταίων ζημιών την επανάληψη:

$\widehat{Var}$ (C_{ik})=$\widehat{Var}$ ($C_{i,k-1}$) $\widehat{f}_k^2$+ $\widehat{C}_{i,k-1} \widehat{\sigma}_k^2$

με την αρχική τιμή :

$\widehat{Var}$ ($C_{i,n+1-i}$)=0

καθώς το $C_{i,n+1-i}$ είναι γνωστό. Ένας αμερόληπτος εκτιμητής του $\widehat{\sigma}_k^2$ δίνεται από τον ακόλουθο τύπο:

$$\widehat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} C_{i,k-1} \left(F_{ik} - \widehat{f}_k \right)^2$$

Παρομοίως,

$\widehat{C}_{ik} = \widehat{C}_{i,k-1} * \widehat{f}_k$

Οπότε παράγεται:

$$Var(\widehat{C}_{ik}) = E\left(Var\left(\widehat{C}_{i,k-1}\widehat{f}_k \mid T_{k-1}\right)\right) + Var\left(E\left(\widehat{C}_{i,k-1}\widehat{f}_k \mid T_{k-1}\right)\right) =$$

$$= E\left(\widehat{C}^2_{i,k-1} Var(\widehat{f}_k \mid T_{k-1})\right) + Var\left(\widehat{C}_{i,k-1}\right) f_k^2$$

Οπότε ο επαναληπτικός τύπος για τον εκτιμητή $\widehat{Var}\left(C_{in}\right)$ του λάθους εκτίμησης $Var(\widehat{C}_{in})$ των τελευταίων ζημιών C_{in} μπορούν να συναχθούν από τον τύπο:

$$\widehat{Var}\left(\widehat{C}_{i,k}\right) = \widehat{Var}\left(\widehat{C}_{i,k-1}\right) f_k^2 + \widehat{C}^2_{i,k-1} \frac{\widehat{\sigma}_k^2}{C_{<,k-1}}$$

Επειδή

$$Var(\widehat{f}_k \mid T_{k-1}) = \frac{\widehat{\sigma}_k^2}{C_{<,k-1}}$$

Εδώ η αρχική τιμή είναι :

$$\widehat{Var}\left(\widehat{C}_{i,n+1-i}\right) = 0$$

επειδή $\widehat{C}_{i,n+1-i}$ είδη παρατηρήσιμο. Οπότε ο επαναληπτικός τύπος για την εκτίμηση όλου του λάθους πρόβλεψης είναι:

Εκτιμούμε το λάθος της πρόβλεψης του συνολικού ποσού των τελευταίων ζημιών Οι ετήσιες αναφορές μιας ασφαλιστικής εταιρείας κάνουν συνήθως εκτιμήσεις μόνο για αποθέματα όλων των ετών ατυχημάτων μαζί .Για να εκτιμήσουμε ένα διάστημα για αυτού του συνολικού ποσού, πρέπει να εκτιμήσουμε το λάθος εκτίμησης και το λάθος πρόβλεψης για όλα τα έτη ατυχημάτων μαζί.

Το C_{1n} είναι ήδη γνωστό και δεν απαιτείται καμία εκτίμηση. Επομένως το πρώτο έτος ατυχήματος δεν προσθέτει κάτι στο τυχαίο λάθος και στο λάθος εκτίμησης. Παίρνοντας αυτό ως δεδομένο, η πρόβλεψη του λάθους πρόβλεψης $mse\left(\sum_{i=2}^{n}\widehat{C}_{in}\right)$ για όλα τα έτη ατυχήματος ορίζεται ως :

$$mse\left(\sum_{i=2}^{n}\widehat{C}_{in}\right) := E\left[\left(\sum_{i=2}^{n}(C_{in}-\widehat{C}_{in})\right)^2 \mid T_n\right]$$

Έτσι έχουμε :

$$mse\left(\sum_{i=2}^{n}\widehat{C}_{in}\right) = Var\left(\sum_{i=2}^{n}C_{in} \mid T_n\right) + \left(\sum_{i=2}^{n}E\left(\widehat{C}_{in} \mid T_{n+1-i}\right)-\widehat{C}_{in}\right)^2$$

$$= Var\left(\sum_{i=2}^{n}C_{in} \mid T_n\right) + \sum_{i=2}^{n}\left(E\left(\widehat{C}_{in} \mid T_{n+1-i}\right)-\widehat{C}_{in}\right)^2$$

$$+2\sum_{2\le i<j\le n}\left(E\left(\widehat{C}_{in} \mid T_{n+1-i}\right)-\widehat{C}_{in}\right)\left(E\left(\widehat{C}_{jn} \mid T_{n+1-j}\right)-\widehat{C}_{jn}\right)$$

$$\approx Var\left(\sum_{i=2}^{n}C_{in} \mid T_n\right) + \sum_{i=2}^{n}Var\left(\widehat{C}_{in} \mid T_{n+1-i}\right)$$

$$+2\sum_{2\le i<j\le n}Cov\left(\widehat{C}_{in},\widehat{C}_{jm} \mid T_{n+1-i}\right)$$

Η αντίστοιχη εκτίμηση είναι :

$$\widehat{mse}\left(\sum_{i=2}^{n}\widehat{C}_{in}\right) = \widehat{mse}\left(\sum_{i=v+3-k}^{n}\widehat{C}_{i,k-1}\right)\widehat{f}_k^{\,2} + \left(\widehat{C}_{\geq,k-1}\right)^2\left(\frac{\widehat{\sigma}_k^{\,2}}{\widehat{C}_{\geq,k-1}} + \frac{\widehat{\sigma}_k^{\,2}}{C_{<,k-1}}\right)$$

Η επανάληψη αρχίζει με k=2.

2.2 Μέθοδος Bootstrap

Όπου το λάθος της εκτίμησης είναι δύσκολο ή αδύνατον να υπολογιστεί αναλυτικά, είναι σύνηθες να υιοθετηθεί η μέθοδος «bootstrap». Στην αποθεματοποίηση ζημιών εμείς ενδιαφερόμαστε να προβλέψουμε το λάθος του συνόλου των τυχαίων μεταβλητών και η μέθοδος «bootstrap» είναι από τις πιο κατάλληλες μεθόδους για αυτή τη διαδικασία. Σε προβλήματα τύπου παλινδρόμησης, είναι σύνηθες να χρησιμοποιείται η μέθοδος «bootstrap» στα υπόλοιπα και όχι στα δεδομένα. Όμως είναι σημαντικό να χρησιμοποιήσουμε ένα κατάλληλο ορισμό για τα υπόλοιπα του προβλήματος που ασχολούμαστε. Στα γενικευμένα γραμμικά μοντέλα τα υπόλοιπα απαιτείται να έχουν προσεγγιστικά τις ιδιότητες της κανονικής κατανομής. Τα πιο συνήθη υπόλοιπα είναι τα Deviance και τα Pearson. Μια Τρίτη κατηγορία υπολοίπων λιγότερο συνήθη είναι τα λεγόμενα Anscombe υπόλοιπα. Παρακάτω παραθέτουμε τα είδη υπολοίπων που αναφέραμε, τα οποία είναι κατάλληλα για το Γενικευμένο Γραμμικό μοντέλο Poisson και είναι τα:

- Μη κλιμακωτό Deviance υπόλοιπο

$$r_D = sign\left(C - m\right)\sqrt{2(C\log(C/m) - C + m}$$

- Μη κλιμακωτό Pearson υπόλοιπο

$$r_P = \frac{C - m}{\sqrt{m}}$$

- Μη κλιμακωτά Anscombe υπόλοιπα

$$r_A = \frac{\frac{3}{2}(C^{2/3} - m^{2/3})}{m^{1/6}}$$

Η διαδικασία «bootstrap» περιλαμβάνει την δειγματοληψία, με επανατοποθέτηση, από τα υπόλοιπα. Ένα «bootstrap» δείγμα δεδομένων δημιουργείται αντιστρέφοντας τον τύπο για τα υπόλοιπα χρησιμοποιώντας τα υπόλοιπα που έχουν παρθεί ως δείγμα. Δεδομένων των r και m, δεν μπορεί να λυθεί αναλυτικά η

$$r_D = sign(C-m)\sqrt{2(C\log(C/m)-C+m}$$

για τις παρατηρούμενες πρόσθετες ζημιές, C, κάνοντας τα Deviance υπόλοιπα λιγότερο κατάλληλα για «bootstrapping». Όμως είναι εύκολο να λύσουμε την εξίσωση

$$r_P = \frac{C-m}{\sqrt{m}}$$

ως προς C. Δεδομένου του υπολοίπου Pearson r_P^* μαζί με την διαφορά της ονομαστικής αξίας από την τρέχουσα τιμή m, οι αντίστοιχες πρόσθετες ζημιές C^*, δίνονται από τον τύπο:

$$C^* = r_P^*\sqrt{m}+m$$

Είναι επίσης πιθανό να λύσουμε τα υπόλοιπα Anscombe ως προς C, αλλά δεν αναφερόμαστε πιο εκτενώς εδώ γιατί χρησιμοποιούνται πιο σπάνια.

Έχοντας αποκτήσει το δείγμα «bootstrap» το μοντέλο μετατρέπεται και τα στατιστικά στοιχεία που μας ενδιαφέρουν. Η διαδικασία επαναλαμβάνεται πολλές φορές, κάθε φορά παρέχοντας ένα νέο «bootstrap» δείγμα και τα στατιστικά στοιχεία που μας ενδιαφέρουν. Το τυπικό λάθος της «bootstrap» μεθόδου είναι η τυπική απόκλιση της «bootstrap» στατιστικής.

Στην στοχαστική αποθεματοποίηση ζημιών, η δειγματοληψία των υπολοίπων (με επανατοποθέτηση) δίνει ένα νέο τρίγωνο από πληρωμές ζημιών. Οφείλουμε να ταιριάξουμε το γενικευμένο γραμμικό μοντέλο «overdispersed» Poisson στο δείγμα «bootstrap» για να εκτιμήσουμε τα «bootstrap» αποθέματα. Όμως μπορούμε να πάρουμε όμοια αποτελέσματα χρησιμοποιώντας την τυπική μεθοδολογία της «chain ladder».

Η χρησιμότητα της μεθόδου «bootstrap» σε αυτό το σημείο γίνεται εμφανής: δεν απαιτείται να σχεδιάσουμε κάποιο λογισμικό , ένα λογιστικό φύλλο είναι αρκετό. Για να υπολογίσουμε το τυπικό λάθος των εκτιμήσεων των αποθεμάτων, είναι αναγκαίο να επαναληφθεί η διαδικασία πολλές φορές (έστω N), κάθε φορά δημιουργώντας ένα νέο «bootstrap» δείγμα και αποκτώντας υπολογισμούς για τα αποθέματα της «chain ladder». Τα τυπικά «bootstrap» λάθη είναι οι τυπικές αποκλίσεις N «bootstrap» εκτιμήσεων των αποθεμάτων. Η διαδικασία είναι γρήγορη, διαρκεί μόνο λίγα δευτερόλεπτα στην οθόνη ενός υπολογιστή.

Το «bootstrap» τυπικό λάθος είναι μια εκτίμηση της τετραγωνικής ρίζας της εκτιμώμενης διασποράς. Όμως αυτό δεν μπορεί να συγκριθεί άμεσα με το ισοδύναμο αναλυτικό το «bootstrap» τυπικό λάθος, γιατί δεν λαμβάνουμε υπόψη τον αριθμό των παραμέτρων που χρησιμοποιούνται στο μοντέλο: η «bootstrap» διαδικασία απλά χρησιμοποιεί τα υπόλοιπα χωρίς να λαμβάνεται υπόψη πως αυτά αποκτήθηκαν. Οι αναλυτικές εκτιμήσεις της εκτίμησης της διασποράς λαμβάνουν υπόψη τους τον αριθμό των εκτιμώμενων μεταβλητών αφού περιλαμβάνουν διασπορά και όρους συνδιακύμανσης που συνεπάγεται ότι περιλαμβάνουν την βαθμιαία παράμετρο φ στους υπολογισμούς τους. Η βαθμιαία παράμετρος εκτιμάται είτε με το μοντέλο της απόκλισης να χωρίζεται από τους βαθμούς της ελευθερίας, είτε η στατιστική Pearson X^2 να χωρίζεται από τους βαθμούς ελευθερίας, η επιλογή έχει μικρή διαφορά. Η απόκλιση και το Pearson X^2 υπολογίζονται από το άθροισμα των τετραγώνων των αντίστοιχων υπολοίπων. Ο αριθμός των βαθμών ελευθερίας υπολογίζονται ως το πλήθος των δεδομένων ελαττωμένα από τον αριθμό των παραμέτρων που τοποθετούνται στο μοντέλο. Επομένως η Deviance βαθμωτή παράμετρος δίνεται από τον τύπο:

$$\Phi_D = \frac{\sum r_D^2}{n-p}$$

Και η βαθμωτή παράμετρος Pearson δίνεται από τον τύπο:

$$\Phi_P = \frac{\sum r_P^2}{n-p}$$

όπου n είναι ο αριθμός των δεδομένων στο δείγμα, p είναι ο αριθμός των παραμέτρων που εκτιμώνται και την άθροιση να γίνεται γύρω από τον αριθμό (n) των υπολοίπων.

Για συνοχή, χρησιμοποιούμε την βαθμιαία παράμετρο Pearson στην αναλυτική εκτίμηση της διασποράς και τα υπόλοιπα Pearson στην διαδικασία «bootstrap». Η εκτίμηση «bootstrap»της διασποράς είναι ανάλογη με την αναλυτική εκτίμηση της διασποράς χωρίς να προσαρμόσουμε με τον αριθμό των παραμέτρων. Για να κάνουμε σύγκριση μεταξύ των διασπορών που υπολογίζονται από τις δύο μεθόδους πρέπει να κάνουμε μία προσαρμογή της εκτίμησης της διασποράς με τη μέθοδο «bootstrap» και να λάβουμε υπόψη τον αριθμό των παραμέτρων που χρησιμοποιούνται στο μοντέλο. Η κατάλληλη προσαρμογή είναι να πολλαπλασιάσουμε την εκτίμηση της διασποράς με τη μέθοδο «bootstrap» με n/(n-p).

Για να έχουμε το λάθος πρόβλεψης με τη μέθοδο «bootstrap», είναι αναγκαίο να προσθέσουμε μία εκτίμηση στην διασπορά της διαδικασίας, η οποία είναι απλά η βαθμιαία παράμετρος πολλαπλασιασμένη με τους εκτιμητές αποθεμάτων. Οι εκτιμητές αποθεμάτων δίνονται από την αρχική πρόβλεψη από την τεχνική «chain ladder», και η βαθμιαία παράμετρος υπολογίζεται από την άθροιση των τετραγώνων των υπολοίπων που χρησιμοποιήθηκαν στην «bootstrap» διαδικασία. Η διασπορά της διαδικασίας μπορεί να υπολογιστεί από ένα λογιστικό φύλλο. Η πρόβλεψη του λάθους «bootstrap» δίνεται από τον τύπο:

$$PE_{bs} = \sqrt{\Phi_P R + \frac{n}{n-p}(SE_{bs}(R))^2}$$

Όπου R είναι ένα έτος ατυχήματος ή το συνολικό απόθεμα, και $SE_{bs}(R)$ είναι το τυπικό λάθος των εκτιμώμενων αποθεμάτων όπως υπολογίστηκαν με τη μέθοδο «bootstrap» (England, 1999).

2.3 Μπεϋζιανές μέθοδοι (Bayesian methods)

Το πρόβλημα

Για να δηλώσουμε το πρόβλημα συμβολίζουμε με X_{ij} είτε το προσαυξημένο ποσό των ζημιών ή τον αριθμό των ζημιών που απορρέουν από το έτος προέλευσης ϊ και πληρώνονται στο έτος ανάπτυξης j και ας υποθέσουμε ότι είμαστε στο έτος n και ότι ξέρουμε όλες τις παρελθοντικές πληροφορίες, οι οποίες είναι X_{ij} για i= 1, 2, …,n και j=1, 2, …,n+1-i.

Έτος συμβάντος	Έτος εξέλιξης						
	1	**2**	**…**	**j**	**…**	**n-1**	**n**
0	X_{11}	X_{12}	…	X_{1j}	…	$X_{1,n-1}$	$X_{1,n}$
1	X_{21}	X_{22}	…	X_{2j}	…	$X_{2,n-1}$	
⋮	⋮	⋮	…	⋮	…		
I	X_{i1}	X_{i2}	…	$X_{i,n+1-i}$			
⋮	⋮	⋮					
n-1	$X_{n-1,1}$	$X_{n-1,2}$					
N	$X_{n,1}$						

Τα δεδομένα σχηματίζουν ένα τρίγωνο και αυτό κάνει πιο εύκολη την διαδικασία, όμως η μέθοδος μπορεί να εφαρμοστεί και σε άλλα σχήματα δεδομένων. Το πρόβλημα αποτελείται από τις προβλέψιμες παρατηρήσεις X_{ij} για i= 1, 2, …,n και j=n+ 2-i, n+3-i,…,n τα οποία αντιστοιχούν στο κάτω-δεξιά τμήμα του τριγώνου που αναπαρίσταται στον παραπάνω πίνακα. Ας συμβολίσουμε με

$$R_i = \sum_{j=n+2-i}^{n} X_{ij}$$

το άθροισμα των ζημιών για κάθε έτος προέλευσης i= 1, 2, …,n. Από την κατανομή του R_i το απαιτούμενο αποθεματικό που αντιστοιχεί σε κάθε έτος προέλευσης μπορεί να υπολογιστεί επιλέγοντας το μέτρο της κεντρικής τάσης ή το επιθυμητό ποσοστημόριο στην κατανομή των εκκρεμών ζημιών. Το συνολικό αποθεματικό μπορεί να υπολογιστεί από την κατανομή

$$R = \sum_{i=2}^{n} R_i$$

Σύμφωνα με τους England και Verall (2002) και τον Pinheiro et al (2003) οι περισσότερες από τις στοχαστικές διαδικασίες βασίζονται στη χρήση γενικευμένων γραμμικών μοντέλων, αυτό όμως δεν είναι αναγκαίο και στην περίπτωση του μοντέλου του Bayes. Τα γενικευμένα γραμμικά μοντέλα μπορούν να περιγραφούν ακολούθως. Αν Y_{ij} είναι ανεξάρτητες τυχαίες μεταβλητές για όλα τα i και j με συνάρτηση πυκνότητας πιθανότητας $f(y/\mu_{ij},\varphi)$ που ανήκει στην εκθετική οικογένεια, $\mu_{ij}= E[X_{ij}]$ και φ μια βαθμωτή παράμετρος. Ο μέσος μ_{ij} σχετίζεται με ένα σύνολο από παράγοντες μέσω της σχέσης $n_{ij}=g(\mu_{ij})$ με $n_{ij}=c+\alpha_i+\beta_j$. Συνήθως παίρνουμε $\alpha_1=\beta_1=0$. Για μοντέλα που βασίζονται στην κατανομή Poisson ή Γάμμα $X_{ij}= Y_{ij}$ και για μοντέλα που βασίζονται στην κανονική κατανομή είναι $Y_{ij} =\log X_{ij}$.

Σε άλλες προσεγγιστικές προσεγγίσεις (Mack, 1991) υποθέτουμε ότι οι πρόσθετες ζημιές ακολουθούν την κατανομή Γάμμα : $X_{ij} \sim G\left(a, \frac{a}{m_{ij}}\right)$. Σε όλες τις περιπτώσεις η βασική υπόθεση είναι ότι τα X_{ij} και τα Y_{ij} θεωρούνται ανεξάρτητα για όλα τα i και j. Αυτό στην πραγματικότητα όμως δεν ισχύει επειδή οι πρόσθετες ζημιές αυξάνονται ή μειώνονται κατά τα έτη ανάπτυξης. Με σκοπό να ξεπεράσουμε αυτόν τον περιορισμό της ανεξαρτησίας, ο Krener (2005) πρότεινε το ακόλουθο μοντέλο:

$$X_{ij} = a_j X_{i,j-1} + \in_{ij}$$

για όλα τα i, j =1, …, n, όπου α_j είναι ο άγνωστος αυξητικός παράγοντας και $\in_{ij}$ είναι ανεξάρτητα λάθη έτσι ώστε $E\left(\in_{ij} | X_{i,j-1}\right) = 0$ και $Var\left(\in_{ij} | X_{i,j-1}\right) = \nu_j X_{i,j-1}$. Αυτό το μοντέλο είναι το πιο απλό και ευρέως χρησιμοποιούμενο όταν υπάρχει συσχέτιση. Σε αυτό το μοντέλο που αναπτύσσουμε υπάρχει συσχέτιση μεταξύ των χρόνων ανάπτυξης στα στις πρόσθετες ζημιές (Alba et al., 2008).

Περίπτωση ανεξαρτησίας

Υποθέτουμε ότι τα πρόσθετα ποσά ζημιών $\{X_{ij}\}$ ακολουθούν μια ανεξάρτητη διαδικασία Γάμμα, δηλαδή $X_{ij} \sim G(\alpha_{ij}, \beta_{ij})$ ανεξάρτητα για i=1, 2, …,n έτσι ώστε $E[X_{ij}]= \alpha_{ij} / \beta_{ij}$. Για να δώσουμε στις παραμέτρους μια ερμηνεία θεωρούμε $\alpha_{ij}= \alpha_i$ για όλα τα j και $\beta_{ij}= \beta_i$ για όλα τα i. Το α_i μπορεί να ερμηνευτεί ως το μέσο συνολικό ποσό που πληρώνεται για τις ζημιές που έχουν έτος προέλευσης i. Για το β_j θεωρούμε ότι ικανοποιεί:

$$\sum_{j=1}^{n} \frac{1}{\beta_j} = 1$$

Έτσι το $\frac{1}{\beta_j}$ μπορεί να ερμηνευτεί ως το μέσο των συνολικών ζημιών του j έτους ανάπτυξης.

Για να εκτιμήσουμε τις παραμέτρους του μοντέλου και να προβλέψουμε τις μελλοντικές ζημιές ακολουθούμε μια Μπεϋζιανή προσέγγιση. Αρχικά πρέπει να ορίσουμε τις κατανομές των παραμέτρων $\alpha=(\alpha_1, \alpha_2, \ldots, \alpha_n)$ και $\beta=(\beta_1, \beta_2, \ldots, \beta_n)$. Θεωρούμε ότι τα $\alpha_i \sim Ga(\alpha_\alpha, \beta_\alpha)$ για i=1, 2, …,n είναι ανεξάρτητα και ότι η κατανομή των $1/\beta_j$ ακολουθεί την κατανομή Dirichlet $(1/\beta_1, 1/\beta_2, \ldots, 1/\beta_n) \sim Dir(c_1, c_2, \ldots, c_n)$. Αυτό σημαίνει πως το β ακολουθεί μια αντίστροφη κατανομή Dirichlet

$\beta \sim I\text{-}Dir(c_1, c_2, \ldots, c_n)$, με συνάρτηση πυκνότητας πιθανότητας:

$$\pi(\beta \mid c) = \frac{\Gamma\left(\sum_{j=1}^{n} c_j\right)}{\prod_{j=1}^{n} \Gamma(c_j)} \prod_{j=1}^{n-1} \beta_j^{-c_j-1} \left(1 - \sum_{j=1}^{n-1} \beta_j^{-1}\right)^{c_n-1}$$

$\beta_j \in (1, \infty)$ και $\sum_{j=1}^{n-1} 1/\beta_j \leq 1$

Οι προηγούμενες πληροφορίες χρειάζεται να ενημερωθούν με πληροφορίες από τα διαθέσιμα δεδομένα $X=\{ X_{ij}, i=1, 2, \ldots, n, j=1, 2, \ldots, n+1-i\}$.

Οι μεταγενέστερες υποθετικές κατανομές των α_i είναι

$$\pi\left(a_i \mid x,\beta\right) \propto \frac{1}{\Gamma\left(a_i\right)} a_i^{a_a-1}\left(e^{-b_a}\prod_{j=1}^{n+1-i}\beta_j x_{ij}\right)^{a_i} I\left(a_i>0\right)$$

με $\beta_j \geq \left(1-\sum_{k\neq j}\beta_j^{-1}\right)^{-1}$ για j=1, …, n-1 και $\beta_n = \left(1-\sum_{j=1}^{n-1}\beta_j^{-1}\right)^{-1}$

Με μπεϋζιανό περιεχόμενο το πρόβλημα της πρόβλεψης των μελλοντικών παρατηρήσεων λύνεται θεωρώντας την προγνωστική κατανομή, η οποία στην περίπτωση αυτή δίνεται από τον τύπο:

$$f\left(x_{ij}^F \mid x\right)=\iint Ga\left(x_{ij} \mid a_i,\beta_i\right)\pi\left(a_j,\beta_j \mid x\right)da_i d\beta_j$$

για i=1, 2, …,n και j=n+2-i, n+3-i, …, n. Ο εκθέτης F χρησιμοποιείται για να δηλώσει ένα μελλοντικό γεγονός.

Περίπτωση εξάρτησης

Για να μοντελοποιήσουμε την συσχέτιση μεταξύ των X_{ij} κατά τα χρόνια ανάπτυξης χρησιμοποιούμε την συσχετισμένη Γάμμα διαδικασία που εισήχθη από τους Nieto-Barajas και Walker (2002). Η συσχετισμένη διαδικασία Γάμμα $\{X_{ij}\}$ κατασκευάζεται μέσω της διαδικασίας $\{Z_{ij}\}$ με τον ακόλουθο τρόπο:

Για κάθε i: $X_{i,1}$~Ga(α_{i1}, β_{i1}) και για κάθε j=2, 3, …

$Z_{ij}|X_{i,j-1}$~P_0 ($\gamma_{ij}X_{i,j-1}$) και επίσης $X_{ij}|$ Z_{ij} ~ Ga ($\alpha_{ij}+Z_{ij}$, $\beta_{ij}+\gamma_{ij}$) όπου οι παράμετροι α_{ij}, β_{ij} και γ_{ij} είναι θετικοί για όλα τα i και j. Για κάθε i η $\{X_{ij}\}$είναι μια Μαρκοβιανή διαδικασία με τις ακόλουθες ιδιότητες:

$E[X_{ij}|X_{i,j-1}]=\dfrac{a_{ij}+\gamma_{ij}x_{i,j-1}}{\beta_{ij}+\gamma_{ij}}$ και $Var[X_{ij}|X_{i,j-1}]=\dfrac{a_{ij}+2\gamma_{ij}x_{i,j-1}}{(\beta_{ij}+\gamma_{ij})^2}$

Κατά την αποθεματοποίηση ζημιών υποθέτουμε ότι :

- για κάθε χρόνο προέλευσης i, το $\{X_{ij}\}$ ακολουθεί μια μη στάσιμη συσχετισμένη διαδικασία Γάμμα κατά το χρόνο ανάπτυξης j και
- για διαφορετικά χρόνια προέλευσης i και i΄, i ≠ i΄, υποθέτουμε ότι οι διαδικασίες $\{X_{ij}\}$ και $\{X_{i'j}\}$ είναι ανεξάρτητες .

Επομένως για κάθε i=1, ..,n οι διαδικασίες $\{X_{ij}\}$ θα ορίζονται μέσω των διαδικασιών $\{Z_{ij}\}$ και τα σύνολα των παραμέτρων $\{\alpha_{ij}\}$, $\{\beta_{ij}\}$και $\{\gamma_{ij}\}$. Θεωρούμε $\alpha_{ij}=\alpha_i$ για όλα τα j, $\beta_{ij}=\beta_j$ και $\gamma_{ij}=\gamma_j$ για όλα τα i . Σύμφωνα με τα παραπάνω συνεπάγεται ότι η προσδοκώμενη τιμή του X_{ij} δεδομένου του $X_{i,j-1}$ έχει τη μορφή :

$$E[X_{ij}|\, X_{i,j-1}]=(1-\lambda_j)\frac{\alpha_i}{\beta_j}+\lambda_j X_{i,j-1}$$

με $\lambda_j=\dfrac{\gamma_j}{\beta_j+\gamma_i}$ και $\lambda_j\in(0,1)$, j=1, ...,n. Από τον τελευταίο τύπο φαίνεται ότι το X_{ij} είναι ένας κυρτός συνδυασμός των ποσοτήτων α_i/β_j και του συνολικού ποσού ζημιών από το προηγούμενο έτος $X_{i,j-1}$. Ο συντελεστής λ_j μπορεί να ερμηνευτεί ως αποτέλεσμα της εξάρτησης, όσο μεγαλύτερο είναι το λ_j τόσο μεγαλύτερη εξάρτηση υπάρχει μεταξύ των X_{ij} και $X_{i,j-1}$. Επιπλέον το Z_{ij} εκφράζει την προσαύξηση στο μέσο συνολικό ποσό α_i, συγκεκριμένα για την χρονιά ανάπτυξης j, το οποίο μας επιτρέπει να μετράμε την εξάρτηση μεταξύ των ετών j-1 και j. Ένα ενδιαφέρον χαρακτηριστικό αυτής της διαδικασίας είναι ότι αν $\gamma_j=0$ για όλα τα j, τότε τα λ_j είναι μηδέν και τότε το μοντέλο παίρνει τη μορφή του ανεξάρτητου μοντέλου.

Έστω $X=\{X_{ij}$, i=1, 2, …,n, j=1, 2, …, n+1-i}, $Z=\{X_{ij}$, i=1, 2, …,n, j= 2, …, n+1-i}, $\alpha=(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\beta=(\beta_1, \beta_2, \ldots, \beta_n)$ και $\gamma=(\gamma_1, \gamma_2, \ldots, \gamma_n)$, τότε το μοντέλο, που προτείνουμε στο πρόβλημα που περιγράψαμε με την τριγωνική μορφή, συνοψίζεται από την ακόλουθη κατανομή :

$$f(x,z|\alpha,\beta,\gamma)=\prod_{i=1}^{n}\left[Ga\left(x_{i1}\,|\,a_i,\beta_1\right)\prod_{j=2}^{n+1-i}\left\{Po\left(z_{ij}\,|\,\gamma_j x_{i,j-1}\right)*Ga\left(x_{ij}\,|\,a_i+z_{ij},\beta_j+\gamma_j\right)\right\}\right]$$

Επίσης ισχύει πάλι: $\sum_{j=1}^{n} \frac{1}{\beta_j} = 1$. Χρησιμοποιούμε τις ίδιες κατανομές για τα **α** και **β** όπως στην εξαρτημένη περίπτωση. Επιπλέον θεωρούμε $\gamma_j \sim Ga(\alpha_\gamma, b_\gamma)$, για j=2, …,n, τα οποία μεταξύ τους είναι ανεξάρτητα.

Οι προηγούμενες πληροφορίες ενημερώνονται με τα διαθέσιμα δεδομένα **X** και **Z.** Τότε η πιθανότητα για (**α**, **β**, **γ**) δίνεται από τον τύπο:

lim (**α**, **β**, **γ| x**, **z**) ∞ f(**x**, **z** | **α**, **β**, **γ**)

Οι προγενέστερες κατανομές για (**α, β, γ**) χαρακτηρίζονται από τρεις υποθέσεις:

i. Οι υποθετικές προγενέστερες κατανομές των α_i είναι:

$$\pi(\alpha_i \mid \mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \propto \frac{\left\{\beta_1 x_{i1} \prod_{j=2}^{n+1-i} (\beta_j + \gamma_j) x_{ij}\right\}^{a_i}}{\Gamma(a_\iota) \prod_{j=2}^{n+1-i} \Gamma(a_i + z_{ij})} Ga(a_i \mid a_a, b_a)$$

Υποθέτουμε ακόμη ότι τα α_i ανεξάρτητα δεδομένων των (**x, z, β, γ**).

ii. Οι υποθετικές προγενέστερες κατανομές των β_j είναι:

$$\pi(\beta_1 \mid \mathbf{x}, \mathbf{z}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}_{-1}) \propto \beta_1^{\sum_{i=1}^{n} a_i - \gamma_i - 1} * e^{-\beta_1 \sum_{i=1}^{n} x_{i1}} \left(1 - \sum_{k=1}^{n-1} \beta_k^{-1}\right)^{\gamma_n - 1}$$

με $\beta_1 \geq \left(1 - \sum_{k \neq 1} \beta_k^{-1}\right)^{-1}$ και

$$\pi(\beta_j \mid \mathbf{x}, \mathbf{z}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}_{-j}) \propto \left(\beta_j + \gamma_j\right)^{\sum_{i=1}^{n+1-i} (a_i + z_{ij})} e^{-\beta_j \sum_{i=1}^{n+1-j} x_{ij}} \beta_j^{-\gamma_j - 1} \left(1 - \sum_{k=1}^{n-1} \beta_k^{-1}\right)^{\gamma_n - 1}$$

με $\beta_j \geq (1 - \sum_{k \neq j} \beta_k^{-1})^{-1}$ για j=2, …, n-1 και $\beta_n = (1 - \sum_{j=1}^{n-1} \beta_j^{-1})^{-1}$

iii. Τέλος, οι υποθετικές προγενέστερες κατανομές των γ_j είναι:

π(γ_j|**x,z,α,γ**)

$$\propto \left(\beta_j+\gamma_j\right)^{\sum_{i=1}^{n+1-i}\left(a_i+z_{ij}\right)} \gamma_j^{\sum_{i=1}^{n+1-j} z_{ij}} e^{-\gamma_j \sum_{i=1}^{n+1-j}\left(x_{ij-1}+x_{ij}\right)} Ga(\gamma_j \mid a_\gamma, b_\gamma)$$

για j=2, .., n.. Αυτές οι προγενέστερες κατανομές του γ_j δεδομένων των (**x, z, α, β**) είναι επίσης ανεξάρτητες.

iv. Οι υποθετικές κατανομές των Z_{ij} δίνονται από τον τύπο :

$$f(z_{ij} \mid \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \propto \frac{\left\{\gamma_j(\beta_j+\gamma_j)x_{i,j-1}x_{ij}\right\}^{z_{ij}}}{\Gamma(1+z_{ij})\Gamma(a_i+z_{ij})}$$

με $z_{ij} \in$ (0, 1, … } για i= 1, …, n και j=1, … , n+1-i

Ορίζουμε με παρόμοιο τρόπο τις μελλοντικές προβλεπόμενες αυξητικές ζημιές X_{ij}^F για i=2, 3, …, n και j= n+2-I, n+3-i, …,n. Ισχύει ότι

$$Z_{ij}^F \mid X_{ij-1}^F \sim Po\left(\gamma_j X_{ij-1}^F\right) \text{ και } X_{ij}^F \mid Z_{ij}^F \sim Ga(a_i+Z_{ij}^F, \beta_j+\gamma_j)$$

Συγκεκριμένα $X_{i,n+1-i}^F = X_{i,n+1-i}$ για i=1, …, n. Η προβλεπόμενη κατανομή σε αυτήν την περίπτωση γίνεται:

$$f\left(x_{ij}^F \mid z_{ij}^F, x\right) = \iiint Ga(x_{ij}^F \mid a_i+z_{ij}^F, \beta_j+\gamma_j)\pi(a_i,\beta_i,\gamma_i \mid x,z)da_i d\beta_j d\gamma_j$$

Παρατηρούμε ότι η προβλεπόμενη κατανομή εξαρτάται από την μεταβλητή Z_{ij}. Έτσι για να μπορέσουμε να κάνουμε τις ζητούμενες προβλέψεις, υπολογίζουμε και την παρακάτω κατανομή:

$$f\left(z_{ij}^F \mid x_{ij-1}^F, x\right) = \int Po(z_{ij}^F \mid \gamma_j x_{i,j-1}^F)\pi(\gamma_i \mid x,z)d\gamma_j$$

Με την χρήση των δύο παραπάνω κατανομών μπορούμε να κάνουμε οποιαδήποτε πρόβλεψη. Αξίζει να σημειωθεί ότι και το ανεξάρτητο και το εξαρτημένο μοντέλο ορίζονται από βασικές παραμετρικές κατανομές.

- Υποθέσαμε ότι οι πρόσθετες ζημιές ακολουθούν την κατανομή Γάμμα, έτσι μας παρέχεται ένα επαρκές μοντέλο για αυτόν τον τύπο δεδομένων που ανταποκρίνεται περισσότερο στην πραγματικότητα.
- Στην περίπτωση της εξαρτημένης περίπτωσης θεωρούμε ότι υπάρχει συσχέτιση μεταξύ των ετών ανάπτυξης και όχι μεταξύ των ετών συμβάντων.
- Το μοντέλο χαρακτηρίζεται ως Μπεϋζιανό γιατί χρησιμοποιεί προγενέστερες κατανομές για όλες τις παραμέτρους.

2.4 Γενικευμένο Γραμμικό Μοντέλο

Οι McCullagh και Nelder (1989) παρείχαν μια λεπτομερής εισαγωγή στα Γενικευμένα Γραμμικά Μοντέλα. Στα βιβλία του Atkins et al. (1989) και Dobson (1990) παρατίθενται πολλά παραδείγματα εφαρμογών της μεθόδου Γενικευμένου Γραμμικού Μοντέλου. Οι Hardin και Hilbe (2007) παρείχαν ένα εγχειρίδιο με οδηγίες για ανάλυση δεδομένων με τη χρήση της μεθόδου του Γενικευμένου Γραμμικού Μοντέλου. Οι Anderson et al. (2005), Coutts (1984), Brockman και Wright (1992) έκαναν πολύ καλές αναφορές στον υπολογισμό του καθαρού κινδύνου από το Γενικευμένο Γραμμικό Μοντέλο.

Το Μοντέλο

Τα Γενικευμένα Γραμμικά Μοντέλα είναι μια γενίκευση του κλασικού γραμμικού μοντέλου σύμφωνα με το οποίο ο μέσος του πληθυσμού εξαρτάται από ένα γραμμικό εκτιμητή μέσω μίας συνάρτησης σύνδεσης.

Ένα γενικευμένο Γραμμικό Μοντέλο αποτελείται από τις ακόλουθες τρεις συνιστώσες:

1) Την Y που ακολουθεί κατανομή από την εκθετική οικογένεια με συνάρτηση πυκνότητας πιθανότητας να έχει τη μορφή

$$f(y;\theta,\varphi) = \exp\left\{\int \frac{[y-\mu(\theta)]}{\varphi V(\mu)}\, d\mu(\theta) + c(\mu,\varphi)\right\}$$

όπου θ ονομάζεται η φυσική παράμετρος, φ είναι μία παράμετρος διασποράς $\mu=\mu(\theta)=E(Y)$ και $V(Y) = \varphi V(\mu)$ για δεδομένη συνάρτηση διασποράς V και γνωστή διμεταβλητή συνάρτηση c. Η εκθετική κατανομή είναι ευέλικτη και μπορεί να μοντελοποιήσει συνεχή, δυαδικά και μετρήσιμα δεδομένα.

2) Για ένα τυχαίο δείγμα $Y_1, \ldots, Y_n$, η γραμμική συνιστώσα ορίζεται ως :

$n_i = X_i^{'}$ **β,** i=1,…, n

3) Μια μονότονη διαφορίσιμη συνάρτηση g περιγράφει πως ο προσδοκούμενος μέσος $\mu_i = E(Y_i)$ σχετίζεται με τον γραμμικό εκτιμητή n_i

$g(\mu_i) = n_i$, i=1,…, n

Η μέθοδος εκτίμησης μέγιστης πιθανοφάνειας χρησιμοποιείται για να υπολογίσουμε την παράμετρο ***β***. Για ένα ανεξάρτητο τυχαίο παρατηρήσιμο δείγμα $y_1, y_2, \ldots, y_n$, θεωρούμε την log πιθανοφάνεια της ***β***.

$$l(\boldsymbol{\beta}) = \ln L(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left\{ \int \frac{[y_i - \mu_i(\theta)]}{\varphi V(\mu_i)} d\mu_i(\theta) + c(y_i, \varphi) \right\}$$

και την παράγωγό του :

$$\frac{dl(\beta)}{d\beta} = \sum_{i=1}^{n} \frac{dl(\beta)}{d\mu_i} \frac{d\mu_i}{d\beta} = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{\varphi V(\mu_i)} \frac{d\mu_i}{dX_i\beta} \frac{dX_i\beta}{d\beta}$$

όπου

$$\frac{d\mu_i}{dX_i\beta} = \frac{dg^{-1}(X_i\beta)}{dX_i\beta} = \frac{1}{g^{'}(\mu_i)}$$

άρα

$$\frac{dl(\beta)}{d\beta} = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{\varphi V(\mu_i)} \frac{1}{g^{'}(\mu_i)} X_i^{'}$$

Αν Y_i ακολουθεί μια κανονική κατανομή, τότε $g'(\mu_i)=1$ και $V(\mu_i)=1$ για όλα τα i. Θέτοντας :

$$\frac{dl(\beta)}{d\beta}=0 \text{ έχει ως αποτέλεσμα } \sum_{i=1}^{n} X_i(y_i - X_i'\beta)=0$$

Σε κάποιες περιπτώσεις που δεν είναι διαθέσιμη λύση με την απλή επίλυση των εξισώσεων που προκύπτουν από τα παραπάνω, χρησιμοποιούνται αριθμητικές μέθοδοι όπως είναι η μέθοδος Newton – Raphson.

Αποθεματοποίηση ζημιών με το γενικευμένο γραμμικό μοντέλο

Έστω ότι η συνάρτηση ατυχημάτων l(t) είναι μια στοχαστική διαδικασία που αντιπροσωπεύει την συχνότητα με την οποία συμβαίνουν οι ζημιές στον χρόνο t. Δηλαδή μια συνάρτηση ατυχημάτων l(t) μας λέει πώς τα ατυχήματα συμβαίνουν.

Οι συνολικές ζημιές $L(t_1, t_2)$ που συμβαίνουν την χρονική περίοδο (t_1, t_2) μπορούν να γραφτούν ως εξής:

$$L\left(t_1, t_2\right)=\int_{t_1}^{t_2} l(t)dt$$

Η συνάρτηση ανάπτυξης των ατυχημάτων D(t) είναι μια στοχαστική διαδικασία το ποσοστό των ατυχημάτων που πληρώνονται μέσα σε t χρόνια μετά το συμβάν του ατυχήματος. Είναι ξεκάθαρο ότι D(t)=0 για t≤0 και $\lim_{t\to\infty} D(t)=1$ Για ένα δεδομένο χρόνο T>t, το l(t) D(T-t) αντιπροσωπεύει το συνολικό ποσό πληρωμών στο χρόνο T

για ζημιές που συμβαίνουν την χρονική στιγμή t.

Δεδομένου μιας συνάρτησης ατυχημάτων l(t) και μια στοχαστική συνάρτηση ανάπτυξης ατυχημάτων D(t), οι συνολικές πληρωμένες ζημιές από τα ατυχήματα που συνέβησαν την χρονική περίοδο (t_1, t_2) στο χρόνο ανάπτυξης $T \geq t_1 \geq t_2$, ορίζονται ως:

$$L\left(T, t_1, t_2\right)=\int_{t_1}^{t_2} l(t)D(T-t)dt$$

Η διαφορά $L(t_1, t_2) - L(T, t_1, t_2)$ είναι οι απλήρωτες ζημιές που ονομάζεται και απόθεμα ζημιών.

Έστω

$$R(T)= L(t_1, t_2) - L(T, t_1, t_2)=\int_{t_1}^{t_2} l(t)\left[1-D(T-t)\right]dt \text{, για } T\geq t_1\geq t_2$$

Ότι είναι το απόθεμα ζημιών για τις υφιστάμενες ζημιές της χρονικής περιόδου (t_1,t_2).

Στην περίπτωση των προεξοφλημένων αποθεμάτων, χρειάζεται να προσθέσουμε ένα προεξοφλητικό παράγοντα στην παραπάνω ανάλυση. Έστω δ(t) η στοχαστική ισχύς του επιτοκίου τη χρονική στιγμή t. Υποθέτουμε ότι η διαδικασία είναι συνεχής τότε ορίζουμε ως $B(t)=\int_0^t \delta(s)ds$ το συνολικό επιτόκιο της περιόδου (0, t) και γενικότερα $B(T+t)-B(t)=\int_T^{T+t} \delta(s)ds$ είναι το συνολικό επιτόκιο για το χρονικό διάστημα (T, T+t) .

Θεωρούμε ότι για μια συγκεκριμένη χρονική στιγμή όπου $t_1<t< t_2$ όπου συμβαίνουν οι ζημιές l(t). Τότε l(t) d(s-t)ds θα αναπτυχθούν στο μέλλον ότι s>t που υποθέτουμε d(t)= D΄(t). Επομένως η διαφορά της τιμής μεταξύ της ονομαστικής αξίας και τις τιμής εκείνη την χρονική στιγμή δίνεται από το $e^{-[B(s)-B(T)]}$ l(t)d(s-t)ds. Τελικά ολοκληρώνοντας για όλες τις μελλοντικές στιγμές ανάπτυξης $s\in(T,\infty)$.

Επίσης ορίζεται:

$$Z(T) = \int_{t_1}^{t_2} l(t)\int_T^{\infty} e^{-[B(s)-B(T)]}\frac{\partial}{\partial t}D(s-t)dsdt$$

ως η διαφορά της ονομαστικής αξίας και της αξίας στο χρόνο T των απλήρωτων αποθεμάτων ατυχημάτων για την περίοδο (t_1, t_2).

Εισάγουμε τις ακόλουθες υποθέσεις για να υπολογίσουμε την προσδοκώμενη τιμή της R(t) και της Z(T) :

(1) Η συνάρτηση ατυχημάτων l(t), η συνάρτηση ανάπτυξης ατυχημάτων D(t) και η ισχύς του επιτοκίου δ(t) είναι ανεξάρτητες. Από την υπόθεση αυτή συνεπάγεται ότι:

Θεώρημα 1: Δεδομένων των E[l(t)], E[D(t)] και E[B(s)], για συγκεκριμένα t και s, τότε

$$E[R(t)]=\int_{t_1}^{t_2} E\left[l(t)\right]\left\{1-E\left[D(T-t)\right]\right\} dt\ ,\quad T\geq t_2 \geq t_1$$

$$E[Z(T)]= \int_{T}^{\infty} e^{-E[B(s)]}\int_{t_1}^{t_2} E\left[l(t)\right]E\left[\frac{\partial}{\partial t}D(s-t)\right]dtds$$

Είναι προφανές ότι τα αποθέματα ζημιών εξαρτώνται μόνο από την προσδοκώμενη συνάρτηση ζημιών και την προσδοκώμενη συνάρτηση ανάπτυξης ζημιών. Χρησιμοποιώντας το Γενικευμένο Γραμμικό Μοντέλο θα υπολογίσουμε αυτές τις δυο συναρτήσεις.

Στην πράξη οι συναρτήσεις ατυχημάτων και ανάπτυξης ατυχημάτων μπορεί να είναι πολύπλοκες. Ακόμη και αν είναι διαθέσιμος μεγάλος αριθμός ιστορικών δεδομένων, η διαδικασία ανάπτυξης μπορεί να αλλάξει με το χρόνο. Η μελλοντική ισχύς του επιτοκίου είναι άγνωστη.

(2) Όλες οι περίοδοι ασφαλιστηρίου είναι ένα έτος και το ποσό έκθεσης στον κίνδυνο του ασφαλιστικού συμβολαίου εξαπλώνεται ομοιόμορφα σε όλη την περίοδο του ασφαλιστηρίου.

(3) Η προσδοκώμενη τιμή της συνάρτησης ανάπτυξης των ατυχημάτων D είναι της μορφής $E[D(t)]= 1-\alpha^{-t}$, όπου το α είναι μια σταθερά και α>1.

(4) Η μελλοντική ισχύς του επιτοκίου είναι μια γνωστή σταθερά δ≥0.

Λήμμα 1: Δεδομένης της τρίτης υπόθεσης ,ο προσδοκώμενος μέσος χρόνος ανάπτυξης είναι $E(\tau)=\frac{1}{\ln a}$, για α>1.

Η πρώτη υπόθεση μας παρακινεί να εκτιμήσουμε τις προσδοκώμενες τιμές της συνάρτησης ατυχημάτων E[l(t)], της συνάρτησης ανάπτυξης ατυχημάτων E[D(t)] και της συνάρτησης προεξοφλητικού επιτοκίου E[B(s)]. Η υπόθεση (2) μας απαλλάσσει από την εποχικότητα άλλων κατανεμημένων σχημάτων. Σύμφωνα με την Τρίτη υπόθεση το E[D(t)] παίρνει τη μορφή της αθροιστικής συνάρτησης κατανομής μιας εκθετικής κατανομής, η οποία είναι κατάλληλη για υψηλής συχνότητας/χαμηλής δριμύτητας επιχειρηματικής γραμμής όπως είναι οι ασφάλειες αυτοκινήτων.

Έχοντας κάνει όλες αυτές τις υποθέσεις μπορούμε να εκτιμήσουμε την προσδοκώμενη τιμή της συνάρτησης ατυχημάτων l(t) και της συνάρτησης ανάπτυξης D(t) μέσα στα πλαίσια του Γενικευμένου Γραμμικού Μοντέλου.

Θεωρούμε ένα σύνολο παρατηρήσιμων ζημιών κάτω από ένα σύστημα ταξινόμησης. Έστω ότι το κελί i συμβολίζει μια γενική τάξη ορισμένη από το σύστημα ταξινόμησης. Το γενικευμένο γραμμικό μοντέλο μπορεί να γραφεί ως ακολούθως. Αν :

- f_i είναι η συχνότητα στο κελί i
- z_i η δριμύτητα των ζημιών στο κελί i
- τ_i είναι η μέση πληρωμή του χρόνου στο κελί i
- $w_i(t)$ ο αριθμός των εκθέσεων στο κελί i και στο χρόνο t.
- n_{fi}, n_{zi} και $n_{\tau i}$ είναι οι γραμμικοί εκτιμητές της συχνότητας των ζημιών, της δριμύτητας και των μέσων πληρωμών στο κελί i αντίστοιχα.
- g_f, g_z και g_τ είναι οι συναρτήσεις σύνδεσης του γενικευμένου γραμμικού μοντέλου, της συχνότητας των ζημιών, της δριμύτητας και των μέσων πληρωμών στο κελί i αντίστοιχα.

Για κάθε κελί i, το γενικευμένο γραμμικό μοντέλο δίνει την προσδοκώμενη τιμή της συχνότητας των ζημιών, της δριμύτητας και των μέσων πληρωμών στο κελί i αντίστοιχα:

$E[f_i]= g_f^{-1}(n_{f_i})$,

$E[z_i]= g_f^{-1}(n_{z_i})$,

$E[\tau_i]= g_f^{-1}(n_{\tau_i})$.

Από το Λήμμα 1, την υπόθεση (2) και την τελευταία σχέση έχουμε ότι :

$$g_{\tau}^{-1}(n_{\tau_i}) = \frac{1}{\ln a_i} \Rightarrow a_i = \exp\left\{\frac{1}{g_{\tau}^{-1}(n_{\tau_i})}\right\}$$

Σύμφωνα με την υπόθεση (2), παίρνουμε ότι η προσδοκώμενη συχνότητα συνολικών ατυχημάτων $E[l_i(t)]$ και η συνάρτηση ανάπτυξης ατυχημάτων $E[D_i(t)]$ στο κελί i και στο χρόνο t είναι:

$$E[l_i(t)] = w_i(t)\, g_f^{-1}(n_{f_i})\, g_s^{-1}(n_{s_i})$$

$$E[D_i(t)] = \exp\left\{\frac{1}{g_{\tau}^{-1}(n_{\tau_i})}\right\}$$

Επομένως

$$E[R_i(t)] = \int_{t_1}^{t_2} l_i(t)\,\{1 - E[D_i(T-t)]\}\,dt$$

$$E[Z_i(t)] = \int_T^{\infty} e^{\delta}(s-t)\int_{t_1}^{t_2} E\left[l_i(t)\right]E\left[\frac{\partial}{\partial s}E\left[D(s-t)\right]\right]dtds$$

Ως εκ τούτου, αθροίζοντας σε όλα τα κελιά του χαρτοφυλακίου, έχουμε τις συνολικές ζημιές και το προεξοφλημένο απόθεμα ζημιών αντίστοιχα :

$$E[R(T)] = \sum_i E\left[R_i(T)\right],$$

$$E[Z(T)] = \sum_i E\left[Z_i(T)\right].$$

Παρατηρήσεις

Το γενικευμένο γραμμικό μοντέλο σε σύγκριση με τα παραδοσιακά που βασίζονταν στην τριγωνική ανάπτυξη ζημιών μας δίνει περισσότερες πληροφορίες, όπως είναι η έκθεση των ασφαλίστρων, τα ατυχήματα και η ανάπτυξη των ατυχημάτων. Επίσης το μοντέλο αυτό μας παρέχει πιο λεπτομερή περιγραφή της εκτίμησης των αποθεμάτων των ζημιών.

2.5 Το γενικευμένο βελτιωτικό μοντέλο (generalized additive model)

Οι Renshaw και Verrall (1994) όρισαν το γενικευμένο γραμμικό μοντέλο που βασιζόταν στην μέθοδο «chain ladder» και πρότειναν και κάποια άλλα γενικευμένα γραμμικά μοντέλα που ήταν χρήσιμα στην αποθεματοποίηση ζημιών. Οι Hastie και Tibshirani, (1990,1993) διατήρησαν την υποκείμενη δομή των γενικευμένων γραμμικών μοντέλων, αλλά αντικατέστησαν τον γραμμικό εκτιμητή από μία μη παραμετρική εξομαλυντική διαδικασία. Παραδείγματα τέτοιων μη παραμετρικών εξομαλυντικών διαδικασιών αποτελούν η εξομάλυνση Whittaker και η εξομάλυνση του Σταθμισμένου Κινητού Μέσου. Αυτού του είδους οι εξομαλυντές χρησιμοποιούνται και στην αποθεματοποίηση των ζημιών. Το γενικευμένο βελτιωτικό μοντέλο μπορεί να εφαρμοστεί σε μια μεγάλη ποικιλία μοντέλων, γεγονός που δεν ισχύει για το γραμμικό «chain ladder» μοντέλο (Verall, 1996).

Ο σκοπός που χρησιμοποιείται το γενικευμένο βελτιωτικό μοντέλο είναι :

- η εξομάλυνση γύρω από τα χρόνια ατυχημάτων. Το μοντέλο «chain ladder» είναι υπερπαραμετρικοποιημένο και έτσι δεν υπάρχουν αρκετές πληροφορίες για τα χρόνια ατυχημάτων. Αυτό έχει ως αποτέλεσμα μια μη αξιόπιστη αποθεματοποίηση ζημιών.
- Τα χρόνια ατυχημάτων είναι περισσότερο συσχετισμένα μεταξύ τους από ότι .συνεπάγεται η παραμετρικοποίηση της τεχνικής «chain ladder». Πολλές προτάσεις έχουν γίνει πάνω σε αυτό, όπως η ομαδοποίηση των χρόνων ατυχημάτων χρησιμοποιώντας πολλαπλές συγκρίσεις t-test, η μέθοδος Bayes και η εμπειρική μέθοδος Bayes, και το φιλτράρισμα Kalman (Verall,1989, De Jong και Zehnwirth, 1983). Το γενικευμένο βελτιωτικό μοντέλο προσφέρει μία μέθοδο που έχει περισσότερες ομοιότητες με το φιλτράρισμα του Kalman και του δυναμικού γενικευμένου γραμμικού μοντέλου, αλλά πλεονεκτεί ως προς αυτές τις μεθόδου.

Το βελτιωτικό μοντέλο βασίζεται στην υπόθεση ότι υπάρχουν γνωστές παράμετροι $\pi_0, \pi_1, \ldots, \pi_n \in (0,\infty)$ και άγνωστες παραμέτρους $\zeta_0, \zeta_1, \ldots, \zeta_n$ έτσι ώστε να ισχύει η ταυτότητα :

$E[Z_{i,k}]=\zeta_k\pi_i$ για όλα τα i, k $\in \{0, 1, \ldots, n\}$.

Αν οι παράμετροι $\pi_1, \ldots, \pi_n$ ερμηνεύονται ως τα ασφάλιστρα ή άλλη μονάδα μέτρησης των ετών ατυχημάτων, τότε η υπόθεση είναι ότι για κάθε έτος ανάπτυξης k, οι προσδοκώμενοι οριακοί συντελεστές ζημιών:

$\zeta_{i,k} := E[\frac{Z_{i,k}}{\pi_i}]$ είναι ταυτόσημοι για όλα τα χρόνια ατυχημάτων.

Αν $\alpha_i := \pi_i \sum_{k=0}^{n}\zeta_k$ και $\gamma_k := \frac{\sum_{l=0}^{k}\zeta_l}{\sum_{l=0}^{n}\zeta_l}$

παρατηρούμε ότι $E[S_{i,k}]= \gamma_k\, \alpha_i$ για όλα τα i, k $\in \{0, 1, \ldots, n\}$ έτσι ώστε $\alpha_i = E[S_{i,n}]$ και οι παράμετροι $\gamma_0, \gamma_1, \ldots,\gamma_n$ σχηματίζουν ένα σχέδιο ανάπτυξης για αθροιστικές ποσοστώσεις.

Οι βελτιωτικοί εκτιμητές των πρόσθετων ζημιών $Z_{i,k}$ με $i+k \geq n$ ορίζονται ως :

$$\widehat{Z}_{ik}^{AD} := \widehat{\zeta}_k^{AD}\pi_i$$

και οι βελτιωτικοί εκτιμητές των αθροιστικών ζημιών $S_{i,k}$ με $i+k \geq n$ ορίζονται ως ;

$$\widehat{S}_{ik}^{AD} := S_{i,n-i} + \sum_{l=n-i+1}^{k}\widehat{Z}_{i,l}^{AD}$$

όπου

$$\widehat{\zeta}_k^{AD} := \frac{\sum_{j=0}^{n-k} z_{i,k}}{\sum_{j=0}^{n-k}\pi_j}$$

είναι η βελτιωτικός οριακός συντελεστής ζημιών το έτος ανάπτυξης k.

Αν

$$\hat{\gamma}_k^{AD} := \frac{\sum_{l=0}^{k} \hat{\zeta}_l^{AD}}{\sum_{l=0}^{n} \hat{\zeta}_l^{AD}}$$

και

$$\hat{a}_i^{AD} := \pi_i \sum_{l=0}^{n} \hat{\zeta}_l^{AD}$$

οι βελτιωτικοί εκτιμητές των μη παρατηρήσιμων αθροιστικών ζημιών μπορούν να γραφτούν ως:

$$\hat{S}_{ik}^{AD} := S_{i,n-i} + (\hat{\gamma}_k^{AD} - \hat{\gamma}_{n-1}^{AD})\hat{a}_i^{AD}$$

Αυτό δείχνει ότι οι βελτιωτικοί εκτιμητές των αθροιστικών ζημιών δεν είναι τίποτα άλλο από τους εκτιμητές Bornhuetter- Ferguson σύμφωνα με τα βελτιωτικές συνολικές ποσοστώσεις $\hat{\gamma}_k^{AD}$ και οι προγενέστεροι εκτιμητές $\hat{a}_i^{AD}$ των τελευταίων προσδοκώμενων ζημιών. Με άλλα λόγια το βελτιωτικό μοντέλο είναι μια ειδική περίπτωση της μεθόδου Bornhuetter- Ferguson με προγενέστερους εκτιμητές των αθροιστικών ποσοστώσεων και των τελευταίων προσδοκώμενων ζημιών να βασίζονται σε εσωτερικές και εξωτερικές πληροφορίες.

Οι προσδοκώμενες συνολικές συχνότητες ατυχημάτων

$$k_i := E\left[\frac{S_{i,n}}{\pi_i}\right]$$

ικανοποιούν την σχέση:

$$k_i = \sum_{l=0}^{n} \zeta_{i,l}$$

Αφού οι προσδοκώμενες πρόσθετες συχνότητες ατυχημάτων είναι όμοιες για όλα τα χρόνια ατυχημάτων, αυτό συνεπάγεται επίσης ότι οι προσδοκώμενες συνολικές συχνότητες ατυχημάτων είναι όμοιες για όλα τα χρόνια ατυχημάτων. Επομένως, η πρόσθετη συχνότητα ατυχημάτων

$$\hat{k}^{AD} := \sum_{l=0}^{n} \hat{\zeta}_l^{AD}$$

Μπορεί να ερμηνευτεί ως ένας εκτιμητής της τελευταίας προσδοκώμενης πρόσθετης συχνότητας ατυχημάτων

$$k := \sum_{l=0}^{n} \zeta_l$$

κοινό για όλα τα χρόνια ατυχημάτων. Επιπλέον οι προγενέστεροι εκτιμητές $\hat{a}_i^{AD}$ μπορούν να γραφούν ως

$$\hat{a}_i^{AD} := \pi_i \hat{k}^{AD}$$

Και άρα ισχύει :

$$\hat{k}^{AD} = \frac{\sum_{j=0}^{n} s_{i,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j}^{AD}}$$

Αυτό δείχνει ότι οι βελτιωτικοί εκτιμητές των μη παρατηρήσιμων αθροιστικών ζημιών δεν είναι τίποτα άλλο από τους Cape – Cod εκτιμητές λαμβάνοντας υπόψη τις βελτιωτικές συνολικές ποσοστώσεις $\hat{\gamma}_k^{AD}$. Με άλλα λόγια το βελτιωτικό μοντέλο είναι μια συγκεκριμένη περίπτωση της μεθόδου Cape – Cod με προγενέστερους εκτιμητές των συνολικών ποσοστώσεων που βασίζονται σε εσωτερικές και εξωτερικές πληροφορίες (Zhou et al., 2006).

2.6 Πολυμεταβλητό βελτιωτικό μοντέλο πρόβλεψης ζημιών

Μεγάλη πρόοδος έχει σημειωθεί στο χώρο της Αναλογιστικής με την επίτευξη της χρήσης του πολυμεταβλητού μοντέλου.

Πρόσφατα οι Radtke και Schmidt (2004) εισήγαγαν τα πολυμεταβλητά μοντέλα στην αποθεματοποίηση ζημιών. Η ανάπτυξη άρχισε με μια συγκεκριμένη διδιάστατη επέκταση της μεθόδου «chain-ladder» που πρότειναν οι Quarg και Mack (2004) που εφαρμόζεται σε υφιστάμενες αποζημιώσεις του ίδιου χαρτοφυλακίου. Ουσιαστικά, ο Braun (2004) πρότεινε ένα άλλο διδιάστατο μοντέλο με σκοπό να κατασκευάσουν συσχετισμένους-εξαρτημένους εκτιμητές για την πρόβλεψη των λαθών των μονοπαραγοντικών εκτιμητών της μεθόδου «chain-ladder». Το μοντέλο του Braun επεκτάθηκε από τους Pröhl και Schmidt (2005) που ανέπτυξε μία πολυμεταβλητή έκδοση της μεθόδου «chain-ladder». Ο Kremer (2005) πρότεινε μια άλλη πολυμεταβλητή επέκταση της μεθόδου «chain-ladder», η οποία είναι λιγότερο κατάλληλη για εφαρμογές, αφού απαιτεί την αντιστροφή πολύ μεγαλύτερων πινάκων (Klaus et al., 2006).

Όμως μέχρι τώρα κανένα επίσημο έγγραφο δεν έχει δικαιολογήσει την επέκταση του πολυμεταβλητού μοντέλου. Εδώ θα περιγράψουμε μια γενική πολυμεταβλητή επέκταση του βελτιωτικού μοντέλου (additive model) και προσδιορίζουμε, κάτω από τις υποθέσεις αυτού του μοντέλου, τους εκτιμητές Gauss-Markov για τις μη παρατηρήσιμες αυξητικές ζημιές.

Έστω (Ω, F, P) είναι ένας χώρος πιθανοτήτων μέσα στον οποίο όλες οι τυχαίες μεταβλητές ορίζονται. Θεωρούμε ότι όλες οι μεταβλητές είναι διπλά ολοκληρώσιμες και ότι τα τυχαία διανύσματα έχουν διπλά ολοκληρώσιμες μεταβλητές.

Πρόβλεψη στο γραμμικό μοντέλο

Θεωρούμε το τυχαίο διάνυσμα **X** και υποθέτουμε ότι το **X** ικανοποιεί το γραμμικό μοντέλο

$E[\mathbf{X}] = \mathbf{A}\boldsymbol{\beta}$

όπου **A** είναι ένας πίνακας και **β** ένα άγνωστο διάνυσμα παραμέτρων.

Υποθέτουμε ότι κάποιες από τις συνιστώσες του **X** είναι παρατηρήσιμες, ενώ μερικές δεν είναι. Τότε το τυχαίο διάνυσμα X_1, το οποίο αποτελείται από τις παρατηρίσιμες συνιστώσες του **X** και το τυχαίο διάνυσμα $\mathbf{X_2}$, το οποίο αποτελείται από τις μη παρατηρήσιμες συνιστώσες του **X** ικανοποιούν τις παρακάτω σχέσεις:

$E[\mathbf{X_1}] = \mathbf{A_1} \boldsymbol{\beta}$

$E[\mathbf{X_2}] = \mathbf{A_2} \boldsymbol{\beta}$

για τους υποπίνακες $\mathbf{A_1}$ και $\mathbf{A_2}$ του **A**.

Υποθέτουμε επίσης ότι οι πίνακες

$\Sigma_{11} := \mathrm{Var}\,[\mathbf{X_1}]$

$\Sigma_{12} := \mathrm{Cov}\,[\mathbf{X_2}, \mathbf{X_1}]$

είναι γνωστοί και ότι ο πίνακας Σ_{11} είναι αντιστρέψιμος. Αφού οι συνιστώσες του $\mathbf{X_2}$ είναι άγνωστες, μόνο το τυχαίο διάνυσμα $\mathbf{X_1}$ μπορεί να χρησιμοποιηθεί για τον υπολογισμό της παραμέτρου **β**. Ο εκτιμητής του **β** είναι :

$$\beta^* := \left(A_1 \sum_{11}^{-1} A_1 \right)^{-1} A_1 \sum_{11}^{-1} X_1$$

Ας θεωρήσουμε τώρα την πρόβλεψη του μη παρατηρήσιμου τυχαίου διανύσματος $\mathbf{DX_2}$, που **D** είναι πίνακας κατάλληλης διάστασης. Μία τυχαία μεταβλητή $\hat{Y}$ είναι μια πρόβλεψη για το $\mathbf{DX_2}$. Για μια πρόβλεψη $\hat{Y}$ του $\mathbf{DX_2}$, ο αριθμός

$$E\left[\left(\hat{Y} - DX_2\right)\left(\hat{Y} - DX_2\right)\right]]$$

λέγεται το αναμενόμενο τετραγωνικό λάθος πρόβλεψης του $\hat{Y}$ **.**

Ένα παρατηρήσιμο τυχαίο διάνυσμα $\hat{Y}$ λέγεται ότι είναι

— Μία γραμμική πρόβλεψη του $\mathbf{DX_2}$ αν υπάρχει πίνακας **Q** έτσι ώστε $\hat{Y} = QX_1$

— Ένας αμερόληπτος εκτιμητής του $\mathbf{DX_2}$ αν E[$\hat{Y}$]=[DX_2]

— Ένας Gauss- Markov εκτιμητής του **DX₂,** αν είναι ένας γραμμικός αμερόληπτος εκτιμητής του **DX₂** και αν ελαχιστοποιεί το προσδοκώμενο τετραγωνικό λάθος της εκτίμησης για όλους τους αμερόληπτους γραμμικούς εκτιμητές του **DX₂**.

Έχουμε το ακόλουθο αποτέλεσμα:

Πρόταση 1: (Θεώρημα Gauss Markov) Υπάρχει ένας μοναδικός εκτιμητής **Y* (DX₂)** του **DX₂** ο οποίος ικανοποιεί:

$$Y^*\left(DX_2\right) = D\left(A_2\beta^* + \Sigma_{21}\Sigma_{11}^{-1}\left(X_1 - A_1\beta^*\right)\right)$$

Συγκεκριμένα, **Y* (DX₂)= DY* (X₂)** .

Εφαρμογή στην Αποθεματοποίηση Ζημιών

Θεωρούμε m χαρτοφυλάκια που όλα έχουν τον ίδιο αριθμό ετών ανάπτυξης.

Τα m χαρτοφυλάκια ίσως ερμηνευτούν ως υποχαρτοφυλάκια ενός συνολικού χαρτοφυλακίου. Για το χαρτοφυλάκιο p ∊ {1, 2, …, m}, συμβολίζουμε με $Z_{i,k}^{(p)}$ το αυξητικό μέγεθος ζημιών του χρόνου ατυχήματος i ∊ {0, 1, …, n} και του έτους ανάπτυξης k ∊ {0, 1, …, n}.

Για i, k ∊ {0, 1, …, n} ,παίρνουμε το τυχαίο διάνυσμα των αυξητικών ζημιών που έχει διάσταση m :

$$Z_{i,k} := (Z_{i,k}^{(p)})_{p\in(1,\ldots,m)}$$

Οι παρατηρήσιμες επαυξήσεις των ζημιών έχουν την ακόλουθη τριγωνική μορφή:

Έτος συμβάντος	Έτος εξέλιξης								
	0	**1**	**…**	**k**	**…**	**n-i**	**…**	**n-1**	**n**
0	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,k}$	…	$Z_{0,n-i}$	…	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1.0}$	$Z_{1.1}$	…	$Z_{1.k}$	…	$Z_{1.n-i}$	…	$Z_{1.n-1}$	
⋮	⋮	⋮	⋯	⋮	⋯	⋮			
I	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,k}$	…	$Z_{i,n-i}$			
⋮	⋮	⋮	⋯	⋮					
n-k	$Z_{n-k,0}$	$Z_{n-k,1}$	…	$Z_{n-k,k}$					
⋮	⋮	⋮							
n-1	$Z_{n-1,0}$	$Z_{n-1,1}$							
N	$Z_{n-1,n}$								

Εμείς τώρα διατυπώνουμε το πολυμεταβλητό βελτιωτικό μοντέλο :

Υπάρχουν θετικά ορισμένοι συμμετρικοί πίνακες $\mathbf{V_0}, \mathbf{V_1}, \ldots, \mathbf{V_n}$ και $\mathbf{\Sigma_0}, \mathbf{\Sigma_1}, \ldots, \mathbf{\Sigma_n}$ και τα άγνωστα διανύσματα $\zeta_0, \zeta_1, \ldots, \zeta_n$ έτσι ώστε

$E[\mathbf{Z}_{i,k}] = \mathbf{V}_i \zeta_k$

και

$$Cov\left[Z_{i,k}, Z_{j,1}\right] = \begin{cases} V_i^{1/2} \Sigma_k V_i^{1/2} & \alpha\nu \text{ i=j } \kappa\alpha\iota \text{ k=l} \\ 0 & \alpha\lambda\lambda\iota\omega\varsigma \end{cases}$$

για όλα τα i, j, k, l $\in$ {0, 1, …, n}.

Το πολυμεταβλητό βελτιωτικό μοντέλο είναι μία γενική επέκταση του μονοπαραγοντικού βελτιωτικού μοντέλου που τεκμηριώθηκε από τον Schmidt (2004b). Συγκεκριμένα, κάθε ένας από τους πίνακες ίσως επιλεγεί να είναι διαγώνιος έτσι ώστε να αντιπροσωπεύει τον όγκο των μέτρων του έτους συμβάντος i.

Υποθέτουμε ότι οι υποθέσεις του βελτιωτικού πολυμεταβλητού μοντέλου πληρούνται.

Εξαιτίας των υποθέσεων στις προσδοκίες των αυξητικών ζημιών, το πολυμεταβλητό βελτιωτικό μοντέλο είναι ένα γραμμικό μοντέλο.

Ορίζουμε:

$$\beta := \begin{pmatrix} \zeta_0 \\ \zeta_1 \\ \dots \\ \zeta_{k-1} \\ \zeta_k \\ \zeta_{k+1} \\ \dots \\ \zeta_n \end{pmatrix}$$

Και για όλα τα i, k $\in$ {0, 1, …, n} ορίζουμε

$A_{i,k} := (\ 0\ \ 0\ \ \dots\ 0\ \ V_i\ \ 0\ \ \dots\ \ 0)$

που ο πίνακας V_i συμβαίνει στην θέση k+1. Τότε έχουμε

$E[Z_{i,k}] = A_{i,k}\,\beta$

για όλα τα i, k $\in$ {0, 1, …, n}. Έστω ότι με $\mathbf{Z_1}$ και $\mathbf{A}_1$ συμβολίζουμε ένα διάνυσμα κορμό και ένα πίνακα κορμό που αποτελούνται από τα διανύσματα $\mathbf{Z}_{i,k}$ και τους πίνακες $\mathbf{A}_{i,k}$ αντίστοιχα με $i+k \leq n$ και $\mathbf{Z_2}$ και $\mathbf{A}_2$ συμβολίζουμε ένα διάνυσμα κορμό και ένα πίνακα κορμό που αποτελούνται από τα διανύσματα $\mathbf{Z}_{i,k}$ και τους πίνακες $\mathbf{A}_{i,k}$ αντίστοιχα με $i+k>n$. Τότε έχουμε :

$E[\mathbf{Z}_1] = \mathbf{A}_1\boldsymbol{\beta}$

$E[\mathbf{Z}_2] = \mathbf{A}_2\boldsymbol{\beta}$

Συνεπώς, το πολυμεταβλητό βελτιωτικό μοντέλο είναι ένα γραμμικό μοντέλο.

Το ακόλουθο αποτέλεσμα παρέχει έναν τύπο για τους Gauss-Markov εκτιμητές των

μη παρατηρήσιμων αυξητικών ζημιών:

Θεώρημα 2: Για όλα τα i, k $\in$ {0, 1, …, n} τέτοια ώστε i+k>n, ο Gauss-Markov εκτιμητής **Y*(Z**$_{i,k}$) του **Z**$_{i,k}$ να ικανοποιεί τη σχέση :

$$Y*\left(Z_{i,k}\right)=V_i\left(\sum_{j=0}^{n-k}V_j^{1/2}\sum_k^{-1}V_j^{1/2}\right)^{-1}\sum_{j=0}^{n-k}\left(V_j^{1/2}\sum_k^{-1}V_j^{1/2}\right)V_j^{-1}Z_{j,k}$$

Το θεώρημα Gauss-Markov συνεπάγεται ότι οι εκτιμητές Gauss-Markov για το σύνολο των μη παρατηρήσιμων αυξητικών ζημιών ενός δεδομένου έτους ατυχήματος και για το σύνολο όλων των μη παρατηρήσιμων αυξητικών ζημιών επιτυγχάνεται από το σύνολο των Gauss-Markov εκτιμητών για χωριστές μη παρατηρήσιμες ζημιές.

Το θεώρημα Gauss-Markov συνεπάγεται ακόμη ότι δεδομένων ενός έτους ατυχήματος και ενός έτους ανάπτυξης, ο Gauss-Markov εκτιμητής για τις μη πατατηρήσιμες αυξητικές ζημιές **1**΄ **Z**$_{i,k}$ του συνολικού χαρτοφυλακίου είναι ίσες με το ποσό **1**΄**Y*** (**Z**$_{i,k}$) των Gauss-Markov εκτιμητών των μη παρατηρήσιμων αυξητικών ζημιών των υποχαρτοφυλακίων, το **1** συμβολίζει μοναδιαίο διάνυσμα διάστασης m με όλες τις συντεταγμένες ίσες με 1.

Εκτίμηση της διασποράς των παραμέτρων

Στην περίπτωση που m=1, δηλαδή στην μονοπαραγοντική περίπτωση, οι παράμετροι της διασποράς **Σ**$_0$**, Σ**$_1$**, …, Σ**$_n$ αποσύρονται από τους τύπους των εκτιμητών Gauss-Markov.

Στην περίπτωση που m≥2, μόνο η παράμετρος της διασποράς **Σ**$_n$ αποσύρεται από τους τύπους των εκτιμητών Gauss-Markov του μοντέλου, οι παράμετροι της διασποράς **Σ**$_0$**, Σ**$_1$**, …, Σ**$_{n-1}$ πρέπει να εκτιμηθούν.

Για κάθε k $\in$ {0, 1, …, n-1}, το τυχαίο διάνυσμα

$$\tilde{\zeta}_k := \left(\sum_{j=0}^{n-k}V_j\right)^{-1}\sum_{j=0}^{n-k}Z_{j,k}$$

είναι ένας αμερόληπτος γραμμικός εκτιμητής του ζ_k

και ο τυχαίος πίνακας

$$\tilde{\Sigma}_k := \frac{1}{n-k} \sum_{j=0}^{n-k} V_j^{-1/2} \left(Z_{j,k} - V_j \tilde{\zeta}_k \right) \left(Z_{j,k} - V_j \tilde{\zeta}_k \right) V_j^{-1/2}$$

είναι ένας θετικά ημιορισμένος εκτιμητής του θετικά ορισμένου πίνακα Σ_k. Επιπλέον όταν οι πίνακες V_j είναι όλοι διαγώνιοι, τότε όλα τα διαγώνια στοιχεία του τυχαίου πίνακα είναι αμερόληπτοι εκτιμητές των διαγώνιων στοιχείων του Σ_k αλλά τα μη διαγώνια στοιχεία του υποτιμούν τα αντίστοιχα στοιχεία του Σ_k. Η αμεροληψία του εκτιμητή του Σ_k δεν είναι απαραίτητη. Όμως ο εκτιμητής του Σ_k είναι απαραίτητο να είναι θετικά ημιορισμένος .

Στις πρακτικές εφαρμογές είναι αναγκαίο να ελέγχουμε αν οι προτεινόμενοι εκτιμητές των παραμέτρων της διασποράς είναι αντιστρέψιμοι ή όχι, και να τροποποιούμε εκείνους τους εκτιμητές που δεν είναι αντιστρέψιμοι. Άλλος πιθανός τρόπος για την εκτίμηση των παραμέτρων διασποράς συνίσταται στη χρήση εξωτερικών πληροφοριών, που δεν βρίσκονται στο τρίγωνο του μοντέλου και προέρχονται από άλλα τρίγωνα από παρόμοια χαρτοφυλάκια ή από αγορές ή από στατιστικές.

2.7 Εκτίμηση του αποθεματικού ζημιών όταν η σφοδρότητα ζημιών ακολουθεί την Λογαριθμοκανονική κατανομή

Είναι απαραίτητο οι αναλογιστές να εκτιμούν και ένα μέτρο διασποράς του αποθέματος ζημιών. Το μέτρο αυτό διασποράς χρησιμοποιείται ως μέτρο για το πόσο ασφαλής είναι μια μέθοδος πρόβλεψης του αποθεματικού. Τα τελευταία είκοσι χρόνια έχει αναπτυχθεί μια ποικιλία μεθόδων που περιγράφουν πως μπορούν να εκτιμηθούν τα συστηματικά λάθη κατά την εκτίμηση του αποθεματικού. Ο Hayne (1985) παρείχε μια εκτίμηση της στατιστικής διασποράς αναπτύσσοντας την μέθοδο παραγόντων, κάνοντας την υπόθεση ότι οι παράγοντες ακολουθούν την Λογαριθμοκανονική κατανομή. Ο Verall (1991) υπολόγισε αμερόληπτες εκτιμήσεις των συνολικών απαιτήσεων καθώς και των συστηματικών λαθών αυτών των εκτιμήσεων. Ο Mack (1993) υπολόγισε χωρίς τη χρήση κατανομών το συστηματικό λάθος του αποθεματικού που υπολογίστηκε με την τεχνική chain ladder. Οι England και Verall έκαναν μια αναλυτική εκτίμηση του προβλεπόμενου λάθους του αποθεματικού. Ο De Alba (2002) έδωσε μια Μπεϋζιανή προσέγγιση για την προσέγγιση των εκτιμήσεων και κάποιων μέτρων όπως είναι η διασπορά και τα τεταρτημόρια (Han et al., 2008).

Η μέθοδος αυτή είναι μια εναλλαγή της τεχνική chain ladder που δίνει αμερόληπτες εκτιμήσεις για τα αποθεματικά και για τα αντίστοιχα συστηματικά λάθη. Η σφοδρότητα ζημιών αφού υπολογίζεται από την συσσώρευση των ζημιών είναι θετική. Σε κάποιες περιπτώσεις όμως που συμβαίνουν καταστροφικά γεγονότα και απαιτούνται άμεσες πληρωμές η σφοδρότητα είναι μεγάλη και ασταθής.

Υποθέσεις

Έστω L_{ij} η σφοδρότητα ζημιών μεταξύ του χρόνου j και j+1 για το i-οστο χρόνο του ασφαλιστηρίου. Ισχύει:

$$L_{ij}=\frac{c_{i,j+1}}{c_{ij}}, \ i=1, 2,\dots ,m-1 \ \ j=1, 2,\dots , n-1$$

Το C_{ij} αντιπροσωπεύει το δεδουλευμένο ποσό των ζημιών στο j-οστο χρόνο και στο i-οστο ασφαλιστήριο χρόνο[1]. Υποθέτουμε:

1. Τα $L_{1j}, L_{2j}, \ldots, L_{m-1j}$ είναι όμοια κατανεμημένα με το Lj για κάθε j.
2. $L_1, L_2, \ldots, L_{n-1}$ είναι ανεξάρτητα
3. Τα Lj ακολουθούν Λογαριθμοκανονική κατανομή με παραμέτρους μ_j και σ_j^2.

Υποθέτουμε ότι c_{ij} είναι εμπειρικές ζημιές ή πραγματοποιήσεις των C_{ij}. Τότε οι εμπειρικές ζημιές συνοψίζονται σε ένα τρίγωνο ανάπτυξης ζημιών που τα δεδομένα έχουν προσαρμοστεί στον πληθωρισμό και στις πιθανότητες ζημιάς.

Έτος έκδοσης	Έτος εξέλιξης						
	1	**2**	**3**	**…**	**n-2**	**n-1**	**n**
1	C_{11}	C_{12}	C_{13}	…	$C_{1,n-2}$	$C_{1,n-1}$	$C_{1,n}$
2	C_{21}	C_{22}	C_{23}	…	$C_{2,n-2}$	$C_{2,n-1}$	
3	C_{31}	C_{32}	C_{33}	…	$C_{3,n-2}$		
⋮	⋮	⋮	⋮	⋮			
m-2	$C_{m-2,1}$	$C_{m-2,2}$	$C_{m-2,3}$				
m-1	$C_{m-1,1}$	$C_{m-1,2}$					
m	$C_{m,1}$						

Εδώ θεωρούμε ότι το ασφαλιστήριο έτος ισοδυναμεί με ένα ημερολογιακό έτος, θα μπορούσε όμως να ήταν ένα τρίμηνο ή ένας μήνας. Ο σκοπός της αποθεματοποίησης ζημιών είναι να προβλέψουμε τους αριθμούς στο χαμηλότερο τμήμα του τριγώνου.

Αμεροληψία

Αφού έχουμε υποθέσει ότι το L_j ακολουθεί Λογαριθμοκανονική κατανομή με παραμέτρους μ_j και σ_j^2, έχουμε το

$$\theta_j = E[L_j] = \exp\left(\mu_j + \frac{\sigma_j^2}{2}\right)$$

[1] Σαν όρος αφορά το χρονικό διάστημα που ισχύει ένα ασφαλ. συμβόλαιο

Όταν m>n οι εκτιμητές μέγιστης πιθανοφάνειας για τις παραμέτρους μ_j και σ_j^2 είναι:

$$\hat{\mu}_j = \frac{1}{m-j}\sum_{i=1}^{m-j} Y_{ij},\ j = 1,2,...,n-1, \quad \text{και}$$

$$\hat{\sigma}_j^2 = \frac{1}{m-j}\sum_{i=1}^{m-j}\left(Y_{ij} - \hat{\mu}_j\right)^2,\ j = 1,2,...,n-1$$

όπου Y_{ij} =log(L_{ij}), i =1, 2, …, m-1 , j=1, 2, …, min{n-1, m-1}

Τότε η μέγιστη πιθανοφάνεια για την εκτιμήτρια $\hat{\theta}_j$ της θ_j δίνεται από:

$$\hat{\theta}_j = \exp(\hat{\mu}_j + \frac{\hat{\sigma}_j^2}{2})$$

Αυτό αναφέρεται από τον Finney(1941). Περισσότερες λεπτομέρειες μας δίνει ο Verrall (1991).

Για να διορθωθεί η μεροληψία ο Finney (1941) εισήγαγε την συνάρτηση

$$g_k(t) = \sum_{n=0}^{\infty} \frac{k^n (k+2n)}{k(k+2)...(k+2n)} \frac{t^n}{n!}$$

Ο δείκτης κ συμβολίζει τους βαθμούς ελευθερίας που σχετίζονται με το $\hat{\sigma}_j^2$. Σε αυτή την περίπτωση κ=m-j-1. Έτσι :

$$\tilde{\theta}_j = \exp\left(\hat{\mu}_j\right) g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)$$

είναι μια αμερόληπτη εκτίμηση του θ_j με j=1,2,…,n-1

Όμως όταν είμαι m=n και j=n-1 υπάρχει ένα μόνο δεδομένο διαθέσιμο για τον υπολογισμό του $\hat{\mu}_{n-1}$ και $\hat{\sigma}_{n-1}^2$. Οι βαθμοί ελευθερίας που σχετίζονται με το $\hat{\sigma}_{n-1}^2$ είναι μηδέν. Χρησιμοποιούμε την εκτίμηση του Mack (1993) :

$$\hat{\sigma}^2_{n-1} = \min\left\{ \frac{\hat{\sigma}^4_{n-2}}{\hat{\sigma}^2_{n-3}}, \min\left\{ \hat{\sigma}^2_{n-2} / \hat{\sigma}^2_{n-3} \right\} \right\}$$

με $g_0\left(\frac{\hat{\sigma}^2_{n-1}}{2}\right)$ να προσεγγίζεται από το $\exp\left(\frac{\hat{\sigma}^2_j}{2}\right)$

Η διασπορά της $\tilde{\theta}_j$ έστω ότι είναι η τ_j^2 , για την οποία ισχύει:

$$\tau_j^2 = E\left[\tilde{\theta}_j^{\,2}\right] - \left(E\left[\tilde{\theta}_j\right]\right)^2$$

όπου μια αμερόληπτη εκτιμήτρια του $E\left[\tilde{\theta}_j^{\,2}\right]$ δίνεται από τον τύπο:

$$\tilde{\theta}_j^{\,2} = \exp(2\hat{\mu}_j)\left(g_{m-j-1}\left(\frac{\hat{\sigma}^2_j}{2}\right)\right)^2$$

Για να βρούμε μια αμερόληπτη εκτιμήτρια του $\left(E\left[\tilde{\theta}_j\right]\right)^2$:

$$\left(E\left[\tilde{\theta}_j\right]\right)^2 = \left[\exp(\hat{\mu}_j + \frac{\hat{\sigma}^2_j}{2})\right]^2 = \exp(2\mu_j + \sigma_j^2)$$

Σύμφωνα με την αναλογία των Finney(1941) και Verrall (1991) μια αμερόληπτη εκτιμήτρια που προκύπτει για την παραπάνω σχέση είναι :

$$\exp\left(2\hat{\mu}_j\right) g_{m-j-1}\left(\left(1 - \frac{2}{m-j}\right) s_j^2\right)$$

όπου $s_j^2 = \frac{m-1}{m-j-1}\hat{\sigma}^2_j$. Έτσι λοιπόν μια αμερόληπτη εκτιμήτρια για το τ_j^2 δίνεται από τον τύπο :

$$\tilde{\tau}_j^2 = \exp\left(2\hat{\mu}_j\right)\left[\left(g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)\right)^2 - g_{m-j-1}\left(\frac{m-j-2}{m-j-1}\hat{\sigma}_j^2\right)\right]$$

Αν m=n και j=n-1 το $\tilde{\tau}_j^2$ προσεγγίζεται σύμφωνα με τον Finney (1941) από τον τύπο:

$$\exp\left(2\hat{\mu}_j\right)\left[\exp(\sigma_{n-1}^2)(1+\sigma_{n-1}^2) - \exp(\sigma_{n-1}^2)\right]$$

Αμερόληπτη εκτίμηση του αποθεματικού

Έστω

$$L_i^{TE\Lambda} = \prod_{j=m+1-i}^{n-1} L_j \,, \qquad i = 1 ,.., m \qquad , \text{και}$$

$$\theta_i^{TE\Lambda} = E\left[L_i^{TE\Lambda}\right] = \prod_{j=m+1-i}^{n-1} \theta_j \qquad i = 1 ,.., m$$

όπου $L_i^{TE\Lambda}$ είναι η σφοδρότητα ζημιών για το i-στο ασφαλιστήριο έτος στο τελευταίο έτος ανάπτυξης . Τότε η προσδοκώμενη τελική ζημιά για το i-στο ασφαλιστήριο έτος υπολογίζεται από το $c_{i,\ m+1-i}$ πολλαπλασιασμένο με το $\theta_i^{TE\Lambda}$, όπου $c_{i,\ m+1-i}$ η συσσωρευμένη ζημιά για το τρέχον έτος. Υποθέτουμε ότι οι ζημιές που έχουν πληρωθεί για το i-στο ασφαλιστήριο έτος είναι $p_{i,\ m+1-\ i}$. Άρα μια αμερόληπτη εκτίμηση για το αποθεματικό για το i-στο ασφαλιστήριο έτος είναι ίσο με:

$$c_{i,m+1-i} * \prod_{j=m+1-i}^{n-1} \theta_j - p_{i,m+1-i}$$

Άρα μια αμερόληπτη εκτίμηση για το συνολικό αποθεματικό (δηλαδή για όλα τα έτη ατυχημάτων) δίνεται από τον ακόλουθο τύπο:

$$\sum_{i=1}^{m}\left[c_{i,m+1-i} * \prod_{j=m+1-i}^{n-1} \theta_j - p_{i,m+1-i}\right]$$

Αν V_i^2 είναι η διασπορά για το i-οστο χρόνο του ασφαλιστηρίου. Ισχύει:

$$E\left(\prod_{j=m+1-i}^{n-1}\tilde{\theta}_j^2\right)-\left(E\left(\prod_{j=m+1-i}^{n-1}\tilde{\theta}_j\right)\right)^2$$

αφού

$$\prod_{j=m+1-i}^{n-1}\tilde{\theta}_j=\prod_{j=m+1-i}^{n-1}\exp\left(\hat{\mu}_j\right)g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)$$

και

$$\prod_{j=m+1-i}^{n-1}\tilde{\theta}_j^2=\prod_{j=m+1-i}^{n-1}\exp\left(\hat{\mu}_j\right)\left(g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)\right)^2$$

Τότε:

$$V_i^2=c_{i,m+1-i}^2*\exp\left(2\sum_{j=m+1-i}^{n-1}\hat{\mu}_j\right)\left[\prod_{j=m+1-i}^{n-1}\left(g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)\right)^2-\prod_{j=m+1-i}^{n-1}g_{m-j-1}\left(\frac{m-j-2}{m-j-1}\hat{\sigma}_j^2\right)\right]$$

Εκτίμηση της διασποράς του συνολικού αποθεματικού

Ο υπολογισμός της διασποράς του συνολικού αποθεματικού δεν είναι εύκολος για αυτό το λόγο χρησιμοποιούμε μια επαναληπτική διαδικασία. Αν N είναι ο δείκτης που δηλώνει τα χρόνια ανάπτυξης και παίρνει τιμές από 2 μέχρι n, τότε η αμερόληπτη εκτίμηση του συνολικού αποθεματικού για όλα τα ατελή ασφαλιστήρια χρόνια από m-N+2 μέχρι το τέλος του χρόνου ανάπτυξης N δίνεται από τον τύπο:

$$\sum_{i=m-N+2}^{m}\left[c_{i,m+1-i}\prod_{j=m+1-i}^{N-1}\tilde{\theta}_j-p_{i,m+1-i}\right]$$

Επειδή τα ποσά $p_{i,m+1-i}$ είναι σταθερά δεν συμπεριλαμβάνονται στην εκτίμηση της διασποράς. Έστω

$$T_N = \sum_{i=m-N+2}^{m}\left[c_{i,m+1-i}\prod_{j=m+1-i}^{N-1}\tilde{\theta}_j\right]$$

όπου T_N είναι η εκτίμηση για τις πρόσθετες ζημιές από το τέλος του έτους ανάπτυξης N. Τότε, το T_N είναι μια αμερόληπτη εκτίμηση του $E[T_N]$ και T_N^2 χρησιμοποιείται ως αμερόληπτη εκτίμηση του $E[T_N^2]$. Υποθέτουμε ότι $T_N^{(2)}$ ορίζεται ως μια αμερόληπτη εκτίμηση του $(E[T_N])^2$. Μια αμερόληπτη εκτίμηση της διασποράς του T_N δίνεται από τη διαφορά $(T_N)^2 - T_N^{(2)}$.

Για τον επόμενο χρόνο N+1 έχουμε :

$$T_{N+1} = \sum_{i=m-(N+1)+2}^{m}\left(c_{i,m+1-i}\prod_{j=m+1-i}^{(N+1)-1}\tilde{\theta}_j\right) = \left[T_N + c_{m-N+1,N}\right]\tilde{\theta}_N$$

Επομένως

$$E\left[T_{N+1}\right] = E\left[\left(T_N + c_{m-N+1,N}\right)\tilde{\theta}_N\right] = E\left[T_N + c_{m-N+1,N}\right]E\left[\tilde{\theta}_N\right]$$

και

$$\left(E\left[T_{N+1}\right]\right)^2 = \left(E\left[T_N + c_{m-N+1,N}\right]E\left[\tilde{\theta}_N\right]\right)^2 = \left(E\left[T_N + c_{m-N+1,N}\right]\right)^2\left(E\left[\tilde{\theta}_N\right]\right)^2$$

Επειδή ισχύει ότι :

$$\tilde{\theta}_j = \exp\left(\hat{\mu}_j\right)g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)$$

Και επειδή μια αμερόληπτη εκτιμήτρια του $\left(E\left[\tilde{\theta}_j\right]\right)^2$ είναι :

$$\exp\left(2\hat{\mu}_j\right)g_{m-j-1}\left(\left(1-\frac{2}{m-j}\right)s_j^2\right)$$

Καταλήγουμε ότι οι αμερόληπτες εκτιμήσεις της διασποράς της συνολικής ζημιάς από το τέλος ανάπτυξης του έτους N+1 δίνονται από τους ακόλουθους επαναληπτικούς τύπους:

$$(1) T_{N+1} = \left(T_N + c_{m-N+1,N}\right)\exp\left(\widehat{\mu}_N\right) g_{m-N-1}\left(\frac{\widehat{\sigma}_N^2}{2}\right)$$

$$(2) T_{N+1}^{(2)} = \left(T_N^{(2)} + 2c_{m-N+1,N}T_N + c_{m-N+1,N}^2\right)\exp\left(2\widehat{\mu}_N\right) g_{m-N-1}\left(\frac{m-N-2}{m-N-1}\widehat{\sigma}_N^2\right)$$

Οι αρχικές τιμές δίνονται :

N=2

$$T_2 = c_{m,1}\exp\left(\widehat{\mu}_1\right) g_{m-2}\left(\frac{\widehat{\sigma}_1^2}{2}\right)$$

$$T_2^{(2)} = c_{m,1}^2\exp\left(2\widehat{\mu}_1\right) g_{m-2}\left(\frac{m-3}{m-2}\widehat{\sigma}_1^2\right)$$

Αφού T_1=0 και $T_1^{(2)} = 0$. Τα βήματα (1) και (2) επαναλαμβάνονται μέχρι το n να πάρει την τιμή N+1. Τότε μια αμερόληπτη εκτίμηση της διασποράς του συνολικού αποθεματικού είναι ίσο με: $(T_n)^2 - T_n^{(2)}$

Εφαρμογή του μοντέλου Bornhuetter- Ferguson

Η Bornhuetter- Ferguson μέθοδος χρησιμοποιεί συνδυασμό της σφοδρότητας ζημιών και των προσδοκώμενων ζημιών για να υπολογίσει το αποθεματικό. Το βασικό πλεονέκτημα του BF μοντέλου είναι ότι μειώνει την επίδραση των μη προσδοκώμενων ζημιών και έτσι αυξάνεται η σταθερότητα του μοντέλου. Η προτεινόμενη μέθοδος μπορεί να εφαρμοστεί για να εκτιμήσει τα συστηματικά λάθη του αποθεματικού που υπολογίστηκε από το BF μοντέλο.Για τα ασφαλιστήρια χρόνια i=1,2, …, m οι προσδοκώμενες ζημιές από άλλες πηγές είναι EL_i. Τότε το εκτιμώμενο αποθεματικό για το ασφαλιστήριο έτος i θα είναι $EL_i\left(1-\frac{1}{L^{\text{ΤΕΛ}}}\right)$. Επιπλέον το $1/L_j$

ακολουθεί μια Λογαριθμοκανονική κατανομή με παραμέτρους μ_j και σ_j^2 . Ως εκ τούτου :

$$\tilde{\theta}_j^{'} = \exp\left(-\hat{\mu}_j\right) g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right), \text{ j=1, 2, …, n-1}$$

είναι αμερόληπτες εκτιμήσεις του $E[1/L_j]$ j=1, 2, …, n-1. Επιπλέον η κατανομή που ακολουθεί το 1/ $L^{TE\Lambda}$ είναι Λογαριθμοκανονική με παραμέτρους $\sum_{j=m+1-i}^{n-1}\mu_j$ και $\sum_{j=m+1-i}^{n-1}\sigma_j^2$.

Συνεπώς μια αμερόληπτη εκτίμηση για το $E[1/ L^{TE\Lambda}]$ για το i-οστό ασφαλιστήριο έτος είναι $\prod_{j=m+1-i}^{n-1}\tilde{\theta}_j^{'}$. Έτσι μια αμερόληπτη εκτίμηση του αποθέματος για το ασφαλιστήριο έτος i δίνεται από τον τύπο:

$$EL_i\left[1-\exp\left(-\sum_{j=m+1-i}^{n-1}\mu_j\right)\prod_{j=m+1-i}^{n-1} g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right)\right]$$

Ενώ η αντίστοιχη αμερόληπτη εκτίμηση της διασποράς για το ασφαλιστήριο έτος i δίνεται από τον τύπο:

$$\tilde{V}^2 = EL_i^2 \exp\left(-2\sum_{j=m+1-i}^{n-1}\mu_j\right) * \left[\prod_{j=m+1-i}^{n-1} g_{m-j-1}\left(\frac{\hat{\sigma}_j^2}{2}\right) - \prod_{j=m+1-i}^{n-1} g_{m-j-1}\left(\frac{m-j-2}{m-j-1}\hat{\sigma}_j^2\right)\right]$$

Όσο για τη διασπορά του συνολικού αποθεματικού, θεωρούμε ότι ο

$$T_N^{'} = \sum_{i=m-N+2}^{m} EL_i \prod_{j=m+1-i}^{N-1}\tilde{\theta}_j^{'}$$

είναι ο όρος που προκαλεί διασπορά στην εκτίμηση του συνολικού αποθεματικού. Στο τέλος του χρόνου N+1 έχουμε:

$$T_{N+1}^{'} = \sum_{i=m-(N+1)+2}^{m} EL_i \prod_{j=m+1-i}^{(N+1)-1}\tilde{\theta}_j^{'} = \left(T_N^{'} + EL_{m-N+1}\right)\tilde{\theta}_N^{'}$$

Και έτσι προκύπτει ένας επαναληπτικός τύπος.

- Εφαρμόζοντας το μοντέλο αυτό κάνουμε την εξής υπόθεση. Αρχικά θεωρούμε ότι το μέγεθος της διασποράς της σφοδρότητας των ζημιών ελαττώνεται καθώς οι ζημιές προσεγγίζουν την τελευταία τιμή. Αν τα δεδομένα παρουσιάζουν σημαντική μεταβλητότητα προς το τέλος των χρόνων ανάπτυξης, χρειάζονται πιο πολλά δεδομένα για μια ακριβή μέτρηση.
- Χρησιμοποιώντας την μέθοδο αυτή μπορούμε να βρούμε αμερόληπτες εκτιμήσεις για όλα τα αποθεματικά και τα αντίστοιχα συστηματικά λάθη. Επιπλέον αυτή η μέθοδος μπορεί να εφαρμοστεί στην μέθοδο Bornhuetter-Ferguson και τα συστηματικά λάθη της μεθόδου μπορούν να ποσοτικοποιηθούν. Επομένως αυτή η μέθοδος είναι ένα χρήσιμο εργαλείο για την αποθεματοποίηση ζημιών.

2.8 Το μοντέλο Poisson

Το μοντέλο

Έστω C_{ij} οι πρόσθετες ζημιές . Και το σύνολο των ζημιών ορίζεται από το:

$$D_{ij} = \sum_{k=1}^{j} c_{ik}$$

Και

$\{\lambda_j: j=2, \ldots, n\}$

ορίζονται οι παράγοντες ανάπτυξης όπως υπολογίστηκαν από τη μέθοδο «chain ladder».

Θεωρούμε :

C_{ij} ~ ανεξάρτητη κατανομή Poisson

Όπου $E[C_{ij}[= x_i y_j$ και $\sum_{k=1}^{n} y_k = 1$. Προφανώς x_i=E[D_{in}], το οποίο αποτελεί το σύνολο των τελευταίων προσδοκώμενων ζημιών (μέχρι το τελευταίο έτος ανάπτυξης). Οι παράμετροι $\{y_j$: j=1, …, n; $\sum_{k=1}^{n} y_k = 1\}$ θα αναφέρονται ως οι «παράμετροι της στήλης». Οι «παράμετροι της στήλης» θα ερμηνεύονται ως τα μέρη των τελευταίων ζημιών που εμφανίζονται σε κάθε έτος ανάπτυξης. Το μοντέλο αυτό μπορεί να ξαναπαραμετροποιηθεί όπως εμφανίζεται ακολούθως:

C_{ij} ~ ανεξάρτητη κατανομή Poisson

με $E\left[c_{ij}\right] = \dfrac{z_i y_i}{\sum_{k=1}^{n-i+1} y_k}$ και $\sum_{k=1}^{n} y_k = 1$.

Σε αυτήν την περίπτωση z_i= E[$D_{i,n-i+1}$], η οποία είναι η προσδοκώμενη τιμή των ζημιών μέχρι το τελευταίο έτος ανάπτυξης όταν το έτος ατυχήματος είναι το i. Χρησιμοποιώντας την παραμετρικοποίηση, η πιθανοφάνεια είναι:

$$L=\prod_{i=1}^{n}\prod_{j=1}^{n-i+1}\left[\frac{\left(\frac{z_i y_i}{s_{n-i+1}}\right)^{c_{ij}} e^{-z_i y_i / s_{n-i+1}}}{c_{ij}}\right]$$, όπου $S_n=\sum_{k=1}^{m} y_k$

Έτσι

$$L=\prod_{i=1}^{n}\left\{z_i^{\sum_{j=1}^{n-i+1} c_{ij}} \exp\left(-\frac{z_i \sum_{j=1}^{n-i+1} y_j}{S_{n-i+1}}\right)\prod_{j=1}^{n-i+1}\left[\frac{\left(y_j / S_{n-i+1}\right)^{c_{ij}}}{c_{ij}\,!}\right]\right\}$$

$$=\prod_{i=1}^{n}\left\{z_i^{D_{i,n-i+1}} e^{-z_i}\prod_{j=1}^{n-i+1}\left[\frac{\left(y_j / S_{n-i+1}\right)^{c_{ij}}}{c_{ij}\,!}\right]\right\}$$

$$=\prod_{i=1}^{n}\left\{\frac{z_i^{D_{i,n-i+1}} e^{-z_i}}{D_{i,n-i+1}\,!}\left[\frac{D_{i,n-i+1}\,!}{\prod_{i=1}^{n-i+1} c_{ij}\,!}\prod_{j=1}^{n-i+1}\left[\left(y_j / S_{n-i+1}\right)^{c_{ij}}\right]\right]\right\}$$

Το L μπορεί να γραφτεί ως $L=L_D L_C$ όπου:

$$L_D=\prod_{i=1}^{n}\left\{\frac{z_i^{D_{i,n-i+1}} e^{-z_i}}{D_{i,n-i+1}\,!}\right\}$$

Και

$$L_C=\prod_{i=1}^{n}\left[\frac{D_{i,n-i+1}\,!}{\prod_{i=1}^{n-i+1} c_{ij}\,!}\prod_{j=1}^{n-i+1}\left[\left(y_j / S_{n-i+1}\right)^{c_{ij}}\right]\right]$$

Έτσι το L είναι το γινόμενο της πιθανοφάνειας του z_i και της πιθανοφάνειας για $\{y_j : j=1, .., n, \ \sum_{j=1}^{n} y_j=1\}$ Ο πρώτος όρος είναι η πιθανότητα για $D_{i,\ n-i+1}$ και ο δεύτερος όρος είναι η πιθανότητα για $\{C_{ij}$: j=1, …, n-i+1$\}$, υποθέτοντας το $D_{i,\ n-i+1}$. Είναι προφανές ότι η μέγιστη πιθανοφάνεια του z_i εκτιμάται ότι είναι $D_{i,\ n-i+1}$. Αυτό που επίσης είναι ξεκάθαρο είναι ότι η ίδια εκτίμηση για το $\{y_j : j=1, .., n, \ \sum_{j=1}^{n} y_j=1\}$ επιτυγχάνεται όταν L ή L_C μεγιστοποιείται.

Η σύνδεση μεταξύ αυτών των εκτιμήσεων και των εκτιμήσεων των παραγόντων ανάπτυξης στην τεχνική «chain ladder» είναι η ακόλουθη. Αποδεικνύεται ότι οι «παράμετροι της στήλης» $\{y_j : j=1, \ldots, n; \sum_{j=1}^{n} y_j = 1\}$, σχετίζονται με τους παράγοντες ανάπτυξης της τεχνικής «chain ladder» μέσω $\lambda_j = S_j / S_{j-1}$. Έτσι οι εκτιμήσεις μέγιστης πιθανοφάνειας των παραγόντων ανάπτυξης $\hat{\lambda}_j$ μπορεί να επιτευχθούν από τις εκτιμήσεις της μέγιστης πιθανοφάνειας των «παραμέτρων της στήλης» στο πολλαπλασιαστικό μοντέλο $\hat{y}_j$:

$$\hat{\lambda}_j = \frac{\sum_{k=1}^{j} \hat{y}_k}{\sum_{k=1}^{j-1} \hat{y}_k}$$

Αυτό μας δίνει τους ίδιους παράγοντες ανάπτυξης με τη μέθοδο «chain ladder»,

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{i,j-1}}$$

Επομένως, η τεχνική «chain ladder» μπορεί να θεωρηθεί ως ένας αλγόριθμος παραγωγής των εκτιμητών της μέγιστης πιθανοφάνειας είτε για την μη εξαρτημένη πιθανοφάνεια $\{C_{ij}: j=1, \ldots, n-i+1; i=1, \ldots, n\}$, L ή για την εξαρτημένη πιθανοφάνεια, θεωρώντας δεδομένο $\{D_{i,n-i+1}: i=1,\ldots,n\}$, L_C. Αξίζει να σημειωθεί ότι:

1. Δεν είναι αναγκαίο να χρησιμοποιήσουμε την τεχνική «chain ladder» για να υπολογίσουμε τους εκτιμητές των παραγόντων ανάπτυξης, υπάρχουν και άλλοι τρόποι που μπορούν να υπολογιστούν αυτοί οι εκτιμητές.

2. Είναι πιθανό να μπορεί αυτό το μοντέλο να εκτιμήσει τους παράγοντες ανάπτυξης και τις «παραμέτρους της στήλης» ακόμη και όταν κάποιες πρόσθετες ζημιές είναι αρνητικές, αν και κάποια λογιστικά πακέτα αντιμετωπίζουν δυσκολίες σε αυτήν την περίπτωση.

2.9 Αποθεματοποίηση ζημιών με την οικογένεια μοντέλων overdispersed[2] Poisson

Όταν εφαρμόζουμε το μοντέλο Poisson για την αποθεματοποίηση ζημιών, υποθέτουμε πως οι πρόσθετες απαιτήσεις είναι ανεξάρτητες και ακολουθούν την κατανομή Poisson με τις προσδοκίες να εξαρτώνται από δυο παράγοντες, το χρόνο συμβάντος και το χρόνο ανάπτυξης.

Στο overdispersed Poisson μοντέλο με παράμετρο α $\in(0,\infty)$, υποθέτουμε ότι οι πρόσθετες απαιτήσεις είναι ανεξάρτητες και ανήκουν στην εκθετική οικογένεια διασπορών με Var[X]/E[X]=α. Κάθε overdispersed Poisson μοντέλο μπορεί α μετατραπεί στο μοντέλο Poisson διαιρώντας όλες τις πρόσθετες απαιτήσεις με την παράμετρο.

Η οικογένεια εκθετικών διασπορών

Θεωρούμε :

- Ένα σύνολο Borel D $\in$ B(R), ένα ανοιχτό διάστημα $\Theta \subseteq R$και ένα μη κενό σύνολο Φ$\subseteq$ (0, ∞)
- Μία συνάρτηση b: $\Theta \to$R και μία συνάρτηση c: D$\times$Φ→R η οποία είναι μετρήσιμη ως προς την πρώτη μεταβλητή και
- Ένα πεπερασμένο μέτρο ν: B(R) → [0, ∞].

Τις περισσότερες περιπτώσεις το ν είναι είτε το μέτρο Lebesgue ή το μετρήσιμο μέτρο συγκεντρωμένο στο N_0.

Η κατανομή P_X της τυχαίας μεταβλητής X ανήκει στην οικογένεια εκθετικών διασπορών (D, ν, Θ, Φ, b, c), εάν υπάρχει $\theta \in \Theta$ και φ$\in$ Φ που ικανοποιούν τη σχέση

$P_X = \int f_{\theta,\varphi}(x)dv(x)$ που η συνάρτηση $f_{\theta,\varphi}(x): R \to R_+$ δίνεται από τον τύπο

[2] Όταν παρουσιάζονται στα δεδομένα μεγαλύτερη διασπορά από ότι προσδοκούταν.

$$f_{\theta,\varphi}(x) := \begin{cases} \exp\left(\dfrac{\theta x - b(\theta)}{\varphi} + c(x;\varphi)\right), \alpha\nu \ \mathrm{x} \in \mathrm{D} \\ 0 \qquad\qquad\qquad\qquad\qquad ,\text{αλλιώς} \end{cases}$$

Η συνάρτηση $f_{\theta,\varphi}(x)$ ονομάζεται πυκνότητα του X.

Η κανονική οικογένεια εκθετικών διασπορών

Μια οικογένεια εκθετικών διασπορών (D, ν, *Θ*, Φ, b, c) ονομάζεται κανονική αν η συνάρτηση b είναι δυο φορές παραγωγίσιμη με b΄΄(θ)>0 και $\int x^2 f_{\theta,\varphi}(x)dv(x) < \infty$ για όλα τα θ$\in$Θ και φ $\in$Φ.

Λήμμα 2: Υποθέτουμε ότι η εκθετική οικογένεια διασπορών (D, ν, *Θ*, Φ, b, c) είναι μια κανονική εκθετική κατανομή διασπορών. Εάν X μια τυχαία μεταβλητή με πυκνότητα $f_{\theta,\varphi}(x)$ τότε :

E[X]=b**΄(**θ**)** και Var[X]=φ b΄΄(θ)

Η οικογένεια overdispersed[3] Poisson

Μία κανονική οικογένεια εκθετικών διασπορών (D, ν, *Θ*, Φ, b, c) λέγεται οικογένεια overdispersed Poisson με παράμετρο α $\in (0,\infty)$ αν η ταυτότητα :

Var[X]=α E[X]

για κάθε τυχαία μεταβλητή X με $P_X \in$ (D, ν, *Θ*, Φ, b, c)

[3] Όταν παρουσιάζονται στα δεδομένα μεγαλύτερη διασπορά από ότι προσδοκούταν.

Λήμμα 3: Υποθέτουμε ότι (D, ν, *Θ*, Φ, b, c) είναι μια οικογένεια overdispersed Poisson με παράμετρο α. Τότε το Φ είναι ένα μονοσύνολο και υπάρχει β $\in (0,\infty)$ και γ $\in$ R τέτοιο ώστε

$$b(\theta) = \beta \frac{\varphi}{\alpha} e^{\left(\frac{\alpha}{\varphi}\right)^{\theta}} + \gamma$$

για κάθε θ $\in$ Θ. Επιπλέον η συνάρτηση της διασποράς ικανοποιεί τη σχέση :

$$V(\mu) = \frac{\alpha}{\varphi} \mu$$

για όλα τα μ $\in$ b΄(Θ).

Θεώρημα 3: Θεωρούμε α $\in (0,\infty)$. Τότε για τυχαία μεταβλητή X, ισχύει η ακόλουθη ισοδυναμία :

1 Η τυχαία μεταβλητή X / α ακολουθεί κατανομή Poisson

2 Η κατανομή του X ανήκει σε μια οικογένεια overdispersed Poisson με παράμετρο α.

Συνέπεια του παραπάνω θεωρήματος είναι ότι η οικογένεια overdispersed Poisson με παράμετρο α=1 είναι ταυτόσημη με την κατανομή Poisson.

2.10 Μοντέλο Αξιοπιστίας

Το μοντέλο αξιοπιστίας έχει ως στόχο την πρόβλεψη ενός γραμμικού συνδυασμού T των πρόσθετων ζημιών από έναν εκτιμητή της μορφής

$$\widehat{T} = \alpha + \sum_{j=0}^{n} \sum_{l=0}^{n-j} \alpha_{j,l} z_{j,l}$$

Αυτοί οι εκτιμητές λέγονται επιτρεπτοί. Ας σημειώσουμε ότι:

— Οι επιτρεπτοί εκτιμητές δεν είναι αναγκαστικά γραμμικοί ως προς τις παρατηρήσιμες πρόσθετες ζημιές

— και δεν υποθέτουμε ότι οι προβλέψεις είναι αμερόληπτες

Η γενική μορφή του προβλήματος πρόβλεψης είναι λογική αφού αυτό περιλαμβάνει την πρόβλεψη των τελευταίων συνολικά ζημιών $S_{i,n}$ οι οποίες είναι άθροισμα των παρατηρήσιμων πρόσθετων ζημιών $Z_{i,0}$, $Z_{i,1}$, …, $Z_{i,n-i}$ και των μη παρατηρήσιμων πρόσθετων ζημιών $Z_{i,n-i+1}$, …, $Z_{i,n}$.

Για ένα ποσό T των πρόσθετων ζημιών, ένας επιτρεπτός εκτιμητής λέγεται αξιόπιστος εκτιμητής του T αν ελαχιστοποιεί το προσδοκώμενο τετράγωνο του λάθους του εκτιμητή

$$E\left[\left(\hat{T} - T\right)^2\right]$$

για όλους τους αποδεκτούς εκτιμητές $\widehat{T}$.

Ισχύει:

(1) Για κάθε ποσό T πρόσθετων ζημιών, υπάρχει ένας αξιόπιστος εκτιμητής $\widehat{T}^{CR}$ και ο αξιόπιστος εκτιμητής είναι μοναδικός.

(2) Αν T_1 και T_2 είναι ποσά των πρόσθετων ζημιών και εάν c_1 και c_2 είναι πραγματικοί αριθμοί, τότε ο αξιόπιστος εκτιμητής του $T := c_1 T_1 + c_2 T_2$ ικανοποιεί

$$\widehat{T}^{CR} = c_1 \widehat{T}_1^{CR} + c_2 \widehat{T}_2^{CR}$$

Δηλαδή ο αξιόπιστος εκτιμητής είναι γραμμικός.

(3) Αν Τ είναι ένα σύνολο πρόσθετων ζημιών τότε ένας αποδεκτός εκτιμητής $\hat{T}^*$ είναι αξιόπιστος εκτιμητής του Τ αν ικανοποιεί την κανονική ισότητα :

$$E\left[\hat{T}^*\right] = E\left[T\right]$$

και

$$E\left[\hat{T}^* Z_{j,l}\right] = E\left[TZ_{j,l}\right]$$ για όλα τα j, l ϵ{0, 1, …, n} έτσι ώστε j+l≤n

(4) Ο αξιόπιστος εκτιμητής κάποιου ποσού πρόσθετων ζημιών είναι αμερόληπτος.

Εξαιτίας του (2) είναι επαρκές να προσδιορίσουμε τους εκτιμητές αξιοπιστίας των πρόσθετων ζημιών $Z_{i,k}$. Στην περίπτωση που i+k≤n, έχουμε

$$\hat{Z}_{i,k}^{CR} = Z_{i,k}$$

Στην περίπτωση που i+k≥n+1:

$$\hat{Z}_{i,k}^{CR} = \alpha_{i,k} + \sum_{b=0}^{n} \sum_{m=0}^{n-b} \alpha_{i,k,b,m} Z_{b,m}$$

Και προσδιορίζουμε τους συντελεστές από τις κανονικές εξισώσεις

$$E\left[\alpha_{i,k} + \sum_{b=0}^{n} \sum_{m=0}^{n-b} \alpha_{i,k,b,m} Z_{b,m}\right] = E\left[Z_{i,k}\right]$$

και

$$E\left[\alpha_{i,k} + \sum_{b=0}^{n} \sum_{m=0}^{n-b} \alpha_{i,k,b,m} Z_{b,m}\right] = E\left[Z_{i,k} Z_{j,l}\right]$$

τα οποία ισοδύναμα μπορούν να γραφούν:

$$\alpha_{i,k} + \sum_{b=0}^{n} \sum_{m=0}^{n-b} \alpha_{i,k,b,m} E[Z_{b,m}] = E\left[Z_{i,k}\right]$$

και

$$\sum_{b=0}^{n}\sum_{m=0}^{n-b}\alpha_{i,k,b,m}Cov[Z_{b,m},Z_{j,l}]=Cov\left[Z_{i,k}Z_{j,l}\right]$$ για όλα τα j, l ϵ{0, 1, …, n}

έτσι ώστε j+l≤n.

Έτσι παρατηρούμε ότι ο αξιόπιστος εκτιμητής μίας μη παρατηρήσιμης πρόσθετης ζημιάς είναι πλήρως προσδιορισμένος από την πρώτη και δεύτερη στιγμή των πρόσθετων ζημιών. Λύνοντας τις δυο κανονικές εξισώσεις σε δυο βήματα:

- Οι κανονικές εξισώσεις που περιλαμβάνουν τις συνδιακυμάνσεις σχηματίζουν ένα σύστημα γραμμικών εξισώσεων για τους συντελεστές $\alpha_{i,k,b,m}$. Το γεγονός ότι ο αξιόπιστος εκτιμητής υπάρχει συνεπάγεται ότι το σύστημα αυτών των γραμμικών εξισώσεων έχει τουλάχιστον μια λύση.

- Καταχωρώντας κάποια λύση στην κανονική εξίσωση που περιλαμβάνει τις προσδοκίες προκύπτει ο συντελεστής $\alpha_{i,k}$.

Πρέπει να σημειωθεί πως το σύστημα των γραμμικών εξισώσεων ίσως έχει αρκετές λύσεις. Αυτό σημαίνει ότι ο αξιόπιστος εκτιμητής $Z_{i,k}$, ο οποίος είναι γνωστό ότι είναι μοναδικός μπορεί να παρουσιαστεί με αρκετούς τρόπους.

Στα περισσότερα μοντέλα αξιοπιστίας για αποθεματοποίηση ζημιών υποτίθεται ότι δύο πρόσθετες ζημιές από διαφορετικά έτη ατυχήματος είναι ασυσχέτιστες. Σε αυτήν την περίπτωση ο αξιόπιστος εκτιμητής μιας μη παρατηρούμενης πρόσθετης ζημιάς $Z_{i,k}$ μπορεί να γραφεί :

$$\widehat{Z}_{i,k}^{CR}=\alpha_{i,k}+\sum_{m=0}^{n-i}\alpha_{i,k,b,m}Z_{i,m}$$

Και οι συντελεστές μπορούν να προσδιοριστούν από τις μειωμένες κανονικές εξισώσεις

$$\alpha_{i,k}+\sum_{m=0}^{n-i}\alpha_{i,k,b,m}E[Z_{i,m}]=E[Z_{i,k}]$$

και

$$\sum_{m=0}^{n-i}\alpha_{i,k,b,m}Cov[Z_{i,m},Z_{j,l}]=Cov[Z_{i,k},Z_{j,l}]$$ για όλα τα l ϵ{0, 1, …, n-i}.

2.11 Μοντελοποίηση αρνητικών δεδομένων στα στοχαστικά μοντέλα αποθεματοποίησης

Τα δεδομένα των πρόσθετων ζημιών συχνά έχουν αρνητικές τιμές. Στην περίπτωση όμως αυτή δεν ικανοποιούνται κάποιες από τις υποθέσεις των στοχαστικών μοντέλων εξαιτίας των αρνητικών τιμών. Τεχνικές έχουν αναπτυχθεί για την μοντελοποίηση των δεδομένων πρόσθετων ζημιών που περιέχουν αρνητικές τιμές. Για παράδειγμα ο Mack (1994) πρότεινε να μην λαμβάνουν υπόψη τους αυτές τις τιμές στην ανάλυση. Άλλα μοντέλα αντιμετώπισαν το πρόβλημα αυτό με τη χρήση της λύσης της σταθεράς. Σύμφωνα με αυτήν την λύση προστίθεται μία σταθερά σε όλες τις προγενέστερες πρόσθετες ζημιές της ανάλυσης, η οποία τις εξαναγκάζει να οριστούν θετικά. Η ίδια σταθερά από τις προβλέψεις που γίνονται (Kunkler, 2006).

Στο Λογαριθμοκανονικό μοντέλο υποθέτουμε ότι οι πρόσθετες ζημιές είναι θετικά ορισμένες εξαιτίας του λογαριθμικού μετασχηματισμού που γίνεται στα δεδομένα των πρόσθετων ζημιών. Η λύση της σταθεράς είχε τροποποιηθεί για το Λογαριθμοκανονικό χρησιμοποιώντας μια Λογαριθμοκανονική κατανομή τριών παραμέτρων. Οι τρεις παράμετροι αποτελούνται από το μέσο , τη διασπορά και μια σταθερή παράμετρο. Αυτή η προσέγγιση δίνει ικανοποιητικά αποτελέσματα όταν δεν είναι πολλές οι αρνητικές πρόσθετες ζημιές, πρέπει όμως να δοθεί προσοχή στο γεγονός ότι η τελική πρόβλεψη ποικίλει εξαρτώμενη από την τιμή της σταθερής παραμέτρου.

Στα μοντέλα Poisson ο περιορισμός για την μη αρνητικότητα των τιμών είναι λιγότερο αυστηρός, αφού υποθέτει όλες οι τιμές των πρόσθετων ζημιών ότι είναι μη αρνητικές. Η λύση στο πρόβλημα αυτό είναι η χρήση του «overdispersed» Poisson μοντέλου.

Θα παρουσιάσουμε ένα εναλλακτικό μοντέλο για την αντιμετώπιση των περιορισμών ως προς τις αρνητικές τιμές που ανέπτυξε ο Kerner to 2004.

Το εναλλακτικό μοντέλο

Ένας εναλλακτικός τρόπος να μοντελοποιηθούν τα αυθεντικά δεδομένα ζημιών **y**={y_{ij}: i=1, 2, …,$n_α$, j=1, 2, …, $n_α$-i+1} που περιλαμβάνουν και θετικές και αρνητικές τιμές είναι να χωριστούν αρχικά σε δυο σύνολα : το πρώτο σύνολο περιέχει αρνητικές πρόσθετες ζημιές $S^{(-)}$={y_{ij}: y_{ij}<0} και το δεύτερο που περιέχει όλες τις θετικές πρόσθετες ζημιές $S^{(+)}$={y_{ij}: y_{ij}>0}. Για χάρη της πληρότητας , το y_{ij} αντιπροσωπεύει μία παρατηρήσιμη πρόσθετη αξίωση στην i-οστό περίοδο ατυχήματος και τη j-οστό περίοδο ανάπτυξης, το $n_α$ αντιπροσωπεύει τον αριθμό των περιόδων ατυχημάτων και n_d αντιπροσωπεύει τον αριθμό των περιόδων ανάπτυξης. Τα σύνολα $S^{(-)}$ και $S^{(+)}$ δημιουργούνται από ένα μίγμα δεδομένων **z**={ z_{ij}: i=1, 2, …,$n_α$, j=1, 2, …, $n_{α-i+1}$}, το οποίο ορίζεται ως εξής:

$$z_{ij} = \begin{cases} -1 & \alpha\nu \ \mathrm{y}_{ij} < 0 \\ 1 & \alpha\nu \ \mathrm{y}_{ij} > 0 \end{cases}$$

Έτσι οι πιθανές τιμές για κάθε z_{ij} είναι R_z={ $z_{ij,}$ z_{ij}=-1, 1}.

Το μοντέλο αυτό βασίζεται στο μίγμα δεδομένων z, που συνδυάζουν ένα αντανακλώμενο (reflected) γενικευμένο γραμμικό μοντέλο και ένα σύνηθες γενικευμένο γραμμικό μοντέλο. Το αντανακλώμενο γενικευμένο γραμμικό μοντέλο χρησιμοποιείται για την μοντελοποίηση των αρνητικών δεδομένων, ενώ το σύνηθες γενικευμένο γραμμικό μοντέλο για την μοντελοποίηση των θετικών δεδομένων .

Κάποιος μπορεί να μοντελοποιήσει τις απόλυτες τιμές των αρνητικών πρόσθετων ζημιών. Έτσι επιτυγχάνεται η δημιουργία ενός μικτού συνόλου δεδομένων

q= {q_{ij}: i=1, 2, …,$n_α$, j=1, 2, …, $n_{α-i+1}$}, το οποίο ορίζεται από q_{ij} =y_{ij} z_{ij}

Τα μικτά δεδομένα **q** δεν υπόκεινται σε κανένα περιορισμό αφού υποχρεούνται από τον ορισμό τους να είναι θετικά ορισμένα ,

q_{ij}>0, όταν y_{ij}>0, αφού z_{ij}=1

q_{ij}>0, όταν y_{ij}<0, αφού z_{ij}=-1

Ο πολλαπλασιασμός των πρόσθετων ζημιών **y** από το μίγμα δεδομένων **z** πραγματοποιεί ένα βασικό μετασχηματισμό που «αντανακλά» τις αρνητικές τιμές στο **y.**

Όλα όσα αναφέραμε απαιτούνται για την παραγωγή προβλέψεων για τις μελλοντικές πρόσθετες ζημιές δεδομένων με

$\tilde{y}$ ={ $\tilde{y}_{ij}$: i=1, 2, …,n_α , j=1, 2, …, n_α-i+1}

να είναι οι μελλοντικές προβλέψεις και για το μελλοντικό μίγμα δεδομένων

$\tilde{z}$ ={ $\tilde{z}_{ij}$: i=1, 2, …,n_α , j=1, 2, …, n_α-i+1}

και για τα σύνθετα δεδομένα

$\tilde{q}$ ={ $\tilde{q}_{ij}$: i=1, 2, …,n_α , j=1, 2, …, n_α-i+1},

αφού η σχέση για τα μελλοντικά δεδομένα είναι :

$$\tilde{y}_{ij} = \frac{\tilde{q}_{ij}}{\tilde{z}_{ij}}$$

2.12 Μοντελοποιώντας τα μηδενικά δεδομένα στην στοχαστική αποθεματοποίηση ζημιών

Πολλά από τα στοχαστικά μοντέλα, όπως το Λογαριθμοκανονικό μοντέλο, υπόκειντα στον περιορισμό ότι οι αυξητικές ζημιές πρέπει να είναι μεγαλύτερες του μηδενός. Πρακτικά όμως αυτό δεν μπορεί να ισχύει γιατί πολλά δεδομένα είναι μηδενικά ή /και αρνητικά. Παρακάτω θα περιγράψουμε ένα μίγμα μοντέλων που επεκτείνει τα στοχαστικά αποθεματικά μοντέλα και για αυξητικές ζημιές που είναι μηδενικές. Θα επικεντρωθούμε στην εφαρμογή της επέκτασης αυτής στο Λογαριθμοκανονικό μοντέλο, παρόλο που τα πλαίσια του μοντέλου είναι ευέλικτα και για εφαρμογή σε άλλα στοχαστικά μοντέλα αποθεματοποίησης (Kunkler, 2004).

Το πρόβλημα

Έστω y= { y_{ij}: i=1, 2, …, n_a, j=1, 2, …, n_a-i+1 } ότι αντιπροσωπεύει ένα τριγωνικό σχεδιάγραμμα δεδομένων αυξητικών ζημιών, που y_{ij} είναι μια πρόσθετη αξίωση στη i-οστή περίοδο συμβάντος ζημιάς και στη j-οστή περίοδο ανάπτυξης, n_a ο αριθμός των περιόδων συμβάντος ζημιών, n_d ο αριθμός των περιόδων των ετών συμβάντος ζημιών. Είναι τυπικό ότι ο αριθμός των περιόδων ανάπτυξης είναι ίσος με τον αριθμό των περιόδων των ετών ανάπτυξης (n_a= n_d). Θα θεωρήσουμε το σύνολο με τις πιθανές τιμές για κάθε y_{ij} είναι R_y ={ y_{ij}: $0\leq y_{ij}<\infty$ }, το οποίο παρατηρούμε ότι περιλαμβάνει και μηδενικά. Έτσι η υπόθεση του Λογαριθμοκανονικού μοντέλου για θετικές αυξητικές ζημιές ικανοποιείται.

Μία λύση

Μία λύση στο πρόβλημα είναι να χωρίσουμε τα δεδομένα σε δύο σύνολα. Όταν το y_{ij} είναι ίσο με το μηδέν, τοποθετείται στο ένα σύνολο, $S^{(0)}$={ y_{ij}: y_{ij}=0}, και όταν y_{ij} είναι μεγαλύτερα από το μηδέν τοποθετείται στο άλλο σύνολο, $S^{(+)}$={ y_{ij}: y_{ij}>0}. Στο μοντέλο χρησιμοποιούμε επιπλέον το πρόσθετο παρατηρήσιμο δείκτη του τριγωνικού σχεδιαγράμματος **z** ={ z_{ij}: i=1, 2, …, n_a, j=1, 2, …, n_a-i+1 }, που ονομάζεται το μικτό τρίγωνο δεδομένων και ορίζεται ως:

$$z_{ij} = \begin{cases} 0 & ,\alpha\nu\ y_{ij} = 0 \\ 1 & ,\alpha\nu\ y_{ij} > 0 \end{cases}$$

όπου το εύρος των πιθανών τιμών για κάθε z_{ij} είναι $R_Z = \{ z_{ij}: z_{ij}=0,1\}$. Τα y_{ij} τότε μοντελοποιούνται υποθέτοντας τα z_{ij}.

Η πρόβλεψη αρχίζει με την συνδεδεμένη μεταγενέστερη προγνωστική κατανομή, η οποία μπορεί να παραγοντοποιηθεί σε :

$$p\left(\tilde{y}_{ij}, \tilde{z}_{ij} \mid y, z\right) = p\left(\tilde{z}_{ij} \mid z\right) p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij}, y\right)$$

Για το πρόβλημα αποθεματοποίησης ζημιών, το ενδιαφέρον για τα μελλοντικά αυξητικά δεδομένα $\tilde{y}$, όπου $\tilde{z}$ είναι μια ενοχλητική παράμετρος και μπορεί να μετακινηθεί παίρνοντας το μέσο όρο αυτού, ώστε να αποκτήσουμε την οριακή μεταγενέστερη προγνωστική κατανομή που δίνεται από:

$$p\left(\tilde{y}_{ij}, \tilde{z}_{ij} \mid y, z\right) = p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij} = 0, y\right) p\left(\tilde{z}_{ij} = 0 \mid z\right) + p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij} = 1, y\right) p\left(\tilde{z}_{ij} = 1 \mid z\right) =$$
$$= (1-\lambda_j) p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij} = 0, y\right) + \lambda_j p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij} = 1, y\right)$$

όπου οι άγνωστοι παράμετροι του μίγματος $\lambda_j = P(z_{ij}=1)$ αντιπροσωπεύουν την πιθανότητα που σχετίζεται με μια αυξητική αξίωση στην j-οστό περίοδο ανάπτυξης που είναι θετικά ορισμένη και $(1-\lambda_j) = P(z_{ij}=0)$ αντιπροσωπεύουν την πιθανότητα που σχετίζεται με μια αυξητική αξίωση στην j-οστό περίοδο ανάπτυξης που είναι μηδέν με $\lambda_j+(1-\lambda_j)=1$.

Η οριακή μεταγενέστερη προφητική κατανομή για $\tilde{y}_{ij}$ είναι ένα σταθμισμένο σύνολο των δύο συστατικών $p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij} = 0, y\right)$ και $p\left(\tilde{y}_{ij} \mid \tilde{z}_{ij} = 1, y\right)$, με τα βάρη τους να προσδιορίζονται από τα λ_j .

Πράγματι αυτό είναι ένα μικτό μοντέλο που επεκτείνει τα στοχαστικά αποθεματικά μοντέλα για δεδομένα αυξητικών ζημιών που περιέχουν μηδέν.

Η κατανομή για τα z_{ij}, που ονομάζεται δειγματική μικτή κατανομή, είναι διακριτή. Υποτίθεται ότι μέσα στη j-οστή περίοδο ανάπτυξης όπου κάθε z_{ij} είναι ανταλλάξιμα και είναι μια τυχαία ανεξάρτητη μεταβλητή Bernoulli που δίνεται από:

$$p\left(z_{ij} \mid \lambda_j\right) = \lambda_j^{z_{ij}} \left(1-\lambda_j\right)^{1-z_{ij}}$$

Έστω $z_{.j}$ αντιπροσωπεύει των θετικά ορισμένων αυξητικών ζημιών στην περίοδο ανάπτυξης j, όπου j=1, 2, …, n_d , έτσι ώστε

$$z_{.j} = \sum_{i=1}^{n_{a(j)}} z_{ij}$$

όπου $n_{a(j)} = n_a - j + 1$ αντιπροσωπεύει τον αριθμό των περιόδων συμβάντων ζημιών (παρατηρήσεις) στη j-οστή περίοδο ανάπτυξης. Κάθε $z_{.j}$ είναι το σύνολο των $n_{a(j)}$ ανεξάρτητων και ομοιόμορφα κατανεμημένων Bernoulli τυχαίων μεταβλητών, όπου μια τυχαία μεταβλητή Bernoulli δίνεται από

$$p\left(z_{.j} \mid \lambda_j\right) = \binom{n_{a(j)}}{z_{.j}} \lambda_j^{z_{.j}} \left(1-\lambda_j\right)^{n_{a(j)}-z_{.j}} \propto \lambda_j^{z_{.j}} \left(1-\lambda_j\right)^{n_{a(j)}-z_{.j}}$$

Υποτίθεται ότι το εύρος των πιθανών τιμών για κάθε πιθανότητα λ_j είναι $R_\lambda = \{ \lambda_j: 0 < \lambda_j < 1\}$. Πολλές αναλύσεις των πιθανοτήτων χρησιμοποιούν μία συνάρτηση σύνδεσης l(.), να μετατρέψουμε την R_λ σε $R_{l(\lambda)} = \{ \lambda_j: -\infty < l(\lambda_j) < \infty \}$, το οποίο είναι πιο κατάλληλο εφαρμογή για τα γραμμικά μοντέλα.

$$\left(\lambda_j\right) = \begin{cases} \log\left(\dfrac{\lambda_j}{1-\lambda_j}\right), \text{logistic model} \\ \Phi^{-1}\left(\lambda_j\right), \text{Probit model} \\ \log\left(-\log\left(\lambda_j\right)\right), \text{Συμπληρωματικό log - log model} \end{cases}$$

όπου Φ^{-1} είναι η αντίστροφη αθροιστική συνάρτηση κατανομής μιας κανονικής κατανομής. Πρακτικά θα προσδοκούσαμε

$$l^{-1}\left(\left(X_{\delta}\delta\right)_{j}\right)=\begin{cases}\dfrac{\exp\left(\left(X_{\delta}\delta\right)_{j}\right)}{1+\exp\left(\left(X_{\delta}\delta\right)_{j}\right)}, \text{logistic model}\\ \Phi\left(\left(X_{\delta}\delta\right)_{j}\right), \text{Probit model}\\ \exp\left(-\exp\left(\left(X_{\delta}\delta\right)_{j}\right)\right), \text{Συμπληρωματικό log - log model}\end{cases}$$

μια σχέση μεταξύ κάθε λ_j . Στην περίπτωση που μελετάμε έχει επιλεγεί κάθε λ_j να ικανοποιεί τη σχέση:

$$l(\lambda_j) = \sum_{d=0}^{j-1}\delta_d$$

Αν και η παραπάνω σχέση προσφέρει μια ευέλικτη δομή που επιτρέπει στις παραμέτρους δέλτα να μετακινηθούν (αν τους θέσουμε με μηδέν) ή να τους θέσουμε ίσους με την παράμετρο της προηγούμενης αναπτυξιακής περιόδου. Για παράδειγμα το απλό γραμμικό μοντέλο είναι μια ειδική περίπτωση που έχουμε θέσει $\delta_d = \delta_0$ όταν d=0 και $\delta_d = \delta_1$ όταν δ>1, με

$$l(\lambda_j) = \delta_0 + (j-1)\delta_1$$

Το μοντέλο $l(\lambda_j) = \sum_{d=0}^{j-1}\delta_d$ μπορεί να εκφραστεί σε μορφή πινάκων :

$$l\,(\boldsymbol{\lambda}) = \mathbf{X}_{\delta}\boldsymbol{\delta}$$

όπου **λ** είναι ένα $n_d \times 1$ διάνυσμα των μικτών παραμέτρων, l(**λ**) ένα $n_d \times 1$ διάνυσμα των μετασχηματισμένων μικτών παραμέτρων, $\mathbf{X}_{\delta}$ ένα $n_d \times k_d$ γνωστός πίνακας και δ ένα $k_d \times 1$ διάνυσμα παραμέτρων.

Η μεταγενέστερη μικτή κατανομή για **λ** δίνεται από

$$p(\lambda \mid z) \propto \prod_{j=1}^{n_d} p(\lambda_j)\, p(z_{.j} \mid \lambda_j)$$

όπου $\rho(\lambda_j)$ μέσα σε κάθε περίοδο ανάπτυξης j είναι η προγενέστερη μικτή κατανομή για λ_j και $\rho(z_{.j}/\lambda_j)$ είναι η δειγματική μικτή κατανομή όπου όπως έχουμε αναφέρει έχει Διωνυμική κατανομή.

Πρέπει επίσης να σημειωθεί ότι η προγενέστερη μικτή κατανομή μπορεί να γραφτεί σε όρους δ με :

$$p(\delta \mid z) \propto p(\delta) \prod_{j=1}^{n_d} p\left(z_{.j} \mid \delta\right)$$

όπου ρ(δ) είναι η προγενέστερη μικτή κατανομή για δ και $\rho(z_{.j}/\lambda_j)$ η δειγματική μικτή κατανομή μέσα στη j περίοδο ανάπτυξης:

$$p\left(z_{.j} \mid \delta\right) \propto \left(l^{-1}\left(\left(X_\delta \delta\right)_j\right)\right)^{z_{.j}} \left(1 - l^{-1}\left(\left(X_\delta \delta\right)_j\right)\right)^{n_{a(j)} - z_{.j}}$$

Η πιθανότητα εκτίμησης του λ_j μπορεί να υπολογιστεί $l^{-1}\left(\left(X_\delta \delta\right)_j\right)$ από :

$$\widehat{\lambda_j} = l^{-1}\left(\left(X_\delta \delta\right)_j\right)$$

όπου $l^{-1}(\cdot)$ είναι η αντίστροφη της επιλεγμένης συνάρτησης σύνδεσης, έτσι ώστε

$$l^{-1}\left(\left(X_\delta \delta\right)_j\right) = \begin{cases} \dfrac{\exp\left(\left(X_\delta \delta\right)_j\right)}{1 + \exp\left(\left(X_\delta \delta\right)_j\right)}, \text{logistic model} \\ \Phi\left(\left(X_\delta \delta\right)_j\right), \text{Probit model} \\ \exp\left(-\exp\left(\left(X_\delta \delta\right)_j\right)\right), \text{Συμπληρωματικό log - log model} \end{cases}$$

όπου Φ είναι η αθροιστική συνάρτηση της κανονικής κατανομής.

Ο αναλυτής έχει τον πλήρη έλεγχο των προγενέστερων μικτών κατανομών και ο προσδιορισμός μπορεί να επιτευχθεί σε όρους του λ_j ή σε όρους του δ. Σε αυτήν την περίπτωση τα περισσότερα Μπεϋζιανά μοντέλα γενικά έχουν τρεις τύπους προγενέστερων κατανομών από τους οποίους επιλέγουμε. Ο πρώτος τύπος είναι μη πληροφοριακός προγενέστερος που $p(\lambda_j)=1$ ή p(**δ**)=1, τα οποία αναγκαία υποδεικνύουν ότι δεν υπάρχουν προγενέστερες πληροφορίες. Ο δεύτερος τύπος μια συζευγμένη προγενέστερη κατανομή, το οποίο σημαίνει ότι ο τύπος της προγενέστερης κατανομής είναι η ίδια με την μεταγενέστερη κατανομή. Ο τρίτος τύπος είναι μια συζευγμένη προγενέστερη κατανομή, η οποία περιλαμβάνει όλες τις κατανομές που δεν ανήκουν στους δυο πρώτους τύπους.

Με πιθανότητες όπως λ_j είναι ευκολότερο να ασχοληθούμε με συντελεστές όπως είναι το δ. Εδώ υποθέτουμε ότι η προγενέστερη μικτή κατανομή για λ_j είναι μια Βήτα κατανομή που δίνεται από

$$p(\lambda_j) = \frac{\Gamma(a_j + b_j)}{\Gamma(a_j)\Gamma(b_j)} \lambda_j^{a_j - 1} (1-\lambda_j)^{b_j - 1} \propto \lambda_j^{a_j - 1} (1-\lambda_j)^{b_j - 1}$$

όπου α_j, $b_j>0$ και Γ(.) είναι η συνάρτηση Γάμμα που δίνεται από :

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

όπου $x \geq 0$. Μία τυχαία μεταβλητή με κατανομή Βήτα έχει αναμενόμενη τιμή

$$a_j / (a_j + b_j)$$

και διασπορά

$$a_j b_j / \left((a_j + b_j)^2 (a_j + b_j + 1) \right)$$

Αξίζει να σημειωθεί ότι όταν τα α_j, b_j αυξάνονται η διασπορά μειώνεται. Επίσης σημειώστε ότι το λ_j έχει μια ομοιόμορφη κατανομή στο [0, 1] όταν $\alpha_j= b_j=1$.

Αντικαθιστώντας την Βήτα προγενέστερη μικτή κατανομή στη σχέση με τη διωνυμική δειγματική μικτή κατανομή προκύπτει

$$p(\lambda | z) \propto \prod_{j=1}^{n_d} \lambda_j^{a_j - 1} (1-\lambda_j)^{b_j - 1} \lambda_j^{z_{.j}} (1-\lambda_j)^{n_{a(j)} - z_{.j}}$$

$$= \prod_{j=1}^{n_d} \lambda_j^{a_j + z_{.j} - 1} (1-\lambda_j)^{b_j + n_{a(j)} - z_{.j} - 1}$$

$$= \prod_{j=1}^{n_d} \lambda_j^{a_j^* - 1} (1-\lambda_j)^{b_j^* - 1}$$

όπου την j-οστή περίοδο ανάπτυξης η μορφή είναι της κατανομής Βήτα με παραμέτρους

$$a_j^* = a_j + z_{.j}$$

και

$$b_j^* = b_j + n_{a(j)} - z_{.j}$$

Έτσι η κατανομή Βήτα είναι συνδυαστική προγενέστερη κατανομή για την Διωνυμική κατανομή. Η Βήτα προγενέστερη κατανομή στη j-οστή περίοδο ανάπτυξης προσθέτει α_j-1 επιπλέον παρατηρήσεις με τον αριθμό 1 και b_j-1 επιπλέον παρατηρήσεις με μηδενικά. Για τριγωνικά σχεδιαγράμματα ανάπτυξης ζημιών, το δείγμα αλλάζει σε κάθε περίοδο ανάπτυξης κα τείνει να είναι μικρό στις τελευταίες περιόδους ανάπτυξης. Στην Βήτα προγενέστερη κατανομή, ο αναλυτής πρέπει να μειώνει και το α_j και το b_j στην ουρά των τριγωνικών δεδομένων έτσι ώστε οι προγενέστερες κατανομές να μην έχουν σημαντική επίδραση στην μεταγενέστερη κατανομή. Περισσότερες λεπτομέρειες μπορούν να βρεθούν στο Gelman et al. (1995).

Η δειγματική κοινή κατανομή για y_{ij} και z_{ij} μπορούν να παραγοντοποιηθούν :

$$p\left(y_{ij}, z_{ij} \mid \theta, \lambda_j\right) = p\left(z_{ij} \mid \lambda_j\right) p\left(y_{ij} \mid z_{ij}, \theta\right)$$

όπου $p(z_{ij} \mid \lambda_j)$ είναι η δειγματική μικτή κατανομή και $p(y_{ij} \mid z_{ij}, \boldsymbol{\theta})$ η υποθετική δειγματική κατανομή για y_{ij}. Το z_{ij} είναι μια βοηθητική παράμετρος και μπορεί να μετακινηθεί παίρνοντας μέσο όρο γύρω από αυτό και αποκτεί την οριακή δειγματική κατανομή

$$p\left(y_{ij}, z_{ij} \mid \theta, \lambda_j\right) = \Pr\left(z_{ij} = 0 \mid \lambda_j\right) p\left(y_{ij} \mid z_{ij} = 0, \theta_0\right) + \Pr\left(z_{ij} = 1 \mid \lambda_j\right) p\left(y_{ij} \mid z_{ij} = 1, \theta_1\right) =$$
$$= \left(1 - \lambda_j\right) p\left(y_{ij} \mid z_{ij} = 0, \theta_0\right) + \lambda_j p\left(y_{ij} \mid z_{ij} = 1, \theta_1\right)$$

Η οριακή δειγματική κατανομή για y_{ij} είναι ένα σταθμισμένο σύνολο δύο δειγματικών κατανομών $p(y_{ij} \mid z_{ij}=0 , \boldsymbol{\theta_0})$ και και $p(y_{ij} \mid z_{ij}=1 , \boldsymbol{\theta_1})$ με τα βάρη να προσδιορίζονται από λ_j. Το τριγωνικό σχεδιάγραμμα που αντιπροσωπεύει τα αυξητικά δεδομένα ζημιών χωρίζεται σε δυο σύνολα ($S^{(0)}$ και $S^{(+)}$), τα οποία μοντελοποιούνται χωριστά . Μια συνάρτηση δείκτης $I_{(A)}$ ορίζεται ως $I_{(A)}=1$ όταν ικανοποιείται η συνθήκη A και $I_{(A)}=0$ αλλιώς.

Η οριακή δειγματική κατανομή για y_{ij} μπορεί να ξαναγραφεί ως

$$p\left(y_{ij} \mid \theta, \lambda_j\right) = \left(1-\lambda_j\right) I_{(z_{ij=0})} + \lambda_j p\left(y_{ij} \mid z_{ij} = 1, \theta_1\right)$$

όπου ο δείκτης για το θ στο δεύτερο συστατικό της δειγματικής κατανομής $p(y_{ij} \mid z_{ij} = 1, \boldsymbol{\theta_1})$ μπορεί τώρα να παραληφθεί αφού δεν υπάρχει ευκαιρία για ασάφεια.

Όταν λ_j είναι μικρότερο του 0, το μικτό μοντέλο θα παράγει μικρότερες προβλέψεις από ένα μοντέλο που οι μηδενικές πρόσθετες ζημιές μετακινήθηκαν όλες μαζί από την ανάλυση.

Το εύρος των πιθανών τιμών για y_{ij} στο δεύτερο μέρος της δειγματικής κατανομής όπως φαίνεται στην τελευταία σχέση είναι θετικά ορισμένη. Πράγματι το Λογαριθμοκανονικό μοντέλο μπορεί τώρα να εφαρμοστεί αφού του μοντέλου η θετικά ορισμένη υπόθεση ικανοποιείται.

Για να απλοποιήσουμε τη σημείωση αυτή, έστω y^+ αντιπροσωπεύει τα μη συμπληρωμένα δεδομένα ζημιών του τριγωνικού σχεδιαγράμματος που αποτελείται από αυξητικές ζημιές που είναι θετικές, $y_{ij} \in S^{(+)}$. Το δεύτερο μέρος της δειγματικής κατανομής μπορεί να εκφραστεί σε όρους $\mathbf{y}^+$ αφού $p(\mathbf{y}^+|\boldsymbol{\theta}) \equiv p(y_{ij}|z_{ij}=1, \boldsymbol{\theta})$.

Υποτίθεται ότ $\mathbf{y}^+$ είναι μια τυχαία μεταβλητή με μια Λογαριθμοκανονική δειγματική κατανομή να δίνεται από

$$\log(y^+) \mid \theta \sim N\left(X_\beta \beta, \sigma^2 I\right)$$

όπου $\boldsymbol{\theta}=[\boldsymbol{\beta}, \sigma^2, \mathbf{I}]$ ένας $n_{y|z} \times n_{y|z}$ ταυτοτικός πίνακας,

$$n_{y/z} = \sum_{i=1}^{n_a} \sum_{j=1}^{n_a - i + 1} I_{\left(z_{ij}=1\right)}$$

ο αριθμός των y_{ij} που περιλαμβάνονται στο Λογαριθμοκανονικό μοντέλο, β ένα $k_\beta \times 1$ διάνυσμα και $\mathbf{X}_\beta$ ένας $n_{y|z} \times k_\beta$ γνωστός πίνακας ο οποίος δεν είναι τυχαίος και έχει κατασταλεί από τις υποθετικές καταστάσεις. Ο τύπος του $\mathbf{X}_\beta\beta$ δίνεται από

$$\left(X_\beta \beta\right)_{ij} = a_i + \sum_{d=1}^{j-1} \gamma_d + \sum_{t=1}^{i+j-2} l_t$$

Αν και η παραπάνω σχέση είναι υπερπαραμετρικοποιημένη, αυτή παρέχει μια ευέλικτη δομή και επιτρέπει στις παραμέτρους να μετακινηθούν (θέτοντας με μηδέν) ή θέτοντας ίσες με την προηγούμενη παράμετρο. Η κοινή προγενέστερη κατανομή για β και σ^2 μπορεί να παραγοντοποιηθεί σε

$$p\left(\beta,\sigma^2 \mid y^+\right) = p\left(\beta \mid \sigma^2, y^+\right) p\left(\sigma^2 \mid y^+\right)$$

Όπου η υποθετική προγενέστερη κατανομή για το **β** είναι

$$\beta,\sigma^2 \mid y^+ \sim N\left(\widehat{\beta},\left(X_\beta^T X_\beta\right)^{-1}\sigma^2\right)$$

Και η οριακή προγενέστερη κατανομή για σ^2 είναι

$$\sigma^2 \mid y^+ \sim Inv - x^2\left(n_{y|z} - k_\beta, s^2\right)$$

Οι εκτιμητές για το **β** και σ^2 υπολογίζονται από

$$\widehat{\beta} = \left(X_\beta^T X_\beta\right)^{-1} X_\beta^T \log\left(y^+\right)$$

και

$$s^2 = \frac{1}{n_{y|z} - k_\beta}\left(\log\left(y^+\right) - X_\beta\widehat{\beta}\right)^T \left(\log\left(y^+\right) - X_\beta\widehat{\beta}\right)$$

αντίστοιχα. Υποτίθεται επίσης ότι p(**β,** σ^2) έχει ένα μη πληροφοριακό παρελθόν που δίνεται από

$$p\left(\beta,\sigma^2\right) \propto \left(\sigma^2\right)^{-1}$$

όπου η εναλλακτική προγενέστερη κατανομή μπορεί να βρεθεί στο Gelman et al. (1995).

Η υιοθέτηση μιας Μπεϋζιανής προσέγγισης επιτρέπει σε έναν αναλυτή να ενσωματώσει την αβεβαιότητα στην παράμετρο της διασποράς σ^2, η οποία υποτίθεται ότι έχει μια κλιμακωτή αντίστροφη κατανομή Χ^2. Από μια μη μπεϋζιανή οπτική η οριακή προγενέστερη κατανομή για σ^2 θα είναι ρ(σ^2|**y**$^+$)=1 έτσι ώστε μόνο η αβεβαιότητα στο **β** να μοντελοποιείται.

3. Όρια εμπιστοσύνης για προεξοφλημένα αποθέματα ζημιών

Επειδή τα προεξοφλημένα αποθέματα IBNR είναι ένα σύνολο εξαρτημένων Λογαριθμοκανονικών τυχαίων μεταβλητών, η συνάρτηση κατανομής δεν μπορεί να προσδιοριστεί αναλυτικά. Επομένως, αντί να υπολογιστεί η ακριβής κατανομή, θα αναζητήσουμε όρια κατά την έννοια του «πιο ευνοϊκού/λιγότερο επικίνδυνου» και «λιγότερο ευνοϊκού /περισσότερο επικίνδυνου», με μια πιο απλή δομή. Αυτή η τεχνική χρησιμοποιείται από την αναλογιστική βιβλιογραφία. Όταν το κατώτερο και ανώτερο όριο είναι κοντά το ένα στο άλλο, μπορούν να παρέχουν αξιόπιστες πληροφορίες για την αυθεντική και πιο περίπλοκη μεταβλητή. Η έννοια «λιγότερο ευνοϊκή» ή «περισσότερο επικίνδυνη» μεταβλητή θα οριστεί με τη βοήθεια της κυρτής διάταξης (Hoedemakers, 2003).

Ορισμός 1: Μία τυχαία μεταβλητή V είναι μικρότερη από μια τυχαία μεταβλητή W σε κυρτή διάταξη αν E[u(V)]≤ E[u(W)] για όλες τις κυρτές συναρτήσεις u: R→R: x→u(x), υπό την προϋπόθεση ότι οι προσδοκίες υπάρχουν.

Αυτό συμβολίζεται ως :

$$V \leq_{cx} W$$

Όταν λέμε

$$V \leq_{cx} W$$

αυτό σημαίνει ότι το W είναι πιο πιθανό να πάρει ακραίες τιμές από το V. Σε όρους θεωρίας χρησιμότητας αυτό σημαίνει ότι η ζημιά V είναι προτιμότερη από τη ζημιά W, αυτό συνεπάγεται Var[V]≤ Var[W].

Θεώρημα 4: Θεωρούμε ένα αυθαίρετο άθροισμα τυχαίων μεταβλητών :

$V=X_1+X_2+\ldots+X_n$ και ορίζουμε τις σχετικές στοχαστικές ποσότητες

$$V_1 = E\left[X_1 \mid Z\right]+E\left[X_2 \mid Z\right]+\ldots+E\left[X_n \mid Z\right],$$
$$V_u^{'} = F_{X_1|Z}^{-1}\left(U\right)+F_{X_2|Z}^{-1}\left(U\right)+\ldots+F_{X_n|Z}^{-1}\left(U\right),$$
$$V_u = F_{X_1}^{-1}\left(U\right)+F_{X_2}^{-1}\left(U\right)+\ldots+F_{X_n}^{-1}\left(U\right)$$

με U μια ομοιόμορφη τυχαία κατανομή (0, 1) και με Z μια αυθαίρετη τυχαία μεταβλητή ανεξάρτητη της U. Η ακόλουθη σχέση τότε ισχύει:

$$V_1 \le_{cx} V \le_{cx} V_u^{'} \le_{cx} V_u$$

Για κάθε j=1, …,n οι όροι στην αρχική μεταβλητή V και οι αντίστοιχοι όροι στα άνω όρια V_u και V_u' είναι όμοια κατανεμημένοι :

$$X_j \stackrel{d}{=} F_{X_j}^{-1}\left(U\right) \stackrel{d}{=} F_{X_j|Z}^{-1}\left(U\right)$$

Για τα χαμηλότερα όρια, οι ισότητες των κατανομών X_j και $E[X_j|Z]$ μόνο ισχύει στην περίπτωση που όλα τα X_j, δεδομένου Z=z, είναι σταθερά για κάθε z.

Αυτά τα αποτελέσματα μπορούν να γενικευτούν στην περίπτωση που V αποτελείται από ένα άθροισμα μονότονων συναρτήσεων φ_j τυχαίων μεταβλητών X_j, απλά αντικαθιστώντας το Y_j για $\varphi_j(X_j)$ και εφαρμόζοντας το Θεώρημα 4.

Θεώρημα 5: Υποθέτουμε ότι τα διανύσματα **X** και **Y,** δεδομένου της τυχαίας μεταβλητής Z, είναι αμοιβαία ανεξάρτητες και το Z ανεξάρτητο του Y. Θεωρείτε δύο ανεξάρτητες ομοιόμορφα κατανεμημένες τυχαίες μεταβλητές U, V στο (0, 1). Αν οι X_i και Y_i είναι μη αρνητικές τυχαίες μεταβλητές, τότε βρίσκουμε ότι ισχύουν οι ακόλουθες σχέσεις :

$$W_1 \leq_{cx} W \leq_{cx} W_u^{'} \leq_{cx} W_u$$

με

$$\begin{aligned}
&W = X_1Y_1 + X_2Y_2 + \ldots + X_nY_n \\
&W_1 = E\left[X_1 \mid Z\right]E\left[Y_1\right] + E\left[X_2 \mid Z\right]E\left[Y_2\right] + \ldots + E\left[X_n \mid Z\right]E\left[Y_n\right], \\
&W_u^{'} = F_{X_1|Z}^{-1}\left(U\right)F_{Y_1}^{-1}\left(V\right) + F_{X_2|Z}^{-1}\left(U\right)F_{Y_2}^{-1}\left(V\right) + \ldots + F_{X_n|Z}^{-1}\left(U\right)F_{Y_n}^{-1}\left(V\right), \\
&W_u = F_{X_1}^{-1}\left(U\right)F_{Y_1}^{-1}\left(V\right) + F_{X_2}^{-1}\left(U\right)F_{Y_2}^{-1}\left(V\right) + \ldots + F_{X_n}^{-1}\left(U\right)F_{Y_2}^{-1}\left(V\right)
\end{aligned}$$

και όπου U, V και Z είναι αμοιβαίως ανεξάρτητα.

Από το παραπάνω θεώρημα τα προηγούμενα αποτελέσματα επεκτείνονται από τα συνήθη ποσά των μεταβλητών στα ποσά των βαθμωτών προϊόντων των ανεξάρτητων τυχαίων μεταβλητών.

Ανώτερα και κατώτερα όρια για το προεξοφλημένο IBNR απόθεμα

Όπως έχουμε περιγράψει στο προηγούμενο θεώρημα για το προεξοφλημένο IBNR απόθεμα:

$$S = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} e^{(\overrightarrow{\hat{B}})_{ij} - Y(i+j-t-1) + \in_{ij}}$$

Εισάγουμε τις τυχαίες μεταβλητές Vij και W_{ij} να ορίζονται από

$$W_{ij} = \left(B\ \overrightarrow{\hat{\beta}}\right)_{ij} - Y\left(i+j-t-1\right), \qquad \mathrm{V}_{ij} = e^{W_{ij}}$$

Υποθέτουμε τώρα μια υποθετική κανονικά κατανεμημένη τυχαία μεταβλητή Ζ οριζόμενη ως ακολούθως:

$$Z = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} v_{ij} Y_{i+j-t-1}$$

Θα υπολογίσουμε το κατώτατο και ανώτερο όριο για την ακόλουθη επιλογή των παραμέτρων:

$$v_{ij} = \sum_{k=i+1}^{t} \sum_{l=t+2-k}^{t} \exp\left((B\overrightarrow{\beta})_{kl} - \left(k+l-t-1\right)\left(\mu + \frac{1}{2}\delta^2\right)\right)$$
$$+ \sum_{l=j}^{t} \exp\left((B\overrightarrow{\beta})_{ij} - \left(i+l-t-1\right)\left(\mu + \frac{1}{2}\delta^2\right)\right)$$

Ζ είναι ένας γραμμικός μετασχηματισμός της πρώτης τάξης προσέγγισης του

$$\tilde{S} = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} e^{(B\overrightarrow{\hat{\beta}})_{ij} - Y(i+j-t-1) + \epsilon_{ij}}$$

Για μια πολυμεταβλητή κανονική κατανομή, κάθε γραμμικός συνδυασμός των συστατικών του έχει μια μονοπαραγοντική κανονική κατανομή, έτσι Ζ είναι κανονικά κατανεμημένο. Επίσης, (W_{ij}, Z) έχει μια διωνυμική κανονική κατανομή. Υποθέτοντας δεδομένου Z=z, W_{ij} έχει μια μονοπαραγοντική κανονική κατανομή με μέσο και διασπορά να δίνονται από

$$E\left[W_{ij} \mid Z = z\right] = E\left[W_{ij}\right] + \rho_{ij} \frac{\sigma w_{ij}}{\sigma z}\left(z - E[Z]\right)$$

Και

$$Var\left[W_{ij} \mid Z = z\right] = \sigma w_{ij}^2 \left(1 - \rho_{ij}^2\right)$$

Όπου ρ_{ij} δηλώνει την συσχέτιση μεταξύ Ζ και W_{ij}.

Πρόταση 2: Έστω S, $S_{l,}$ S_u΄ και S_u ορίζονται όπως ακολούθως:

$$S = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \exp\left(W_{ij} + \in_{ij}\right),$$

$$S_1 = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \exp\left(E[W_{ij}] + \rho_{ij}\sigma w_{ij}\Phi^{-1}\left(U\right) + \frac{1}{2}\left(1-\rho_{ij}^{2}\right)\sigma w_{ij}^{2} + \frac{1}{2}\sigma_{ij}^{2} \right),$$

$$S_u^{'} = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \exp\left(E[W_{ij}] + \rho_{ij}\sigma w_{ij}\Phi^{-1}\left(U\right) + \sqrt{1-\rho_{ij}^{2}}\sigma w_{ij}\Phi^{-1}\left(W\right) + \sigma_{\in_{ij}}\Phi^{-1}\left(V\right)\right),$$

$$S_u = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \exp\left(E[W_{ij}] + \sigma w_{ij}\Phi^{-1}\left(U\right) + \sigma_{\in_{ij}}\Phi^{-1}\left(V\right)\right)$$

Όπου U, V και W εείναι αμοιβαίως ανεξάρτητες ομοιόμορφες κατανεμημένες στο (0, 1) τυχαίες μεταβλητές και Φ είναι η αθροιστική συνάρτηση πιθανότητας της N(0, 1) κατανομής. Τότε έχουμε:

$$S_1 \leq_{cx} S \leq_{cx} S_u^{'} \leq_{cx} S_u$$

4. Επιλογή μοντέλου και εκτίμηση προβλέψεων

Στην ερώτηση ποιο είναι το καλύτερο μοντέλο, δεν απαιτείται να απαντήσουμε κάποιο συγκεκριμένο, αλλά μπορούμε να επιλέξουμε ένα σύνολο μοντέλων και να χρησιμοποιούμε ένα γραμμικό συνδυασμό αυτών.

Αρχίζουμε με ένα σύνολο επεξηγηματικών μεταβλητών, αυτές μπορεί να είναι είτε στοχαστικές, είτε ντετερμινιστικές. Στην περίπτωση των στοχαστικών εμείς ίσως ανατρέξουμε στις παρελθοντικές παρατηρήσεις. Ένας τρόπος είναι να τρέξουμε όλες τις δυνατές παλινδρομήσεις (2^r παλινδρομήσεις – για r μεταβλητές), αυτή η διαδικασία όμως δεν είναι βολική για μεγάλο αριθμό μεταβλητών.

Ο πιο απλός τρόπος είναι η σταδιακή διαδικασία παλινδρόμησης, η οποία εισάγει όλες τις παλινδρομούσες μεταβλητές βήμα προς βήμα, αρχίζοντας με αυτή που δίνει την καλύτερη προσαρμογή, έπειτα προσθέτουμε εκείνη που δίνει την καλύτερη βελτίωση για το R^2 κ.ο.κ. Η διαδικασία σταματάει όταν καμία επιπλέον μεταβλητή δεν αποφέρει σημαντική βελτίωση. Το πρόβλημα είναι ότι σε αυτήν την περίπτωση δεν εφαρμόζεται της t-κατανομής και της F οι πίνακες και αυτό έχει ως συνέπεια να πάρουμε λάθος αποφάσεις.

Η προς τα πίσω διαδικασία αποκλεισμού αρχίζει με το μεγαλύτερο μοντέλο και σε κάθε βήμα χάνει μια μεταβλητή της οποίας η κατανομή έχει την μικρότερη απόλυτη τιμή της t-στατιστικής. Με μια τέτοια διαδικασία ο χρήστης έχει μικρότερο κίνδυνο χρησιμοποιώντας t στατιστική. Αν κάποιος χρησιμοποιεί επίπεδο σημαντικότητας το ποσοστό 100α η συνολική μετά από p τεστ θα είναι ίση με $1-(1-\alpha)^p$. Ένα μειονέκτημα αυτής της διαδικασίας είναι ότι στα πολυσυγγραμμικά προβλήματα στα αρχικά βήματα, σε σχέση με το Χ΄Χ που δεν αντιστρέφεται. Μια καλή λύση είναι να αρχίσουμε με τη διαδικασία της σταδιακής παλινδρόμησης που συζητήθηκε παραπάνω και να κάνουμε ένα μεγάλο αριθμό από τεστ σε κάθε βήμα.

Έπειτα πρέπει να αποφασίσουμε πιο κριτήριο θα χρησιμοποιήσουμε. Η πιθανότητα του να χρησιμοποιήσουμε F-test ή t-test έχει ήδη αναφερθεί. Τώρα θα συνεχίσουμε με έξω από το δείγμα πρόβλεψη είναι ο στόχος. Ένα κριτήριο που παράγεται από το στόχο αυτό είναι το C_p, γνωστό και ως Mallow κριτήριο, MC,

$$MC = C_p = \frac{RSS_P}{s^2} - \left(N - 2p\right)$$

όπου RSS_P είναι το άθροισμα των τετραγώνων των υπολοίπων που περιλαμβάνει p παραμέτρους. N είναι ο συνολικός αριθμός των παρατηρήσεων και S^2 είναι ίσο με RSS/(N-Q) , με Q τον αριθμό των ερμηνευτικών μεταβλητών.

Τέλος, σύντομα αναφέρουμε ότι υπάρχει μια σχολή από αυτούς που υποστηρίζουν ότι η επιλογή του μοντέλου είναι η επιλογή του βέλτιστου. Οι ποσοστώσεις γύρω από το γραμμικό πρέπει να προστεθούν γιατί τα βάρη εξαρτώνται από τις παρατηρήσεις. Ένα σύνολο μοντέλων έχει μικρότερο κίνδυνο από ότι μόνα τους τα μοντέλα.

5. Συμπεράσματα

Από την μελέτη των βασικών στοχαστικών μοντέλων που έχουν αναπτυχθεί για την αποθεματοποίηση ζημιών καταλήγουμε στα εξής συμπεράσματα που αφορούν σημεία που πλεονεκτούν, αλλά και μειονεκτούν αυτές οι μέθοδοι, καθώς επίσης και σημεία που διαφέρουν ή μοιάζουν.

- Στην ερώτηση «ποιο μοντέλο είναι το καλύτερο;» δεν υπάρχει καμία άμεση απάντηση. Θα πρέπει τα δεδομένα να εξεταστούν με λεπτομέρεια ώστε να βρεθεί το πιο κατάλληλο μοντέλο.

- Οι στοχαστικές μέθοδοι θεωρείται ότι χρησιμοποιούνται στις περιπτώσεις που αποτυγχάνουν οι ντετερμινιστικές. Στην πραγματικότητα όμως η προσφορά των στοχαστικών είναι μεγαλύτερη αφού μπορούν να μας παρέχουν πληροφορίες που είναι πολύ χρήσιμες στην διαδικασία αποθεματοποίησης και γενικότερα στην διοίκηση της επιχείρησης.

- Με την μέθοδο «chain ladder» μπορούμε να κάνουμε πρόβλεψη για τις ζημιές, όμως κάθε πρόβλεψη πρέπει να συνοδεύεται και από τον υπολογισμό της διασποράς που είναι η εκτίμηση του λάθους πρόβλεψης. Αυτή η ανάγκη οδήγησε στην ανάπτυξη στοχαστικών μεθόδων που δίνουν ίδιες προβλέψεις με την τεχνική της «chain ladder», αλλά επιπλέον παρέχουν και ένα μέτρο για το λάθος της εκτίμησης. Η εκτίμηση του λάθους μπορεί να γίνει και τον προσδιορισμό της κατανομής που ακολουθούν τα δεδομένα (όπως στη μέθοδο overdispersed Poisson) ή με τον προσδιορισμό μόνο της κατανομής των δύο πρώτων στιγμών (μέθοδος Mack).

- Παρόλο που το μοντέλα Poisson (ή overdispersed Poisson) και το μοντέλο του Mack ξαναπαράγουν τους ιστορικούς εκτιμητές της αποθεματοποίησης ζημιών του μοντέλου «chain ladder», διαφέρουν μεταξύ τους. Για παράδειγμα τα πραγματικά προσδοκώμενα αποθέματα ζημιών που περιγράφονται από τα δυο μοντέλα είναι διαφορετικά, όπως επίσης και το γεγονός ότι το μοντέλο

Poisson αποκλίνει από τον ιστορικό αλγόριθμο «chain ladder» σε αρκετές πλευρές γεγονός που δεν ισχύει για το μοντέλο του Mack.

- Τα Μπεϋζιανά μοντέλα πλεονεκτούν στο γεγονός ότι με τη χρήση της κατανομής πρόγνωσης των αποθεματικών εκτιμήσεων μπορεί να υπολογιστεί και το τυπικό λάθος της πρόβλεψης με τον υπολογισμό απλά της τυπικής απόκλισης των εκτιμήσεων του αποθέματος, χωρίς τη χρήση πολύπλοκων τύπων. Τα Μπεϋζιανά μοντέλα έχουν επίσης το πλεονέκτημα ότι η αναλογιστική κρίση μπορεί να αναγνωριστεί μέσα από την επιλογή των προγενέστερων πληροφοριακών κατανομών. Αυτό είναι επίσης ένα μεγάλο μειονέκτημα αφού επιτρέπει την κατάχρηση μεθόδων. Πρακτικές δυσκολίες σχετιζόμενες με τα Μπεϋζιανά μοντέλα περιλαμβάνουν την επιλογή κατανομής και ασφάλειας ώστε το λογισμικό να συγκλίνει στη βέλτιστη λύση. Μειονεκτεί επίσης επειδή δεν μπορεί να εφαρμοστεί από το Excel, το οποίο το χρησιμοποιούν οι περισσότεροι επαγγελματίες.

- Οι στοχαστικές μέθοδοι προσομοίωσης είναι πιο εύκολα εφαρμόσιμοι, γιατί δεν περιλαμβάνουν πολύπλοκους τύπους και άρα δεν απαιτούν πολύπλοκους υπολογισμούς. Η μέθοδος προσομοίωσης bootstrapping όμως είναι δύσκολο να εφαρμοστεί, όταν η ανάπτυξη είναι αρνητική η ανάπτυξη των ζημιών. Επίσης, δεν περιλαμβάνει η μέθοδος αυτή την κρίση του αναλογιστή και προγενέστερες πληροφορίες.

- Το λάθος πρόβλεψης μπορεί να έχει μεγάλη τιμή, αυτό μπορεί να οφείλεται στον μικρό αριθμό δεδομένων ή στη χρήση μοντέλου εκτίμησης που δεν ταιριάζει στην περίπτωση.

- Η ευελιξία των Γενικευμένων Γραμμικών Μοντέλων δίνει στον χρήστη τη δυνατότητα να διερευνά διαφορετικές υποκείμενες κατανομές ώστε να μπορεί να κατανοεί καλύτερα τα δεδομένα. Αυτό μπορεί να γίνει με τη χρήση βασικών στατιστικών πακέτων και με τη μέθοδο αυτήν παίρνουμε καλύτερες προσεγγίσεις για θετικά δεδομένα. Τα Γενικευμένα Γραμμικά Μοντέλα δίνουν περισσότερο λεπτομερείς πληροφορίες όπως η έκταση των ζημιών και η ανάπτυξη των ζημιών, γενικότερα δίνουν μια πιο λεπτομερή εκτίμηση του αποθέματος των ζημιών.

- Περιορισμούς ως προς τα αρνητικά δεδομένα παρουσιάζουν η Λογαριθμοκανονική και η Poisson μέθοδος, στην περίπτωση τέτοιων δεδομένων χρησιμοποιείται εναλλακτική μέθοδος.

Βιβλιογραφία

1. Απόφαση 3/13/18-11-2008 της Επιτροπής Εποπτείας Ιδιωτικής Ασφάλισης με Θέμα: «Ρύθμιση Θεμάτων Ασφαλίσεων κατά Ζημιών».
2. Klaus D. Schmidt, (CAS 2006) Methods and Models of Loss Reserving Based on Run-Off Triangles: A Unifying Survey: 270-298
3. Alba, E., Nietro-Baragas, L., E., (2008). Claims reserving : A correlated Bayesanian model. Insurance, Mathematics and Economics 43: 368-376
4. England, P., Verall, R., (1999). Analytic and boostrap estimates of prediction errors in claims reserving. Insurance, Mathematics and Economics 25: 281-293
5. Han, Z., Gau, W., C., (2008). Estimation of loss reserves with lognormal development factors. Insurance, Mathematics and Economics 42: 389-395
6. Hoedemakers, T., Beirlant, J., Goovaerts, M., J., Dhaene, J., (2003). Confidence bounds for discounted loss reserves. Insurance, Mathematics and Economics 33 : 297-316.
7. Klaus, T., H., Klauss, D., S., ZZocher, M., (2006). Multivariate loss prediction in the multivariate additive model. Insurance, Mathematics and Economics 39 : 185-191
8. Kunkler, M., (2004). Modelling zeros in stochastic reserving models. Insurance, Mathematics and Economics 34 : 23-35
9. Kunkler, M., (2006). Modelling negatives in stochastic reserving models. Insurance, Mathematics and Economics 38 : 540-555
10. Verall, R., (1996). Claims reserving and generalized additive models. Insurance, Mathematics and Economics 19 : 31-43
11. Zhou, J. & Garrido, J., (2006). A loss reserving model within the framework of generalized linear models. NSERC Discovery Grant 36860-06.