



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**UNIVERSITY OF PIRAEUS**

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**



**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ  
ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ**

**ΣΤΟΧΑΣΤΙΚΑ ΜΟΝΤΕΛΑ ΑΠΟΘΕΜΑΤΩΝ ΓΙΑ ΤΗΝ ΕΚΤΙΜΗΣΗ  
ΤΩΝ ΔΙΑΓΩΝΙΩΝ ΕΠΙΔΡΑΣΕΩΝ ΚΑΙ ΠΡΟΒΛΕΨΗ ΤΩΝ RBNS  
ΚΑΙ IBNR ΑΠΑΙΤΗΣΕΩΝ**

**Διπλωματική Εργασία**

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου

**ΜΑΡΚΟΥ Α. ΓΕΩΡΓΙΟΣ**

**ΠΕΙΡΑΙΑΣ**

**ΟΚΤΩΒΡΙΟΣ 2020**

Η παρούσα διπλωματική εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη Γ.Σ.Ε.Σ. του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμόν ..... συνεδρίαση του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνων.

Τριμελής επιτροπή:

1. Αναπληρωτής Καθηγητής, Χατζηκωνσταντινίδης Ευστάθιος (Επιβλέπων)
2. Αναπληρώτρια Καθηγήτρια, Βερροπούλου Γεωργία
3. Αναπληρωτής Καθηγητής, Σεβρόγλου Βασίλειος

Η έγκριση της Διπλωματική Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**UNIVERSITY OF PIRAEUS**

**UNIVERSITY OF PIRAEUS**

**DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE**



**POSTGRADUATE PROGRAM IN**

**ACTUARIAL SCIENCE AND RISK MANAGEMENT**

**STOCHASTIC CLAIMS RESERVING FOR DIAGONAL EFFECTS  
AND PREDICTION OF RBNS AND IBNR CLAIMS**

**MSc Dissertation**

submitted to the Department of Statistics and Insurance Science of the University of Piraeus in partial fulfilment of the requirements for the degree of Master of Science in Actuarial Science and Risk Management

**by MARKOU A. GEORGIOS**

**PIRAEUS  
OCTOBER 2020**

The committee consists of:

1. Associate Professor , Chadjikonstantinidis Efstathios (Supervisor)
2. Associate Professor, Verropoulou Georgia
3. Associate Professor, Sevroulou Vasilios

Diploma thesis submitted to the Department of Statistics and Insurance Science of the University of Piraeus as part of the requirements for obtaining the Postgraduate Diploma of Specialization in Actuarial Science and Risk Management.

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Χατζηκωνσταντινίδη Ευστάθιο, Αναπληρωτή Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς, για την καθοδήγηση και την επίβλεψη που χωρίς αυτά η πραγματοποίηση της παρούσας διπλωματικής εργασίας θα ήταν αδύνατη.

Την οικογένειά μου, την μητέρα μου Βασιλική, τον πατέρα μου Αντώνιο και τον αδερφό μου Κωνσταντίνο, για την στήριξη και την αγάπη που μου έχουν προσφέρει και ήταν δίπλα μου σε κάθε εμπόδιο.

## ΠΕΡΙΛΗΨΗ

Θεωρητικά μπορούμε να προβλεψουμε μελλοντικές ζημιές, αυτό μπορεί να επιτευχτεί με την χρήση της απλής μεθόδου αποθεματοποίησης Chain Ladder. Θα αναλύσουμε την θεωρία της καθώς και τους ορισμούς της. Η μέθοδος Chain Ladder χρησιμοποιείται για την πρόβλεψη του μελλοντικού αποθέματος που χρειάζεται για την κάλυψη των ζημιών, οι οποίες έχουν συμβεί αλλά δεν έχουν ανακοινωθεί και αυτές που έχουν ανακοινωθεί αλλά δεν έχουν εξοφληθεί. Είναι η πιο δημοφιλής μέθοδος για την αποθεματοποίηση των ζημιών (claims reserve) καθώς εκτιμώνται οι αποζημιώσεις που αναμένεται να γίνουν με την χρήση ιστορικών δεδομένων. Κυρίως προβλέπει τις αποζημιώσεις που έχουν γίνει αλλά δεν έχουν ανακοινωθεί (IBNR) και το τελικό ποσό ζημιάς που θα προκύψει. Θα παρατηρήσουμε την γενικευμένη μορφή της Chain Ladder μεθόδου καθώς και διάφορες επεκτάσεις της. Όπως για παράδειγμα μια εκδοχή της σε ένα Poisson μοντέλο ή τον συνδυασμό της με την μέθοδο Bornheutter – Ferguson. Επίσης, πως μπορούν να εισχωρήσουν οι διαγώνιες επιδράσεις στις ασφαλίσεις ζημιών που βασίζονται σε τρίγωνα εξέλιξης. Ακόμα, θα παρουσιαστεί ένα στοχαστικό μοντέλο σύνθετης Poisson που περιλαμβάνει την καθυστέρηση από τη στιγμή που προκύπτει μια απαίτηση και μέχρι την αναφορά της (καθυστέρηση IBNR) και την καθυστέρηση από τη στιγμή που αναφέρεται μια απαίτηση και μέχρι την πλήρη εξόφληση (η καθυστέρηση RBNS). Τέλος, θα παρουσιάσουμε ένα στοχαστικό μοντέλο συμπλέγματος Poisson που αποσκοπεί στην πρόβλεψη των αριθμών των αποζημιώσεων και των συνολικών ποσών των αποζημιώσεων για ένα έτος.

## **SUMMARY**

Theoretically we can predict future losses, this can be achieved by using the Chain Ladder method. We will analyze its theory as well as its definitions. The Chain Ladder method is used to predict the future reserve needed to cover losses that have incurred but have not been reported and those that have been reported but have not been settled. It is the most popular method for claim reserve as claims are estimated to be expected to be made using historical data. It mainly provides the claims that has been incurred but has not been reported (IBNR) and the final claim amount that will incur. We will observe the generalized form of the Chain Ladder method as well as its various extensions. For example, a version of it in a Poisson model or its combination with the Bornheutter-Ferguson method. Also, how the diagonal effects can be included in the non-life insurance based on triangles of historic data. Also, a Stochastic Poisson model will be introduced that includes the delay from the time a claim incurs to its report (IBNR delay) and the delay from the time a claim is reported to its full settlement (the RBNS delay). Finally, we present a thoughtful cluster Poisson model that predicts number of claims and total claim amounts for a year.

# Περιεχόμενα

## Κεφάλαιο 1<sup>ο</sup> : Ντετερμινιστική Μέθοδος Chain Ladder - Τρίγωνο Εξέλιξης Ζημιών

- 1.1 Σκοπός
- 1.2 Δημιουργία
- 1.3 Ορισμός
- 1.4 Ορισμός Επιμέρους Κομματιών
- 1.5 Μέθοδος Chain Ladder
- 1.6 Παράδειγμα Chain Ladder Μεθόδου

## Κεφάλαιο 2<sup>ο</sup> : Η Μέθοδος Chain Ladder και οι Ελεκτάσεις της

- 2.1 Περίληψη
- 2.2 Εισαγωγή
- 2.3 Αποθεματοποίηση Απαιτήσεων
- 2.4 Ένα μοντέλο Poisson για αριθμούς αποζημιώσεων
  - 2.4.1 Εκτίμηση και σύνδεση με την Μέθοδο Chain Ladder
  - 2.4.2 Ασυμπτωτικά Αποτελέσματα
  - 2.4.3 Ταυτοποίηση και Πρόβλεψη
  - 2.4.4 Ένα εκτεταμένο μοντέλο με υπερβολική διασπορά
  - 2.4.5 Η Μέθοδος Bornheutter-Ferguson
  - 2.4.6 Ένα μοντέλο αξιοπιστίας που συνδυάζει τις μεθόδους Chain Ladder και Bornheutter-Ferguson
  - 2.4.7 Διαγώνιες Επιδράσεις
- 2.5 Ποσά Αποζημιώσεων
  - 2.5.1 Μοντέλο για Ποσά Αποζημιώσεων
  - 2.5.2 Εκτίμηση
  - 2.5.3 Ασυμπτωτικά Αποτελέσματα
  - 2.5.4 Η Μέθοδος bootstrap
  - 2.5.5 Η μέθοδος Bornheutter-Ferguson και ένα μοντέλο αξιοπιστίας που συνδυάζει τις μεθόδους Chain Ladder και Bornheutter-Ferguson
  - 2.5.6 Διαγώνιες Επιδράσεις

## Κεφάλαιο 3<sup>ο</sup> : Διαγώνιες Επιδράσεις Στην Αποθεματοποίηση Απαιτήσεων

- 3.1 Περίληψη
- 3.2 Εισαγωγή



- 3.3 Μέθοδος Bornhuetter – Ferguson συμπεριλαμβανομένων των διαγώνιων επιδράσεων
- 3.4 Ένα μοντέλο αξιοπιστίας που περιλαμβάνει διαγώνιες επιδράσεις
- 3.5 Έρευνα σε δεδομένα
  - 3.5.1 Έρευνα σε δεδομένα : Μέθοδος Bornhuetter - Ferguson που περιλαμβάνει διαγώνιες επιδράσεις
  - 3.5.2 Έρευνα σε δεδομένα : Ένα μοντέλο αξιοπιστίας που περιλαμβάνει διαγώνιες επιδράσεις

## Κεφάλαιο 4<sup>ο</sup> : Πρόβλεψη RBNS και IBNR Απαιτήσεων χρησιμοποιώντας Ποσά Απαιτήσεων και Μετρήσεις Απαιτήσεων

- 4.1 Περίληψη
- 4.2 Εισαγωγή
- 4.3 Δεδομένα και Σημειογραφία
  - 4.3.1 Δεδομένα και Σημειογραφία
  - 4.3.2 Σημειογραφία
  - 4.3.3 IBNR και RBNS Απαιτήσεις
- 4.4 Το Στοχαστικό Μοντέλο
  - 4.4.1 Μετρήσεις Απαιτήσεων
  - 4.4.2 Ποσά Αποζημιώσεων : Μία αποζημίωση ανά απαίτηση
- 4.5 Εκτίμηση
- 4.6 Πρόβλεψη
  - 4.6.1 Πρόβλεψη των συνολικών IBNR Απαιτήσεων
  - 4.6.2 Πρόβλεψη των συνολικών RBNS Απαιτήσεων
- 4.7 Έρευνα σε δεδομένα
  - 4.7.1 Εκτίμηση στο μοντέλο (4.4.1) και (4.4.2)
  - 4.7.2 Εγκυρότητα Μοντέλου
  - 4.7.3 Εκτίμηση των IBNR και RBNS απαιτήσεων
  - 4.7.4 Σύγκριση με τα αποτελέσματα της Chain Ladder Μεθόδου
- 4.8 Παράρτημα Α
  - 4.8.1 Η συνάρτηση πιθανοφάνειας για τα  $\mathfrak{X}_m$
  - 4.8.2 Η συνάρτηση πιθανοφάνειας για  $\Delta_m$  δοθέντος  $\mathfrak{X}_m$
- 4.9 Παράρτημα Β

## Κεφάλαιο 5<sup>ο</sup> : Πρόβλεψη αποζημιώσεων σε ένα μοντέλο συμπλέγματος Poisson

- 5.1 Περίληψη
- 5.2 Εισαγωγή
- 5.3 Το σφάλμα της πρόβλεψης
  - 5.3.1 Πρόβλεψη ενός βήματος μπροστά

- 5.3.2 Πρόβλεψη πολλαπλών βημάτων μπροστά
- 5.3.3 Υποθετικές ανεξάρτητες αποζημιώσεις
- 5.4 Πρόβλεψη στην  $(a,b)$  – τάξη
- 5.5 Η ασυμπτωτική συμπεριφορά της πρόβλεψης
  - 5.5.1 Η συμπεριφορά των  $R_k(\gamma)$  όταν  $k \rightarrow \infty$ .
    - 5.5.1.1 Η Αρνητική Διωνυμική Κατανομή
    - 5.5.1.2 Η Διωνυμική Κατανομή
    - 5.5.1.3 Η Poisson Κατανομή
  - 5.5.2 Η συμπεριφορά των  $R_k(\gamma)$  για μεγάλα  $\gamma$
- 5.6 Πρόβλεψη σε ένα σύνολο δεδομένων από ασφαλίσεις ζημιών

## Βιβλιογραφία

# Κεφάλαιο Πρώτο

## 1. Ντετερμινιστική Μέθοδος Chain Ladder - Τρίγωνο Εξέλιξης Ζημιών

### 1.1 Σκοπός

Αρχικά καλό θα ήταν να αναφερθούμε στον σκοπό που έχει αυτή η μέθοδος. Η μέθοδος Chain Ladder χρησιμοποιείται για την πρόβλεψη του μελλοντικού αποθέματος που χρειάζεται για την κάλυψη των ζημιών, οι οποίες έχουν συμβεί αλλά δεν έχουν ανακοινωθεί και αυτές που έχουν ανακοινωθεί αλλά δεν έχουν εξοφληθεί. Είναι η πιο δημοφιλής μέθοδος για την αποθεματοποίηση των ζημιών (claims reserve) καθώς εκτιμώνται οι αποζημιώσεις που αναμένεται να γίνουν με την χρήση ιστορικών δεδομένων. Κυρίως προβλέπει τις αποζημιώσεις που έχουν γίνει αλλά δεν έχουν ανακοινωθεί (IBNR) και το τελικό ποσό ζημιάς που θα προκύψει.

### 1.2 Δημιουργία – Γέννηση

Την δημιουργία της μεθόδου Chain Ladder ανέλαβε ο James N. Stanard το 1985 στο άρθρο του (Proceedings) "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques" [15]. Αν και ο ίδιος αναφέρει πως η πρώτη φορά φέρνει στον κόσμο την Chain Ladder τον Μάιο του 1980 στο paper "Experience Rates at Estimators: A Simulation of their Bias and Variance" [15]. Βέβαια, σαν μέθοδος έχει αναλυθεί πιο πολύ από τον Thomas Mack στα παρακάτω papers "A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves" το 1991 και "Distribution – Free Calculations Of The Standard Error Of Chain Ladder Reserve Estimates" το 1993.

### 1.3 Ορισμός

Σαν ορισμό θα μπορούσαμε να πούμε πως είναι η "ανάλυση" του ονόματος της μεθόδου. Μιας και είναι η χρήση της στην πρόβλεψη των IBNR απαιτήσεων που δημιουργούνται χρόνο με τον χρόνο. Μπορούμε επίσης να πούμε χαριτολογώντας πως η Chain Ladder είναι μία σκάλα που βασιζόμενοι στον προηγούμενο χρόνο βρισκόμαστε στον επόμενο και αλυσίδα γιατί συνδέονται άρρηκτα οι χρονικές περίοδοι και οι απαιτήσεις. Άρα είναι μια αλυσιδωτή σκάλα.

### 1.4 Ορισμοί Επιμέρους Κομματιών

Πριν ξεκινήσουμε στη μεθοδολογία της Chain Ladder Μεθόδου καλό θα ήταν να δώσουμε ορισμούς για όλα τα στοιχεία που χρησιμοποιούμε και αναλύουμε.

- 1) Απόθεμα : μία ποσότητα χρημάτων η οποία έχει φυλαχθεί για μελλοντική χρήση.
- 2) Ζημία : το κόστος από ένα ατύχημα σε μία ασφάλιση.

- 3) Απαιτήση : η διεκδίκηση της ασφαλιστικής κάλυψης εφόσον έχει επέλθει ατύχημα.
- 4) Αποζημίωση : το χρηματικό ποσό που καταβάλλεται στον ασφαλισμένο εφόσον έχει επέλθει το ατύχημα.
- 5) Ιστορικά Δεδομένα : είναι οι πληροφορίες για τα έτη, τα κόστη, τα μεγέθη και λοιπά στοιχεία από γεγονότα που έχουν συμβεί με αποτέλεσμα την χρήση τους για την εκτίμηση μελλοντικών γεγονότων.
- 6) IBNR Απαιτήσεις : οι απαιτήσεις που έχουν συμβεί αλλά δεν έχουν ανακοινωθεί στην ασφαλιστική εταιρεία.
- 7) RBNS Απαιτήσεις : οι απαιτήσεις που έχουν ανακοινωθεί στην ασφαλιστική εταιρεία αλλά δεν έχουν εξοφληθεί πλήρως.
- 8) Έτος Ατυχήματος : το ημερολογιακό συνήθως έτος στο οποίο λαμβάνει πράξη ένα γεγονός, στην περίπτωσή μας το ατύχημα.
- 9) Έτος Εξέλιξης : το ασφαλιστικό έτος δηλαδή το έτος στο οποίο από την στιγμή που πραγματοποιείται η ασφαλιστική κάλυψη έχει ισχύ για 365 ημέρες.

## 1.5 Chain Ladder Μέθοδος

Αρχικά η μέθοδος αυτή, θέλει να εκτιμήσει τις αναμενόμενες αποζημιώσεις οπότε γράφουμε τα ιστορικά δεδομένα με τον εξής τρόπο:

Καταγραφή των ιστορικών δεδομένων μας σε έναν τριγωνικό πίνακα.

$i/j$	1	2	...	$j$	...	$n-1$	$n$
1	$c_{11}$	$c_{12}$		$c_{1j}$		$c_{1,n-1}$	$c_{1n}$
2	$c_{21}$	$c_{22}$	...	$c_{2j}$	...	$c_{2,n-1}$	
⋮		⋮	⋱	⋮	⋱		
$i$	$c_{i1}$	$c_{i2}$	...	$c_{ij}$			
⋮		⋮	⋱	⋱			
$n-1$	$c_{n-1,1}$	$c_{n-1,2}$					
$n$	$c_{n1}$						

Πίνακας 1.1.

Το  $c$  υποδηλώνει το κόστος των αναμενόμενων απαιτήσεων για μία χρονική στιγμή. Ισχύει ότι  $1 \leq i, j \leq n$  καθώς και  $i, j \in \{1, 2, 3, \dots, n\}$ . Το  $i$  δείχνει τα έτη στα οποία γίνεται η έρευνά μας - ζημία, έτη εμφάνισης - ατυχήματος, και το  $j$  είναι το έτος στο οποίο εξελίσσεται η εκάστοτε απαίτηση. Συνήθως αντί για τον Πίνακα 1.1 χρησιμοποιείται ο Πίνακας 1.2. Η μόνη διαφορά τους είναι πως ενώ στον Πίνακα 1.1 έχουμε την κάθε απαίτηση για το έτος που συνέβη και το πως εξελίχθηκε στην έρευνά μας, στον Πίνακα 1.2 έχουμε την αθροιστική καταγραφή των δεδομένων μας.

$i/j$	1	2	...	$j$	...	$n-1$	$n$
1	$C_{11}$	$C_{12}$		$C_{1j}$		$C_{1,n-1}$	$C_{1n}$
2	$C_{21}$	$C_{22}$	...	$C_{2j}$	...	$C_{2,n-1}$	
$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\ddots$		
$i$	$C_{i1}$	$C_{i2}$	...	$C_{ij}$			
$\vdots$		$\vdots$	$\ddots$				
$n-1$	$C_{n-1,1}$	$C_{n-1,2}$					
$n$	$C_{n1}$						

Πίνακας 1.2.

Η μόνη διαφορά στους παραπάνω πίνακες είναι πως στον 2<sup>ο</sup> πίνακα χρησιμοποιείται το άθροισμα (συσσώρευση) των προηγούμενων ετών εξέλιξης στο κάθε έτος ατυχήματος. Πιο αναλυτικά στον Πίνακα 1.2 μπορούμε να δώσουμε τον ακόλουθο τύπο :  $C_{ij} = \sum_{k=0}^j C_{ik}$ .

Το επόμενο βήμα στην μέθοδο αυτή είναι να δημιουργήσουμε τους παράγοντες εξέλιξης αναγγελθεισών αποζημιώσεων διαδοχικών ετών. Στην πράξη είναι να διαιρέσουμε, στο ίδιο έτος ατυχήματος πάντα, το επόμενο προς το προηγούμενο έτος εξέλιξης. Πιο αναλυτικά το εκάστοτε κελί  $i, j + 1$  προς το προηγούμενο κελί  $i, j$  όπου  $i \geq 0$  και  $j \geq 1$ .

Δηλαδή θα έχουμε τον παρακάτω πίνακα.

$i/j$	1 - 2	2 - 3	...	$j - j + 1$	...	$n - 1 - n$
1	$C_{12}/C_{11}$	$C_{13}/C_{12}$		$C_{1,j+1}/C_{1j}$		$C_{1n}/C_{1,n-1}$
2	$C_{22}/C_{21}$	$C_{23}/C_{22}$	...	$C_{2,j+1}/C_{2j}$	...	
$\vdots$	$\vdots$		$\ddots$	$\vdots$	$\ddots$	
$i$	$C_{i2}/C_{i1}$	$C_{i3}/C_{i2}$	...	$C_{i,j+1}/C_{ij}$		
$\vdots$	$\vdots$		$\ddots$			
$n-1$	$C_{n-1,2}/C_{n-1,1}$					
$n$						

Πίνακας 1.3.

Παρατηρούμε ότι ο παράγοντας εξέλιξης των αναγγελθεισών αποζημιώσεων μεταξύ δύο διαδοχικών ετών  $j, j + 1$  (Loss Development Factor) ή age-to-age factor είναι  $LDF_{j,j+1} = C_{j+1}/C_j$ . Με την τροποποίηση του LDF δηλαδή την εξής  $100 * (LDF_{j,j+1} - 1)\%$  παρίσταται σαν ποσότητα το ποσοστό μεταβολής των αναγγελθεισών αποζημιώσεων από το έτος εξέλιξης  $j$  στο έτος εξέλιξης  $j+1$ . Πιο απλουστευμένα δείχνει πόσο τις εκατό αυξήθηκαν οι αποζημιώσεις του συγκεκριμένου έτους ατυχήματος από το έτος εξέλιξης  $j$  στο  $j + 1$ . Λόγω εξωραϊσμού και διευκόλυνσης θα θεωρήσουμε το κάτωθι  $l_{ij} =$

$C_{ij+1}/C_{ij}$ . Οπότε με την αλλαγή αυτή ο Πίνακας 1.3 θα μετασχηματιστεί στον Πίνακα 1.4.

$i/j$	1 – 2	2 – 3	...	$j - j + 1$	...	$n - 1 - n$
1	$l_{11}$	$l_{12}$		$l_{1j}$		$l_{1,n-1}$
2	$l_{21}$	$l_{22}$	...	$l_{2j}$	...	
⋮	⋮		⋮	⋮	⋮	
$i$	$l_{i1}$	$l_{i2}$	...	$l_{ij}$		
⋮	⋮		⋮			
$n - 1$	$l_{n-1,1}$					
$n$						

*Πίνακας 1.4*

Έπειτα θέλουμε να επιλέξουμε τους παράγοντες για κάθε διαδοχικό έτος εξέλιξης. Άρα από τον Πίνακα 1.4 θέλουμε για κάθε στήλη έναν δείκτη. Αυτός ο δείκτης ποικίλει ανάλογα με την περίσταση, την έρευνα και τη χρήση που έχει. Μερικές επιλογές είναι ο αριθμητικός μέσος, ο γεωμετρικός μέσος, σταθμισμένος μέσος με βαρύτητα υψηλότερη στα τελευταία έτη.

Οπότε με τη επιλογή του αριθμητικού μέσου στην περίπτωση της εργασίας αυτής έχουμε το παρακάτω.

$i/j$	1 – 2	2 – 3	...	$j - j + 1$	...	$n - 1 - n$
A.M.	$f_1$	$f_2$	...	$f_j$	...	$f_{n-1}$

*Πίνακας 1.5.*

Τώρα μπορούμε να κάνουμε τις εξής δύο κινήσεις είτε να συμπληρώσουμε ολόκληρο τον πίνακα είτε να συμπληρώσουμε μόνο την τελευταία στήλη για ταχύτητα. Παρακάτω θα δειχθούν και οι δύο κινήσεις για να υπάρχει μια ολοκληρωμένη εικόνα της μεθόδου Chain Ladder.

$i/j$	1	2	...	$j$	...	$n - 1$	$n$
1	$C_{11}$	$C_{12}$		$C_{1j}$		$C_{1,n-1}$	$C_{1n}$
2	$C_{21}$	$C_{22}$	...	$C_{2j}$	...	$C_{2,n-1}$	$C_{2,n-1} * f_{n-1}$
⋮		⋮	⋮	⋮	⋮		⋮
$i$	$C_{i1}$	$C_{i2}$	...	$C_{ij}$			$\hat{C}_{i,n-1} * f_{n-1}$
⋮		⋮	⋮				⋮
$n - 1$	$C_{n-1,1}$	$C_{n-1,2}$	...				$\hat{C}_{n-1,n-1} * f_{n-1}$
$n$	$C_{n1}$	$C_{n1} * f_1$	...				$\hat{C}_{n,n-1} * f_{n-1}$

*Πίνακας 1.6.*

Στον Πίνακα 1.6 έχουμε συμβολίσει με  $\hat{C}_{ij} = C_{i,j-1} * f_{j-1}$  που στη συγκεκριμένη περίπτωση ισχύει  $i \geq 1$  και  $j \geq 2$ . Εδώ έχουμε χρησιμοποιήσει από τον Πίνακα 1.2 τις συσσωρευμένες αναγγελθείσες αποζημιώσεις πολλαπλασιασμένες με τον αριθμητικό μέσο του κάθε έτους εξέλιξης. Με αυτή τη διαδικασία καταφέραμε από τον τριγωνικό πίνακα που διαθέταμε στην αρχή να έχουμε έναν ολοκληρωμένο πίνακα που με βάση τα ιστορικά δεδομένα «προβλέψαμε» τις μελλοντικές συσσωρευμένες αναγγελθείσες αποζημιώσεις. Έτσι τώρα ο πίνακας είναι ένας πίνακας  $n \times n$  διαστάσεων.

Αντίστοιχα εάν θέλαμε την στήλη  $n$  μόνο θα ήταν ο πίνακας αυτής της μορφής.

$i/j$	1	2	...	$j$	...	$n-1$	$n$
1	$C_{11}$	$C_{12}$		$C_{1j}$		$C_{1,n-1}$	$C_{1n} * p_n$
2	$C_{21}$	$C_{22}$	...	$C_{2j}$	...	$C_{2,n-1}$	$C_{2,n-1} * p_{n-1}$
$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\ddots$		$\vdots$
$i$	$C_{i1}$	$C_{i2}$	...	$C_{ij}$			$C_{ij} * p_{n-i-1}$
$\vdots$		$\vdots$	$\ddots$				$\vdots$
$n-1$	$C_{n-1,1}$	$C_{n-1,2}$		...			$C_{n-1,2} * p_2$
$n$	$C_{n1}$			...			$C_{n1} * p_1$

Πίνακας 1.7.

Αρχικά, ας εξηγήσουμε τους υπάρχοντες συμβολισμούς για να δοθεί καλύτερη εξήγηση του νέου πίνακα καθώς και τους νέους συμβολισμούς. Στην στήλη  $n$  όπως είναι εμφανές είναι στην αριστερή πλευρά του γινομένου η αθροιστική αναγγελθείσα αποζημίωση όπως προϋπήρχε και στον Πίνακα 1.2. Η πληροφορία αυτή είναι σε κάθε έτος ατυχήματος το τελευταίο στοιχείο της διαγώνιου. Άρα, οι αποζημιώσεις στα κελιά από την διαγώνιο έχουν έρθει στην στήλη  $n$ . Στο δεύτερο κομμάτι του πολλαπλασιασμού, έχουμε το νέο σύμβολο  $p_i$ . Το σύμβολο αυτό είναι ο τελικός παράγοντας εξέλιξης τελικών αποζημιώσεων δηλαδή το γινόμενο των αριθμητικών μέσων που είχαμε υπολογίσει προηγουμένως. Παρακάτω φαίνεται η μορφή του τύπου.

$$p_i = \prod_i^n f_i$$

Με τις εξής συνθήκες :

1.  $p_n = f_n = 1$ , ανεξάρτητο του  $n$ .
2. Το  $i \in \{1, 2, \dots, n\}$

Φτάσαμε στο τελευταίο βήμα που είναι ο υπολογισμός των εκτιμώμενων αποζημιώσεων.

Έτος Ατυχήματος	Έτος Εξέλιξης	Αναγγελθείσες Αποζημιώσεις	Εκτιμώμενες Τελικές Αποζημιώσεις
1	1	$C_{1n}$	$\hat{C}_{1n}$
2	2	$C_{2,n-1}$	$\hat{C}_{2n}$
⋮	⋮	⋮	⋮
$i$	$i$	$C_{i,n-i+1}$	$\hat{C}_{in}$
⋮	⋮	⋮	⋮
$n-1$	$n-1$	$C_{n-1,2}$	$\hat{C}_{n-1,n}$
$n$	$n$	$C_n$	$\hat{C}_{nn}$

Πίνακας 1.8.

Όπως είχαμε δείξει προηγουμένως το  $\hat{C}_{in} = C_{i,n-i+1} * p_{n-i+1}$ , σύμφωνα με τον τελικό παράγοντα εξέλιξης τελικών αποζημιώσεων.

Στην παρούσα φάση που βρισκόμαστε μπορούμε να υπολογίσουμε τις RBNS και τις IBNR απαιτήσεις.

Αθροίζουμε όλες τις Εκτιμώμενες Τελικές Αποζημιώσεις καθώς και όλες τις Αναγγελθείσες Αποζημιώσεις. Αφαιρούμε το Σύνολο των Αναγγελθεισών Αποζημιώσεων από το Σύνολο των Εκτιμώμενων Τελικών Αποζημιώσεων.

**Το Σύνολο των Αναγγελθεισών Αποζημιώσεων είναι οι RBNS Απαιτήσεις.**

**Το αποτέλεσμα της αφαίρεσης αυτής είναι οι IBNR Απαιτήσεις.**

Παρακάτω είναι ένα παράδειγμα το οποίο θα βοηθήσει στην κατανόηση της μεθόδου Chain Ladder, θα γίνει ολόκληρη η διαδικασία και στο τέλος θα υπολογιστούν οι RBNS και οι IBNR Απαιτήσεις.

## 1.6 Παράδειγμα Chain Ladder Μεθόδου

Έχουμε μια Ασφαλιστική Εταιρεία με τις παρακάτω ιστορικές αναφορές.

Έτος Ατυχήματος	Πληρωμές Ζημιών ανά Έτος Εξέλιξης (σε μήνες)				
	12	24	36	48	60
2010	223	112	232	156	90
2011	243	125	68	154	
2012	289	121	108		
2013	320	130			
2014	437				

Σε αυτό το παράδειγμα θα αναλύσουμε περεταίρω όλα τα στοιχεία για την βέλτιστη κατανόηση της μεθόδου.



Στο κελί (1,1) έχουμε το ποσό των 223 μονάδων, αυτό μας δείχνει ότι οι αποζημιώσεις που πληρώθηκαν στο ημερολογιακό έτος 2010 και αναλογούσαν στο ημερολογιακό έτος 2010 ανέρχονταν στο ποσό των 223 μονάδων. Επίσης, στο κελί (1,3) παρατηρείται το ποσό των 232 μονάδων, που σημαίνει πως οι αποζημιώσεις του ημερολογιακού έτους 2010 πληρώθηκαν στο ημερολογιακό έτος 2012 ανέρχονταν στο ποσό των 232 μονάδων. Τέλος στο κελί (4,2) βλέπουμε το ποσό των 130 μονάδων, αυτό το ποσό υποδηλώνει πως για το ημερολογιακό έτος 2013 αποζημιώθηκαν στο ημερολογιακό έτος 2014 απαιτήσεις ύψους 130 μονάδων.

Οι απαιτήσεις για το ημερολογιακό έτος 2010 είναι το άθροισμα της γραμμής του έτους ατυχήματος 2010, δηλαδή 813 μονάδες. Αντίστοιχα για το έτος 2011 είναι 590, το έτος 2012 είναι 518, για το έτος 2013 είναι 450 και τέλος για το έτος 2014 είναι 437. Αυτές οι απαιτήσεις είναι υπολογισμένες στο σήμερα το οποίο είναι το έτος 2015. Μία σημαντική παρατήρηση είναι πως οι ζημιές – απαιτήσεις έχουν υπολογιστεί με το πέρασ του εκάστοτε ημερολογιακού έτους.

Από τον προηγούμενο πίνακα φτιάχνουμε τις αθροιστικές πληρωμές ζημιών ανά έτος.

	Αθροιστικές Πληρωμές Ζημιών ανά Έτος Εξέλιξης (σε μήνες)				
Έτος Ατυχήματος	12	24	36	48	60
2010	223	335	567	723	813
2011	243	368	436	590	
2012	289	410	518		
2013	320	450			
2014	437				

Η διαγώνιος του παραπάνω πίνακα ανέρχεται στις 2808 μονάδες, το οποίο σημαίνει πως στην αρχή του έτους 2015 οι συνολικές απαιτήσεις από τα προηγούμενα 5 χρόνια έχουν φτάσει το παραπάνω ποσό.

Επιπρόσθετα, είναι εμφανές πως σε κάθε κελί βλέπουμε τις συνολικές απαιτήσεις που είχε η εταιρεία για το κάθε ημερολογιακό έτος ανάλογα και με το έτος που έχει γίνει η πληρωμή τους. Δηλαδή, στο κελί (2,3) βλέπουμε πως για το ημερολογιακό έτος 2011 οι συνολικές απαιτήσεις που είχε η ασφαλιστική εταιρεία μέχρι και το ημερολογιακό έτος 2013 είναι 436 μονάδες. Αντίστοιχη εξήγηση έχουν όλα τα υπόλοιπα στοιχεία του πίνακα.

Τώρα θα φτιάξουμε τους παράγοντες εξέλιξης καθώς και τους μέσους που θα χρησιμοποιήσουμε. Έχει γίνει η επιλογή του αριθμητικού μέσου.

	Παράγοντες Εξέλιξης Αναγγελθεισών Αποζημιώσεων Διαδοχικών Ετών				
Έτος Ατυχήματος	12-24	24-36	36-48	48-60	60-ULT
2010	1.5022	1.6925	1.2751	1.1245	
2011	1.5144	1.1848	1.3532		
2012	1.4187	1.2634			
2013	1.4063				
2014					

Στο κελί (1,2) του παραπάνω πίνακα βλέπουμε πως οι συνολικές πληρωμές που έγιναν για το ημερολογιακό έτος 2010 πληρωμένες μέχρι και το ημερολογιακό έτος 2012 είναι 69,25% υψηλότερες από τις συνολικές πληρωμές που έγιναν για το ημερολογιακό έτος 2010 πληρωμένες μέχρι και το ημερολογιακό έτος 2011.

Έπειτα, βλέπουμε τον αριθμητικό μέσο για κάθε έτος εξέλιξης.

	Αριθμητικοί Μέσοι Ανά Έτη Εξέλιξης			
Έτος Εξέλιξης	12-24	24-36	36-48	48-60
Αριθμητικός Μέσος	1.4604	1.3802	1.3142	1.1245

Τώρα μπορούμε να υπολογίσουμε είτε όλα τα κελιά του πίνακα με τις συσσωρευμένες αναγγελθείσες απαιτήσεις είτε μόνο την τελευταία στήλη. Για να δειχθεί πως καταλήγουμε στο ίδιο αποτέλεσμα θα προστεθούν και οι δύο πίνακες.

	Πληρωμές Ζημιών ανά Έτος Εξέλιξης (σε μήνες)				
Έτος Ατυχήματος	12	24	36	48	60
2010	223	335	567	723	813
2011	243	368	436	590	663
2012	289	410	518	681	765
2013	320	450	621	816	918
2014	437	638	881	1158	1302

Στον πρώτο πίνακα έγινε χρήση της τελευταίας αθροιστικής απαίτησης για κάθε ημερολογιακό έτος στην οποία πολλαπλασιάστηκε ο κατάλληλος αριθμητικός μέσος. Δηλαδή στο κελί (3,4) δείχνει τις εκτιμώμενες συνολικές αποζημιώσεις που θα έχει το ημερολογιακό έτος 2012 στο τέλος του ημερολογιακού έτους 2015, που το ποσό των απαιτήσεων είναι 681 μονάδες. Για το κελί (5,3) έχουμε τις εκτιμώμενες συνολικές αποζημιώσεις του έτους 2014 πληρωμένες στο ημερολογιακό έτος 2015 και αυτές τις πολλαπλασιάζουμε με τον αριθμητικό μέσο του για να μας δείξει τις συνολικές

συσσωρευμένες αποζημιώσεις του ημερολογιακού έτους 2014 πληρωμένες στο ημερολογιακό έτος 2016.

	Πληρωμές Ζημιών ανά Έτος Εξέλιξης (σε μήνες)				
Έτος Ατυχήματος	12	24	36	48	60
2010	223	335	567	723	813
2011	243	368	436	590	663
2012	289	410	518		765
2013	320	450			918
2014	437				1302

Στην περίπτωση του δεύτερου πίνακα γίνεται η χρήση του τελικού παράγοντα εξέλιξης τελικών συσσωρευμένων αποζημιώσεων.

Εδώ αναγάγετε το κάθε κελί της διαγώνιου για να φτάσει στο σύνολο της η συνολική αποζημίωση του κάθε ημερολογιακού έτους που παρατηρούμε. Δηλαδή για το ποσό των 1302 μονάδων πολλαπλασιάστηκαν όλοι οι αριθμητικοί μέσοι (δηλαδή έγινε η χρήση του τελικού παράγοντα εξέλιξης) μαζί με το ποσό των 437 μονάδων, που είναι οι αποζημιώσεις του έτους 2014 πληρωμένες στο έτος 2014. Αντίστοιχα, αιτιολογούνται όλα τα υπόλοιπα στοιχεία που συμπληρώθηκαν τώρα.

Μιας και έχουμε συμπληρώσει τον πίνακα μας με τις εκτιμώμενες τελικές αποζημιώσεις για όλα τα έτη ατυχήματος που μας είχαν δοθεί, μπορούμε να τα συνοψίσουμε σε έναν πίνακα που από τον οποίο θα βγάλουμε το συμπέρασμά μας.

Έτος Ατυχήματος	Έτος Εξέλιξης (σε μήνες)	Αναγγελθείσες Αποζημιώσεις	Εκτιμώμενες Τελικές Αποζημιώσεις
2010	12	813	813
2011	24	590	663
2012	36	518	765
2013	48	450	918
2014	60	437	1302

Τώρα, μιας και έχουμε όλα τα στοιχεία μας μπορούμε να υπολογίσουμε τις RBNS και IBNR απαιτήσεις μας.

Το σύνολο των Εκτιμώμενων Τελικών Αποζημιώσεων είναι 4461 μονάδες.

Το σύνολο των Αναγγελθεισών Αποζημιώσεων είναι 2808 μονάδες.

Αν αφαιρέσουμε από το σύνολο των Εκτιμώμενων Τελικών Αποζημιώσεων το σύνολο των Αναγγελθεισών Αποζημιώσεων τότε καταλήγουμε με 1653 μονάδες.

Σύμφωνα με την μέθοδο Chain Ladder καταλήγουμε πως οι RBNS απαιτήσεις είναι 2808 μονάδες και οι IBNR απαιτήσεις είναι 1653 μονάδες.

Τις RBNS απαιτήσεις τις έχουμε πάντα από την διαγώνιο του τριγωνικού πίνακα. Το κυρίως θέμα είναι, πως θα βρούμε την βέλτιστη εκτίμηση για τις IBNR απαιτήσεις που σύμφωνα με την μέθοδο Chain Ladder τις προβλέπουμε με την χρήση ιστορικών δεδομένων. Σε παρακάτω κεφάλαια θα δούμε επεκτάσεις της μεθόδου Chain Ladder που αποσκοπούν στην βέλτιστη πρόβλεψη αποθεμάτων.

## Κεφάλαιο Δεύτερο

### 2. Η Chain Ladder Μέθοδος και οι Ελεκτάσεις της

#### 2.1 Περίληψη

Η μέθοδος Chain Ladder είναι η πιο συχνά χρησιμοποιημένη διαδικασία για τον υπολογισμό αποθέματος σε μία ασφαλιστική εταιρεία ζημιών (όχι ζωής). Αν και η αρχική ιδέα της Chain Ladder μεθόδου είναι στοιχειώδης και βασίζεται σε ευρετική συλλογιστική σκέψη αλλά η μέθοδος Chain Ladder έχει δημιουργηθεί σε ακριβή μαθηματική δομή. Στη δομή αυτή ο αριθμός των μετατροπών και τον ελεκτάσεων έχει εμφανιστεί στην βιβλιογραφία. Πολλές από αυτές τις ελεκτάσεις χρησιμοποιούνται στη πράξη χωρίς αναφορές σε συγκεκριμένες επιστημονικές εργασίες, και έχουν γίνει μέρος της παράδοσης της αναλογιστικής άσκησης. Ο στόχος αυτού του κεφαλαίου είναι να συλλέξει αποτελέσματα σχετικά με την μέθοδο Chain Ladder που είναι κοινής χρήσεων από τους αναλογιστές. Θα χρησιμοποιήσουμε τους Chain Ladder εκτιμητές όποτε μπορούμε, αν και αυτή δεν είναι η περίπτωση, θα προτιμούσαμε να δώσουμε αποτελέσματα όπου μπορούν να παραδοθούν από κλειστού τύπου εκφράσεις.

#### 2.2 Εισαγωγή

Αναλογιστές συχνά αναφέρουν την μέθοδο Chain Ladder (CLM) σαν έναν τρόπο εκτίμησης αποθέματος για μία γραμμή παραγωγής σε μία ασφαλιστική εταιρεία ζημιών. Η πρότυπη μέθοδος βασίζεται σε ευρετική συλλογιστική σκέψη και πιθανόν μία από τις πολλές ανταγωνιστικές μεθόδους για τον καθορισμό του αποθέματος. Η μέθοδος Chain Ladder αποδεικνύεται σταθερή και διαισθητικά ελκυστική και μέχρι τώρα η CLM σε ένα βαθμό καθιερωμένη σαν μέθοδος. Η απλότητα της φύσης της μεθόδου αυτής επιτρέπει για εύκολες σε αυτήν τροποποιήσεις της διαδικασίας του υπολογισμού της. Ακόμη, ο αναλογιστής μπορεί ακόμα να προσθέσει την πείρα της εργασίας στην φύση της εργασίας που χρειάζεται. Ο τρόπος που χειρίζεται τα αποθέματα απαιτήσεων την καθιστούν ευέλικτη και χρήσιμη σε μεγάλο αριθμό δεδομένων.

Η παράδοση του αναλογισμού ωστόσο χρειάζεται πιο επιστημονική προσέγγιση για τον υπολογισμό των αποθεμάτων. Πέρα από συλλογιστική σκέψη και αναφορές στα εμπειρικά στοιχεία είναι πλεονέκτημα να γίνει η ανάλυση της Chain Ladder μεθόδου καθώς και οι ελεκτάσεις της σε μία δομή θεωρίας πιθανοτήτων, π.χ. ένα στοχαστικό μοντέλο χρειάζεται να δημιουργηθεί σε τέτοιο τρόπο ώστε η μέθοδος Chain Ladder να εμφανίζεται σαν φυσική απορροή από τις προδιαγραφές του μοντέλου. Μοντέλα που σχετίζονται με την Chain Ladder είχαν φτιαχτεί από τον Kremer (1982, 1985), Mack (1991, 1993, 1994), Hachemeister και Stanard (1975). Εξετάζουμε αυτά τα μοντέλα και συζητούμε πιθανές ελεκτάσεις και τροποποιήσεις των αρχικών ιδεών με δομημένο τρόπο. Για γνωστά αποτελέσματα δίνουμε αναφορές. Ωστόσο, σε ορισμένες περιπτώσεις, επεξεργαζόμαστε ή διευκρινίζουμε τα αποτελέσματα. Εάν δεν μπορούσαμε να βρούμε κατάλληλες αναφορές στη βιβλιογραφία, δίνουμε αποδείξεις για τα αποτελέσματα.

Προκειμένου τα αποτελέσματα να είναι εφαρμόσιμα στην πράξη, οι υπολογισμοί δεν πρέπει να είναι αριθμητικά απαιτητικοί ή χρονοβόροι, καθώς αυτή η πτυχή θα αποκλείσει πολλές ιδέες απλώς και μόνο επειδή ένας πολυάσχολος αναλογιστής δεν έχει το χρόνο να τα εφαρμόσει. Γι' αυτό, κάνουμε μια προσπάθεια να χρησιμοποιήσουμε τις εκτιμήσεις της Chain Ladder όποτε είναι δυνατόν. Σε ορισμένες περιπτώσεις, αυτή η προσέγγιση θα κάνει την εκτίμηση μη βέλτιστη, αλλά αυτό είναι μία ανταλλαγή που είμαστε πρόθυμοι να αποδεχτούμε.

Προχωράμε ως εξής. Στην Παράγραφο 3 παρουσιάζουμε κάποια σημειογραφία και εξηγούμε τη γενική ιδέα της αποθεματοποίησης απαιτήσεων και της CLM. Στην Παράγραφο 4 ξεκινάμε εξετάζοντας την αρχική ιδέα του Mack (1991), όπου τα δεδομένα μετρήσεων είναι το αντικείμενο ενδιαφέροντος. Στην Παράγραφο 5 προτείνουμε πώς μπορούν να γίνουν κατανοητά τα δεδομένα πληρωμών στο πλαίσιο της CLM. Στη συνέχεια υποστηρίζουμε ότι τα περισσότερα αποτελέσματα από την Παράγραφο 4 μεταφέρονται στην Παράγραφο 5. Τέλος, στην Ενότητα 5 εξετάζουμε την αρχική ιδέα του Mack (1993) και τη συνδέουμε με το πλαίσιο της Παραγράφου 4. Θα χρησιμοποιήσουμε την ενσωμάτωση σε μια χρονολογική σειρά για να βρούμε επεκτάσεις του μοντέλου και ακόμη και να κάνουμε μια σύνδεση μεταξύ της CLM και μιας διαδικασίας διάκρισης.

### 2.3 Αποθεματοποίηση Απαιτήσεων

Σε αυτήν την ενότητα δίνουμε μια σύντομη εισαγωγή για την αποθεματοποίηση απαιτήσεων και τη CLM. Θεωρούμε μια ασφαλιστική εταιρεία ζημιών - εκτός της ζωής που πουλάει συμβόλαια (σε έναν κλάδο επιχειρήσεων) σε μια χρονική περίοδο, ας πούμε ένας χρόνος. Αυτό το έτος αναφέρεται ως έτος ατυχήματος ή ασφαλιστικό έτος. Οι απαιτήσεις σχετικά με ένα ασφαλιστικό έτος δεν θα είναι απαραίτητα όλες αποζημιώσεις εντός αυτού του έτους. Λόγω νομικών ζητημάτων, γενικής εξέτασης της απαίτησης, καθυστέρησης από το χρόνο εμφάνισης της απαίτησης έως το χρόνο αναφοράς κ.λπ. Ωστόσο, σε κάποια στιγμή δεν θα υπάρχουν άλλες πληρωμές ανά ασφαλιστικό έτος, τότε λέμε ότι το έτος αυτό έχει κλείσει.

Επισημώς, λέμε  $X_{ij}$  να είναι οι συνολικές απαιτήσεις σχετικά με την ασφαλιστική περίοδο  $i$  που έχουν εξοφληθεί με καθυστέρηση περιόδων  $j$ . Τα ποσά των απαιτήσεων  $X_{ij}$ , με  $i + j = n$ , έχουν καταβληθεί κατά την ίδια ημερολογιακή περίοδο (έτος), δηλαδή σε περίοδο  $n \in \mathbb{N}$ , αλλά αφορούν διαφορετικές περιόδους αναδοχής (περίοδος  $i$ ). Στο τέλος της περιόδου  $m \in \mathbb{N}$ , παρατηρεί έτσι τις συνολικές αποζημιώσεις

$$X_{ij}, \quad (i, j) \in A_m, \quad (2.3.1)$$

όπου

$$A_m = \{(i, j) \in \mathbb{N} \times \mathbb{N}_0 : 1 \leq i + j \leq m\}.$$

Ο τριγωνικός πίνακας (2.3.1) αναφέρεται ως τρίγωνο εξέλιξης.

Ας υποθέσουμε ότι όλες οι περιόδους ασφάλισης εκτελούνται πλήρως με  $m \in \mathbb{N}$  περιόδους, παραδείγματος χάριν,  $X_{ij} = 0$  για  $j > m_0$ . Στη συνέχεια, το απόθεμα για την ασφαλιστική περίοδο ορίζεται ως πρόβλεψη του μη παρατηρούμενου ποσού  $X_{i,m-1} + \dots + X_{i,m_0}$  με βάση τις παρατηρήσεις (2.3.1). Το συνολικό απόθεμα, ή απλά το απόθεμα, είναι μια πρόβλεψη του αθροίσματος των μεταβλητών  $X_{ij}$ ,  $(i, j) \in B_{m,m_0}$ , όπου

$$B_{m,m_0} = \{(i, j) \in \{1, \dots, m\} \times \{0, \dots, m_0\} : i + j \geq m + 1\}.$$

Συχνά υποθέτουμε ότι  $m = m_0$  οπότε γράφουμε  $B_{m,m_0} = B_m$ . Η CLM λειτουργεί στις συσσωρευμένες αποζημιώσεις

$$Y_{ij} = X_{i0} + \dots + X_{ij} \quad (i, j) \in A_m \cup B_m.$$

Σε αυτήν τη ρύθμιση, η πρόβλεψη των τυχαίων μεταβλητών  $Y_{ij}$ ,  $(i, j) \in B_m$ , είναι ισοδύναμη με την πρόβλεψη του  $X_{ij}$  ως  $X_{ij} = Y_{ij} - Y_{i,j-1}$ ,  $(i, j) \in B_m$  όπου  $Y_{i,-1} = 0$ . Η CLM βασίζεται στους παράγοντες Chain Ladder που δίδονται ως

$$\hat{f}_j = \frac{\sum_{i=1}^{m-j-1} Y_{i,j+1}}{\sum_{i=1}^{m-j-1} Y_{ij}}, \quad 0 \leq j \leq m - 2. \quad (2.3.2)$$

Οι παράγοντες Chain Ladder μπορούν να θεωρηθούν ως σταθμισμένοι μέσοι όροι των αυξήσεων στις αποζημιώσεις από καθυστέρηση  $j$  έως  $j + 1$  ως

$$\hat{f}_j = \sum_{i=1}^{m-j-1} W_{ij} \frac{Y_{i,j+1}}{Y_{ij}}, \quad W_{ij} = \frac{Y_{i,j}}{\sum_{i=1}^{m-j-1} Y_{ij}}$$

Ισχύει ότι  $\sum_{i=1}^{m-j-1} W_{ij} = 1$ . Δεδομένης αυτής της ερμηνείας, είναι ελκυστικό να καθοριστούν οι αντίστοιχοι παράγοντες Chain Ladder

$$\hat{Y}_{ij} = Y_{i,m-i} \prod_{k=m-i}^{j-1} \hat{f}_k, \quad (i, j) \in B_m \quad (2.3.3)$$

Φυσικά, η λέξη «πρόβλεψη» δεν έχει ακόμη γίνει ακριβής με μαθηματική έννοια, καθώς δεν έχει διατυπωθεί στοχαστικό μοντέλο. Υποθέτοντας ότι  $m = m_0$ , το απόθεμα που υπολογίζεται χρησιμοποιώντας τη CLM καθορίζεται από

$$\hat{R}_m = \sum_{i=2}^m [\hat{Y}_{i,m-1} - Y_{i,m-i}] = \sum_{i=2}^m \sum_{j=m-i+1}^{m-1} [\hat{Y}_{ij} - \hat{Y}_{i,j-1}] = \sum_{(i,j) \in B_m} \hat{X}_{ij},$$

όπου  $\hat{X}_{ij} = \hat{Y}_{ij} - \hat{Y}_{i,j-1}$  για  $(i, j) \in B_m$ .

Τώρα έχει διατυπωθεί τι κατανοείται από την CLM. Στην επόμενη ενότητα παρουσιάζουμε ένα στοχαστικό μοντέλο που αναπαράγει τους παραπάνω τύπους.

## 2.4 Ένα μοντέλο Poisson για αριθμούς αποζημιώσεων

Σε αυτήν την ενότητα εξετάζουμε το πολλαπλασιαστικό μοντέλο Poisson που προτείνει ο Mack (1991), σελ. 105-106. Στο πλαίσιο του μοντέλου του Mack, η CLM εμφανίζεται με φυσικό τρόπο για να προβλέψει αποθέματα.

Ας υποθέσουμε ότι οι  $N_{ij}$ ,  $(i, j) \in A_m \cup B_m$  είναι αμοιβαία ανεξάρτητες και

$$N_{ij} \sim \text{Poisson}(e^{\mu_{ij}}) \quad (2.4.1)$$

όπου

$$\mu_{ij} = \delta + \alpha_i + \beta_j, \quad (i, j) \in A_m \cup B_m.$$

Σε ένα πλαίσιο αποθεματοποίησης, το  $N_{ij}$  δηλώνει τον αριθμό των απαιτήσεων από την ασφαλιστική περίοδο  $i$  που αναφέρεται με καθυστέρηση περιόδων  $j$ . Το διάνυσμα παραμέτρων του μοντέλου (2.4.1) δίνεται από

$$\theta_m = (a_1, \dots, \beta_0, \dots, \beta_{m-1}, \delta) \in \mathbb{R}^{2m+1}, \quad (2.4.2)$$

δηλαδή, η διάσταση του χώρου παραμέτρων αυξάνεται με  $m$ .

Πριν περάσουμε από το αποτέλεσμα του Mack (1991), υπάρχουν κάποιες παρατηρήσεις σχετικά με την ταυτότητα του μοντέλου. Το μοντέλο που καθορίζεται στην (2.4.1) προφανώς δεν αναγνωρίζεται πλήρως επειδή

$$\delta + \alpha_i + \beta_j = (\delta + a + b) + (a_i - a) + (\beta_j - b)$$

Για όλα τα  $(i, j) \in A_m \cup B_m$  και  $a, b \in \mathbb{R}$ . Διάφορες μελέτες προτείνουν διαφορετικά σχήματα ταυτοποίησης όπως για παράδειγμα  $a_1 = \beta_0 = 0$  από τους England και Verrall (1999) ή  $(\delta, \sum_{j=0}^{m-1} e^{\beta_j}) = (0, 1)$  από τον Mack (1991), σελ. 97. Σε αυτήν την ενότητα αφήνουμε  $a_1 = \beta_0 = 0$ . Θα αναλύσουμε λεπτομερέστερα το πρόβλημα ταυτοποίησης στην Ενότητα 3.3.

### 2.4.1 Εκτίμηση και σύνδεση με την Μέθοδο Chain Ladder

Σε αυτήν την ενότητα η σύνδεση μεταξύ του μοντέλου (2.4.1) και του CLM (2.3.3) γίνεται ακριβής. βλ. Mack (1991), Παράρτημα Α.

Γράφουμε :

$$M_{ij} = N_{ij} + \dots + N_{i0}, \quad (i, j) \in A_m \cup B_m.$$



Δεδομένου ότι τα  $N_{ij}$  είναι αμοιβαία ανεξάρτητα και Poisson κατανομημένες οι εξισώσεις πιθανοφάνειας με βάση το  $N_{ij}, (i, j) \in A_m$ , δίνονται από τις παρακάτω εξισώσεις

$$\sum_{i=1}^{m-j} N_{ij} = \sum_{i=1}^{m-j} e^{\delta+a_i+\beta_j}, \quad 1 \leq j \leq m-1 \quad (2.4.3)$$

$$M_{i,m-i} = \sum_{j=0}^{m-i} N_{ij} = \sum_{j=0}^{m-i} e^{\delta+a_i+\beta_j}, \quad 2 \leq i \leq m \quad (2.4.4)$$

$$\sum_{i=1}^m M_{i,m-i} = \sum_{i=1}^m \sum_{j=0}^{m-i} N_{ij} = \sum_{i=1}^m \sum_{j=0}^{m-i} e^{\delta+a_i+\beta_j} \quad (2.4.5)$$

Όπως επισημάνθηκε από τους Schmidt και Wünsche (1998) οι εξισώσεις (2.4.3) - (2.4.5) αντιστοιχούν στο βασικό κεφάλαιο των οριακών συνόλων. Συνεπάγεται απευθείας από (2.4.3) - (2.4.5) ότι

$$M_{i,m-i} = \sum_{j=0}^{m-i} e^{\delta+a_i+\beta_j} = e^{\delta+a_i} \sum_{j=0}^{m-i} e^{\beta_j}, \quad 2 \leq i \leq m \quad (2.4.6)$$

Οι παράγοντες Chain Ladder, εδώ θεωρούνται παράμετροι, καθορίζονται από

$$f_j = \frac{\sum_{k=0}^{j+1} e^{\delta+a_i+\beta_k}}{\sum_{k=0}^j e^{\delta+a_i+\beta_k}} = \frac{\sum_{k=0}^{j+1} e^{\beta_k}}{\sum_{k=0}^j e^{\beta_k}}, \quad 0 \leq j \leq m-2.$$

Τώρα προκύπτει από τον Mack (1991), Παράρτημα Α, ότι οι εκτιμήσεις μέγιστης πιθανοφάνειας για  $f_j$ ,  $0 \leq j \leq m-2$ , δίνονται από

$$\hat{f}_j = \frac{\sum_{k=0}^{j+1} e^{\beta_k}}{\sum_{k=0}^j e^{\beta_k}} = \frac{\sum_{i=1}^{m-j-1} M_{i,j+i}}{\sum_{k=0}^{m-j-1} M_{ij}}, \quad 0 \leq j \leq m-2,$$

σύμφωνα με την (2.3.2). Μαζί με (2.4.6) αυτός ο τύπος καταλήγει σε

$$\sum_{k=0}^j e^{\delta+a_i+\beta_k} = M_{i,m-i} \prod_{k=j}^{m-i} (\hat{f}_k)^{-1}, \quad (i, j) \in A_m,$$

$$\sum_{k=0}^j e^{\delta+a_i+\beta_k} = M_{i,m-i} \prod_{k=m-i}^j \hat{f}_k, \quad (i, j) \in B_m.$$

Για τον υπολογισμό των μεμονωμένων εκτιμητών παραμέτρων χρησιμοποιούμε τον περιορισμό  $a_1 = \beta_0 = 0$  μαζί με τον τύπο

$$e^{\delta+a_i+\beta_j} = \sum_{k=0}^j e^{\delta+a_i+\beta_k} - \sum_{k=0}^{j-1} e^{\delta+a_i+\beta_k}, \quad (i, j) \in A_m \cup B_m. \quad (2.4.7)$$

Ακολουθεί κατευθείαν από (2.4.7) ότι, με  $(i, j) = (1, 0)$ , λαμβάνουμε τον εκτιμητή μέγιστης πιθανότητας του  $e^\beta$ . Τότε για  $i = 1$  λαμβάνουμε τους εκτιμητές του

$$e^{\delta+\beta_j}, 0 \leq j \leq m-1. \quad (2.4.8)$$

όμοια με  $j = 0$  παίρνουμε

$$e^{\delta+\alpha_i}, 2 \leq i \leq m. \quad (2.4.9)$$

Τέλος, εάν διαιρέσουμε όλες τις εκφράσεις στα (2.4.8) και (2.4.9) με το  $e^\delta$ , φτάνουμε στο πλήρες σύνολο των εκτιμητών μέγιστης πιθανότητας για τα στοιχεία του  $\theta_m$  στο (2.4.2). Έτσι, οι παράγοντες της Chain Ladder εισάγονται για να βοηθήσουν στον υπολογισμό των εκτιμητών του  $\theta_m$  στο (2.4.2).

Συνεχίζουμε μελετώντας την ασυμπτωτική συμπεριφορά των τελευταίων εκτιμητών παραμέτρων.

## 2.4.2 Ασυμπτωτικά Αποτελέσματα

Σε αυτήν την ενότητα εξετάζουμε την ασυμπτωτική συμπεριφορά των εκτιμητών μέγιστης πιθανότητας που δίδονται στην Υποπαράγραφο 2.4.1, ιδίως των παραγόντων της Chain Ladder.

Για απλότητα ξεκινάμε εξετάζοντας τα  $(\hat{f}_0, \hat{f}_1)$ . Έχουμε

$$t = \sum_{i=1}^{m-1} e^{\delta+\alpha_i+\beta_0} \quad \text{και} \quad s = \frac{\sum_{i=1}^{m-2} e^{\alpha_i}}{\sum_{i=1}^{m-1} e^{\alpha_i}}$$

τέτοια ώστε :

$$\begin{aligned} \sum_{i=1}^{m-1} M_{i0} &\sim \text{Poisson}(t), & \sum_{i=1}^{m-1} M_{i1} &\sim \text{Poisson}(tf_0), \\ \sum_{i=1}^{m-2} M_{i1} &\sim \text{Poisson}(tsf_0), & \sum_{i=1}^{m-2} M_{i2} &\sim \text{Poisson}(tsf_0f_1). \end{aligned}$$

Από τον Mikosch (2006), σελ. 65, ακολουθεί ότι

$$t^{-1} \begin{bmatrix} \sum_{i=1}^{m-1} M_{i0} \\ \sum_{i=1}^{m-1} M_{i1} \\ \sum_{i=1}^{m-2} M_{i1} \\ \sum_{i=1}^{m-2} M_{i2} \end{bmatrix} \xrightarrow{\text{asympt}} \mathcal{N}(\xi, t^{-1}\Sigma), \quad t \rightarrow \infty,$$

όπου

$$\xi = \begin{bmatrix} 1 \\ f_0 \\ s f_0 \\ s f_0 f_1 \end{bmatrix} \quad \text{και} \quad \Sigma = \begin{bmatrix} 1 & 1 & s & s \\ 1 & f_0 & f_0 s & f_0 s \\ s & f_0 s & f_0 s & f_0 s \\ s & f_0 s & f_0 s & f_0 f_1 s \end{bmatrix}.$$

Ορίζουμε τη συνάρτηση  $h : \mathbb{R}_+^4 \rightarrow \mathbb{R}_+^2$  με  $h((x_1, \dots, x_4)') = (x_1 x_2^{-1}, x_3 x_4^{-1})'$ .

Στη συνέχεια, η μέθοδος Δέλτα, βλέπε για παράδειγμα Casella και Berger (2002), που χρησιμοποιήθηκε στο  $h$  υπονοεί ότι

$$\begin{bmatrix} \hat{f}_0 \\ \hat{f}_1 \end{bmatrix} \text{asympt} \mathcal{N} \left( \begin{bmatrix} f_0 \\ f_1 \end{bmatrix}, t^{-1} \begin{bmatrix} f_0(f_0 - 1) & 0 \\ 0 & (sf_0)^{-1} f_1(f_1 - 1) \end{bmatrix} \right), t \rightarrow \infty.$$

Αυτό το αποτέλεσμα μπορεί να γενικευτεί σε διαστάσεις  $m$  όπως διατυπώνονται παρακάτω.

### Πρόταση 2.1

Αν  $N_{ij}, (i, j) \in A_m$ , ακολουθεί την υπόθεση (2.4.1) με  $\alpha_1 = \beta_0 = 0$ . Τότε  $\delta \rightarrow \infty$ ,

$$\begin{bmatrix} \hat{f}_0 \\ \vdots \\ \hat{f}_{m-2} \end{bmatrix} \text{asympt} \mathcal{N} \left( \begin{bmatrix} f_0 \\ \vdots \\ f_{m-2} \end{bmatrix}, \text{diag} \left( \begin{array}{c} (\sum_{i=1}^{m-1} e^{\delta+\alpha_i+\beta_0})^{-1} f_0(f_0 - 1) \\ \vdots \\ (\sum_{j=0}^{m-2} e^{\delta+\alpha_1+\beta_j})^{-1} f_{m-2}(f_{m-2} - 1) \end{array} \right) \right).$$

Εφόσον  $M_{i,m-i}$  είναι ανεξάρτητο των  $\hat{f}_{m-i}, \dots, \hat{f}_{m-2}$  και

$$M_{i,m-i} \text{asympt} \mathcal{N}(\sum_{j=0}^{m-i} e^{\delta+\alpha_i+\beta_j}, \sum_{j=0}^{m-i} e^{\delta+\alpha_i+\beta_j}), \delta \rightarrow \infty,$$

προκύπτει απευθείας από την Πρόταση 1 ότι

$$\sup_{x \in \mathbb{R}} |P(M_{i,m-i} \prod_{j=m-i}^{m-2} \hat{f}_j \leq x) - P(\prod_{j=m-i}^{m-2} U_j \leq x)| \rightarrow 0, \quad \delta \rightarrow \infty,$$

όπου  $U, U_{m-i}, \dots, U_{m-2}$  είναι αμοιβαία ανεξάρτητες,

$$U \sim \mathcal{N} \left( \sum_{j=0}^{m-i} e^{\delta+\alpha_i+\beta_j}, \sum_{j=0}^{m-i} e^{\delta+\alpha_i+\beta_j} \right)$$

και

$$U_j \sim \mathcal{N} \left( f_j, (\sum_{i=1}^{m-j-1} e^{\delta+\alpha_i+\beta_j})^{-1} f_j(f_j - 1) \right), \quad m-i \leq j \leq m-2.$$

Ένα παρόμοιο αποτέλεσμα προκύπτει από τον Mikosch (2006), Theorem 2.2.4, σελ. 60, το οποίο αποδίδει αμέσως αυτό :

$$\hat{f}_j = \frac{\sum_{i=1}^{m-j-1} M_{i,j+i}}{\sum_{k=0}^{m-j-1} M_{i,j}} = f_j \frac{(\sum_{i=1}^{m-j-1} \sum_{k=0}^{j+1} e^{\mu_{ik}})^{-1} \sum_{i=1}^{m-j-1} M_{i,j+1}}{(\sum_{i=1}^{m-j-1} \sum_{k=0}^j e^{\mu_{ik}})^{-1} \sum_{i=1}^{m-j-1} M_{i,j}} \xrightarrow{a.s.} f_j \quad \text{όταν} \quad \delta \rightarrow \infty.$$

Ο Taylor (2003) δείχνει ότι οι εκτιμήσεις της Chain Ladder είναι ανοδικές το μέτρο πιθανότητας υπό όρους όπου το  $\hat{f}_j$  είναι πεπερασμένο. Πιο συγκεκριμένα, ο Taylor (2003) δείχνει ότι όταν οι παράγοντες της Chain Ladder καθορίζονται καλά τότε

$$E(\hat{f}_j | \hat{f}_j < \infty) > f_j$$

για  $0 \leq j \leq m-1$ . Με  $M_{ij} = N_{i0} + \dots + N_{ij}$  αυτό υπονοεί ότι  $\hat{M}_{i,m-1}$  όπως ορίζεται στην (2.3.3) τείνει προς τα πάνω με την έννοια ότι

$$E\left(\hat{M}_{i,m-1} \mid \prod_{j=m-i}^{m-2} \hat{f}_j < \infty\right) = E\left(M_{i,m-i} \mid \prod_{j=m-i}^{m-2} \hat{f}_j \mid \prod_{j=m-i}^{m-2} \hat{f}_j < \infty\right) > EM_{i,m-1}$$

για  $1 \leq i \leq m$ , παραπάνω λεπτομέρειες σχετικά με την απόδειξη Taylor (2003).

### 2.4.3 Ταυτοποίηση και Πρόβλεψη

Σε αυτήν την ενότητα ακολουθούμε τη γενικότερη προσέγγιση που δίνεται από Kuang (2008) όπου επισημαίνεται ότι μια σταθερά μπορεί να αφαιρεθεί από  $a_i$  και  $\beta_j$  λαμβάνοντας διαφορές όπως  $\Delta a_i = a_i - a_{i-1}$ . Συγκεκριμένα επιλέγοντας  $a = a_1$  και  $\beta = \beta_0$  παίρνουμε

$$\mu_{ij} = \mu_{10} + \sum_{k=2}^i \Delta a_k + \sum_{k=1}^j \Delta \beta_k,$$

όπου  $\sum_{k=2}^1 \Delta a_k = \sum_{k=1}^j \Delta \beta_k = 0$  λόγω συνθήκης. Μια κανονική παραμετροποίηση,  $\theta_m^*$ , μπορεί έτσι να δημιουργηθεί χρησιμοποιώντας διαφορές και ένα αρχικό σημείο,

$$\theta_m^* = (\Delta a_2, \dots, \Delta a_m, \Delta \beta_1, \dots, \Delta \beta_{m-1}, \mu_{10}).$$

Σε ομαδικούς-θεωρητικούς όρους αυτό μπορεί να διατυπωθεί ως εξής. Θεωρούμε

$$\mu = \{\mu_{ij}, (i, j) \in A_m \cup B_m\}$$

μια συνάρτηση  $\theta_m$  και προσδιορίζει το σύνολο

$$g : \begin{pmatrix} a_i, & 1 \leq i \leq m \\ \beta_j, & 0 \leq j \leq m-1 \\ \delta \end{pmatrix} \rightarrow \begin{pmatrix} a_i - a, & 1 \leq i \leq m \\ \beta_j - b, & 0 \leq j \leq m-1 \\ \delta + a + b \end{pmatrix}.$$

Τότε το  $\mu$  είναι αμετάβλητο κάτω από  $g$  επειδή  $\mu(g(\theta_m)) = \mu(\theta_m)$ . Στο Kuang (2008) αποδεικνύεται ότι το  $\theta_m^*$  είναι η μέγιστη αμετάβλητη συνάρτηση του  $\theta_m$  κάτω από  $g$ .

Το σχήμα ταυτοποίησης που προτείνεται από England και Verrall (1999),  $a_1 = \beta_0 = 0$ , λαμβάνεται με

$$\delta = \mu_{10}, \quad \alpha_i = \sum_{k=2}^i \Delta a_k, \quad \beta_j = \sum_{k=1}^j \Delta \beta_k.$$

Θα χρησιμοποιήσουμε αυτήν τη συνθήκη σε αυτήν την υποπαράγραφο, εκτός εάν αναφέρεται διαφορετικά. Το σύστημα ταυτοποίησης του Mack (1991),  $(\delta, \sum_{j=0}^{m-1} e^{\beta_j}) = (0,1)$ , μπορεί να ληφθεί από την κανονική παράμετρο με

$$\delta = 0, \quad \alpha_i = C + \mu_{10} + \sum_{k=2}^i \Delta a_k, \quad \beta_j = -C + \sum_{k=1}^j \Delta \beta_k,$$

όπου

$$C = \log \left( \sum_{j=1}^{m-1} e^{\sum_{k=1}^j \Delta \beta_k} \right).$$

Σημείωση, το  $e^{\beta_j}$  είναι το ποσοστό του συνολικού αριθμού των απαιτήσεων που αναφέρεται με καθυστέρηση περιόδων  $j$ , ενώ το  $e^{\alpha_i}$  είναι το (αναμενόμενο) συνολικό ποσό απαίτησης για την ασφαλιστική περίοδο  $i$ . Αυτό το σχήμα ταυτοποίησης συμφωνεί με την ιδέα της μεθόδου Bornheutter-Ferguson που εξετάζεται στις ενότητες 2.4.5 και 2.4.6 παρακάτω.

### Εκτός δείγματος πρόβλεψη

Στα μοντέλα που έχουν εξεταστεί έως αυτό το σημείο έχει υποτεθεί ότι  $m_0 = m$  (χρησιμοποιώντας τη σημείωση της Παραγράφου 2.3). Δηλαδή, θεωρείται ότι  $EN_{ij} = 0$  για  $j \geq m$ . Σε πολλές περιπτώσεις αυτή η υπόθεση δεν ισχύει. Εάν η αποζημίωση των απαιτήσεων που σχετίζονται με το ασφαλιστικό έτος χρειάζονται περισσότερες από  $m$  περιόδους, θα πρέπει κατά κάποιο τρόπο να εκτιμήσουμε τα  $EN_{ij}$  για  $m \leq j \leq m_0$ . Ο εκτιμητής του  $EN_{ij}$  για το  $m \leq j \leq m_0$  δεν θα πρέπει ωστόσο να εξαρτάται από την επιλογή του σχήματος ταυτοποίησης και το ακόλουθο αποτέλεσμα εξασφαλίζει ότι αυτό δεν ισχύει.

### Πρόταση 2.2

Ένας εκτιμητής του  $EN_{ij}$ ,  $m \leq j \leq m_0$ , που είναι συνάρτηση του

$$\theta_m^* = (\Delta \alpha_2, \dots, \Delta \alpha_m, \Delta \beta_1, \dots, \Delta \beta_{m-1}, \mu_{10})$$

δεν εξαρτάται από την επιλογή του σχήματος ταυτοποίησης.

Η απόδειξη της Πρότασης 4 (σε λίγο πιο γενική ρύθμιση) δίνεται από Kuang (2008).

Μια απλή συνέπεια της Πρότασης 4 είναι ότι οποιαδήποτε εκτίμηση των  $f_m, \dots, f_{m_0-1}$  με βάση τα  $\hat{f}_0, \dots, \hat{f}_{m-1}$  είναι αμετάβλητη υπό την επιλογή του σχήματος ταυτοποίησης. Ένας αριθμός τρόπων πρόβλεψης από τους παράγοντες ανάπτυξης

$f_0, \dots, f_{m_0-1}$  δίνονται στον Boor (2006). Ένας άλλος απλός τρόπος εκτίμησης του  $EN_{ij}$ ,  $m \leq j \leq m_0$ , είναι απλώς με την εκτίμηση των διαφορών  $\Delta\beta_m, \dots, \Delta\beta_{m_0-1}$  με βάση τα  $\Delta\hat{\beta}_1, \dots, \Delta\hat{\beta}_{m-1}$ .

#### 2.4.4 Ένα εκτεταμένο μοντέλο με υπερβολική διασπορά

Η υπερδιασπορά παρατηρείται συχνά εμπειρικά στα δεδομένα. Αυτό είναι

$$VN_{ij} = \gamma EN_{ij}$$

για μερικά  $\gamma > 1$ . Ένα μοντέλο συμβατό με αυτήν την ιδιότητα δίνεται εάν το  $N_{ij}$  έχει αρνητική διωνυμική κατανομή. Υποθέτουμε αυτήν την κατανομή σε αυτήν την υποπαραγράφο.

Έχουμε  $\theta_{ij}, (i, j) \in A_m \cup B_m$ , να είναι αμοιβαία ανεξάρτητες τυχαίες μεταβλητές με κατανομή

$$\theta_{ij} \sim \Gamma\left(\frac{\gamma^{-1}}{1 - \gamma^{-1}} e^{\delta + \alpha_i + \beta_j}, \frac{\gamma^{-1}}{1 - \gamma^{-1}}\right).$$

Ας υποθέσουμε περαιτέρω ότι δίνεται  $\theta_{ij}, (i, j) \in A_m \cup B_m$ , οι μεταβλητές  $N_{ij}, (i, j) \in A_m \cup B_m$ , είναι υπό όρους ανεξάρτητοι και

$$N_{ij} \sim \text{Poisson}(\theta_{ij}).$$

Με

$$r_{ij} = \frac{\gamma^{-1}}{1 - \gamma^{-1}} e^{\mu_{ij}}$$

Ακολουθεί, βλέπε Mikosch (2006), σελ. 72, αυτό

$$N_{ij} \sim \text{NegBin}(\gamma^{-1}, r_{ij}).$$

Ο απλός υπολογισμός αποδίδει  $EN_{ij} = EE(N_{ij} | \theta_{ij}) = e^{\mu_{ij}}$  που είναι παρόμοιο με το μοντέλο (2.4.1) και

$$VN_{ij} = EV(N_{ij} | \theta_{ij}) + VE(N_{ij} | \theta_{ij}) = e^{\mu_{ij}} + \frac{(1 - \gamma^{-1})e^{\mu_{ij}}}{\gamma^{-1}} = \gamma EN_{ij},$$

που είναι η επιθυμητή ιδιότητα υπερδιασποράς. Αυτές οι συνθήκες καθορίζουν μια ρύθμιση Γενικευμένου Γραμμικού Μοντέλου όπως στους McCullagh και Nelder (1989), Πίνακας 9.1. Οι McCullagh και Nelder προτείνουν μια συνάρτηση πιθανότητας Poisson για την εκτίμηση των παραμέτρων  $\alpha_i, \beta_j$  και  $\delta$ . Αυτή η διαδικασία οδηγεί στους ίδιους

εκτιμητές όπως στην (2.4.4). McCullagh and Nelder (1989), σελ. 328, δίνουν έναν εκτιμητή ροής  $\gamma$  από

$$\hat{\gamma} = \frac{2}{m(m-1)} \sum_{i=1}^m \sum_{j=0}^{m-i} \frac{(N_{ij} - e^{\mu_{ij}})^2}{e^{\mu_{ij}}}$$

που προτείνει επίσης έναν εκτιμητή για το  $r_{ij}$ , δηλαδή

$$\hat{r}_{ij} = \frac{\hat{\gamma}^{-1}}{1-\hat{\gamma}^{-1}} e^{\mu_{ij}}, \quad (i, j) \in A_m \cup B_m.$$

### 2.4.5 Η Μέθοδος Bornheutter-Ferguson

Σε αυτήν την ενότητα εξετάζουμε μια απλή τροποποίηση της μεθόδου Chain Ladder που ονομάζεται μέθοδος Bornheutter-Ferguson. Δείτε για παράδειγμα Schmidt και Zocher (2008) για μια λεπτομερή περιγραφή αυτής της μεθόδου και των σχετικών θεμάτων. Θα ήταν βολικό να χρησιμοποιήσουμε το σχήμα ταυτοποίησης

$$\left( \delta, \sum_{j=0}^{m-1} e^{\beta_j} \right) = (0, 1)$$

Στην περίπτωση αυτή το  $e^{\beta_j}$  είναι το ποσοστό του συνολικού αριθμού των απαιτήσεων που αναφέρονται με καθυστέρηση περιόδων  $j$ , ενώ το  $e^{\alpha_i}$  είναι η (αναμενόμενη) συνολική απαίτηση για την ασφαλιστική περίοδο  $i$  (το καθαρό ασφάλιστρο για την περίοδο  $i$ ). Η μέθοδος Bornheutter-Ferguson εκμεταλλεύεται απλώς την ιδέα ότι το  $e^{\alpha_i}$  αντικαθίσταται από την προηγούμενη αναμενόμενη αξία του, το καθαρό ασφάλιστρο,  $T_i$ . Με αυτόν τον τρόπο η εκτίμηση αποθέματος γίνεται

$$\hat{R}_m = \sum_{i=2}^m T_i \sum_{j=m-i+1}^{m-1} e^{\beta_j}.$$

Αυτή η απλή ιδέα αποδεικνύεται ότι είναι ένας εύκολος και πρακτικός τρόπος τροποποίησης των εκτιμήσεων της Chain Ladder, εάν οι εκτιμήσεις του  $e^{\alpha_i}$  φαίνονται απροσδιόριστες και πολύ μακριά από την  $T_i$ . Η μέθοδος Bornheutter-Ferguson βασίζει την εκτίμηση αποθέματος αποκλειστικά σε προηγούμενες πληροφορίες, ενώ η μέθοδος Chain Ladder ξεκινά μια αναδρομή καθαρά από τα παρατηρούμενα δεδομένα, δηλαδή από  $M_{i,m-i} = N_{i0} + \dots + N_{i,m-i}$ . Ένας προφανής τρόπος συνδυασμού αυτών των δύο εξωτερικών θέσεων βασίζεται σε μια Μπεϋζιανή ιδέα η οποία χρησιμοποιεί τόσο τις προηγούμενες πληροφορίες όσο και τις πραγματικές τιμές που παρατηρήθηκαν. Μια τέτοια προσέγγιση περιγράφεται στην επόμενη υποπαράγραφο.

## 2.4.6 Ένα μοντέλο αξιοπιστίας που συνδυάζει τις μεθόδους Chain Ladder και Bornheutter-Ferguson

Η ιδέα του συνδυασμού των μεθόδων Chain Ladder και Bornheutter-Ferguson έχει εξεταστεί σε διάφορα άρθρα. βλέπε για παράδειγμα Benktander (1976), Hovinen (1981), Neuhaus (1992), Mack (2000) και Wüthrich (2008). Ωστόσο, θα επιλέξουμε μια εναλλακτική προσέγγιση που σχετίζεται στενά με τον Verrall (2004) αλλά διατυπώθηκε στο πλαίσιο αυτής της εργασίας. Σε αυτό το πλαίσιο μπορεί να εξαχθεί μια απλή εκτίμηση αξιοπιστίας. Το μοντέλο αξιοπιστίας είναι ένα υπό όρους μοντέλο Chain Ladder που καθορίζεται ως εξής. Έχουμε

$$\theta_i \sim \Gamma(\alpha T_i, a)$$

να είναι αμοιβαία ανεξάρτητα για το  $i \geq 1$ , και υποθέτουμε ότι έχουμε δεδομένο  $\theta_i$ ,

$$N_{ij} \sim \text{Poisson}(\theta_i e^{\beta_j}), \quad \sum_{j=0}^{m-1} e^{\beta_j} = 1 \quad (2.4.10)$$

Τότε  $E\theta_i = T_i$ ,  $V\theta_i = a^{-1}T_i$  και

$$\log E(N_{ij} | \theta_i) = \log \theta_i + \beta_j.$$

Αυτή η σχέση ισούται με (2.4.1) με την προϋπόθεση ότι εφαρμόζεται το σχήμα ταυτοποίησης  $(\delta, \sum_{j=0}^{m-1} e^{\beta_j}) = (0, 1)$ .

Η ιδέα είναι να εκτιμηθεί το  $\theta_i$  από τον μπεϋζιανό εκτιμητή,  $\hat{\mu}_i$ , που ορίζεται ως τον εκτιμητή που ελαχιστοποιεί την  $E(\hat{\mu}_i - \theta_i)^2$  σε όλες τις μετρήσιμες, πεπερασμένων διακυμάνσεων συναρτήσεις του  $N_{ij}$ ,  $(i, j) \in A_m$ . Ακολουθώντας μια προσέγγιση παρόμοια με, για παράδειγμα, Mikosch (2006), σ. 197, λαμβάνουμε τον μπεϋζιανό εκτιμητή

$$\hat{\mu}_i = \omega_i T_i + (1 - \omega_i) \frac{\sum_{j=0}^{m-i} N_{ij}}{\sum_{j=0}^{m-i} e^{\beta_j}}, \quad (2.4.11)$$

όπου  $\omega_i = (\alpha^{-1} \sum_{j=0}^{m-i} e^{\beta_j} + 1)^{-1}$ . Παρατηρούμε ότι

$$\frac{\sum_{j=0}^{m-i} N_{ij}}{\sum_{j=0}^{m-i} e^{\beta_j}} = \left( \sum_{j=0}^{m-1} N_{ij} \right) \frac{\sum_{j=0}^{m-i+1} e^{\beta_j}}{\sum_{j=0}^{m-i} e^{\beta_j}} \frac{\sum_{j=0}^{m-1} e^{\beta_j}}{\sum_{j=0}^{m-i+1} e^{\beta_j}} \cdots \frac{\sum_{j=0}^{m-1} e^{\beta_j}}{\sum_{j=0}^{m-2} e^{\beta_j}} = \left( \sum_{j=0}^{m-1} N_{ij} \right) \prod_{j=m-i}^{m-1} f_j \quad (2.4.12)$$

με  $\sum_{j=0}^{m-1} e^{\beta_j} = 1$  από ορισμό. Αυτό αποδεικνύει ότι, εάν οι εκτιμήσεις της Chain Ladder χρησιμοποιούνται για  $f_j$ ,  $0 \leq j \leq m - 1$ , τότε η δεξιά πλευρά στο (2.4.11) είναι στην πραγματικότητα η εκτίμηση του αποθέματος της Chain Ladder.

Για την εκτίμηση της παραμέτρου  $\alpha$  χρησιμοποιούμε το ακόλουθο αποτέλεσμα.

### Πρόταση 2.3



Ορίζουμε

$$\widehat{\alpha}^{-1} = \frac{1}{m} \sum_{i=1}^m \frac{(\sum_{j=0}^{m-i} N_{ij})^2 - T_i(\sum_{j=0}^{m-i} e^{\beta_j}) - T_i^2(\sum_{j=0}^{m-i} e^{\beta_j})^2}{T_i(\sum_{j=0}^{m-i} e^{\beta_j})^2}.$$

Τότε  $E\widehat{\alpha}^{-1} = \alpha^{-1}$  και αν

$$\frac{1}{m^2} \sum_{i=1}^m T_i^{k-2} \left( \sum_{j=0}^{m-i} e^{\beta_j} \right)^{k-4} \rightarrow 0, \quad m \rightarrow \infty,$$

για  $1 \leq \kappa \leq 4$  υποστηρίζει επίσης ότι  $\widehat{\alpha}^{-1} \xrightarrow{P} \alpha^{-1}$  όταν  $m \rightarrow \infty$ .

### Απόδειξη

Πρώτα παρατηρούμε ότι

$$E(N_i^2) = T_i^2 \left( \sum_{j=0}^{m-i} e^{\beta_j} \right)^2 + \frac{T_i}{a} \left( \sum_{j=0}^{m-i} e^{\beta_j} \right)^2 + T_i \left( \sum_{j=0}^{m-i} e^{\beta_j} \right)$$

με  $N_i = N_{i0} + \dots + N_{i,m-i}$ . Αυτή η σχέση βγάζει άμεσα ότι ο εκτιμητής

$$\widehat{\alpha}^{-1} = \frac{1}{m} \sum_{i=1}^m \frac{(N_i)^2 - T_i(\sum_{j=0}^{m-i} e^{\beta_j}) - T_i^2(\sum_{j=0}^{m-i} e^{\beta_j})^2}{T_i(\sum_{j=0}^{m-i} e^{\beta_j})^2}$$

είναι αμερόληπτος. Η ανισότητα του Chebyshev τείνει τότε σε

$$P(|\widehat{\alpha}^{-1} - \alpha^{-1}| > \epsilon) \leq \frac{V\widehat{\alpha}^{-1}}{\epsilon^2}$$

όπου

$$\begin{aligned} V\widehat{\alpha}^{-1} &= \frac{1}{m^2} \sum_{i=1}^m \frac{EE(N_i^4|\theta_i) - [EE(N_i^2|\theta_i)]^2}{T_i^2(\sum_{j=0}^{m-i} e^{\beta_j})^4} \\ &\leq \frac{1}{m^2} \sum_{i=1}^m \frac{EE(N_i^4|\theta_i)}{T_i^2(\sum_{j=0}^{m-i} e^{\beta_j})^4} \\ &\leq \frac{1}{m^2} \sum_{i=1}^m \sum_{k=1}^4 \frac{c_k T_i^{k-2}}{(\sum_{j=0}^{m-i} e^{\beta_j})^{4-k}}. \end{aligned}$$

Οι σταθερές  $c_k$  υπολογίζονται με τον ακόλουθο τρόπο. Από τις παραγοντικές στιγμές

$$E(N_i^{(k)}|\theta_i) = \theta_i^k \text{ δίνει ότι}$$

$$E(N_i^4|\theta_i) = (\theta_i \sum_{j=0}^{m-i} e^{\beta_j})^4 + 6(\theta_i \sum_{j=0}^{m-i} e^{\beta_j})^3 + 7(\theta_i \sum_{j=0}^{m-i} e^{\beta_j})^2 + (\theta_i \sum_{j=0}^{m-i} e^{\beta_j}). \quad (2.4.13)$$

Και καθώς το  $\theta_i$  ακολουθεί την κατανομή Γάμμα

$$E\theta_i^k = \frac{\prod_{l=0}^{k-1} (aT_i + l)}{a^k}.$$

Ακολουθεί άμεσα ότι  $T_i \mapsto EE(N_i^4|\theta_i)$  είναι ένα πολυώνυμο τέταρτης τάξεως όπου οι σταθερές

$c_k > 0$  καθορίζονται συνδέοντας τις ροπές του  $\theta_i$  στο (2.4.13).  $\square$

Ο αρχικός εκτιμητής αξιοπιστίας του  $\theta_i$  που πρότεινε ο Benktander (1976) δίνεται στη μορφή (2.4.11) με  $\omega_i = 1 - \sum_{j=0}^{m-i} e^{\beta_j}$ . Ο λεγόμενος εκτιμητής Benktander του  $\theta_i$  δίνεται έτσι από

$$\hat{\mu}_i^B = T_i(1 - \sum_{j=0}^{m-i} e^{\beta_j}) + \sum_{j=0}^{m-i} N_{ij}. \quad (2.4.14)$$

Το αντίστοιχο απόθεμα που αναφέρεται στα (2.4.11) και (2.4.14) δίνεται ως

$$\hat{R}_m = \sum_{i=2}^m \hat{\mu}_i^B \sum_{j=m-i+1}^{m-1} e^{\beta_j}.$$

## 2.4.7 Διαγώνιες Επιδράσεις

Σε αυτήν την ενότητα εξετάζουμε τις διαγώνιες επιπτώσεις. Θυμηθείτε ότι τα  $N_{ij}$  δηλώνουν τον αριθμό των αναφερόμενων απαιτήσεων σχετικά με την ασφαλιστική περίοδο  $i$  που έχουν αναφερθεί με καθυστέρηση περιόδων  $j$ . Ως κίνητρο υποθέστε ότι θεωρούμε δεδομένα από τις πολιτικές αυτόματης ευθύνης τρίτου μέρους, όπου μια αγωγή για την ύπαρξη επιδράσεων ως αποδεκτές απαιτήσεις έχει κερδηθεί στην περίοδο  $n \leq m$ . Στη συνέχεια, είναι πιθανό ότι ορισμένες άλλες επιδράσεις-απαιτήσεις θα αναφερθούν σε αυτήν την ημερολογιακή περίοδο  $i + j = n$  (και πιθανώς τις επόμενες περιόδους,  $n + 1, n + 2, \dots$ , επίσης). Ο μέσος αριθμός των απαιτήσεων που αναφέρθηκαν κατά μήκος της  $n$ -οστής διαγώνιας θα είναι επομένως σχετικά υψηλός. Ωστόσο, το μοντέλο (2.4.1) δεν επιτρέπει διαγώνιες επιδράσεις.

Για διάφορους λόγους, οι επιδράσεις του ημερολογίου εμφανίζονται σε όλα τα είδη δεδομένων. Σε αυτήν την υποπαράγραφο, το (2.4.1) επεκτείνεται ώστε να περιλαμβάνει διαγώνιες επιδράσεις όπως περιγράφεται παραπάνω. Τα ακόλουθα βασίζονται στα

Kuang, Nielsen και Nielsen (2008a, 2008b). Αρχικά, υποθέστε ότι τα  $N_{ij}$  είναι αμοιβαία ανεξάρτητα, έτσι ώστε το  $N_{ij} \sim Poisson(e^{\mu_{ij}})$  όπου

$$\mu_{ij} = \delta + a_i + \beta_j + \pi_{i+j}. \quad (2.4.15)$$

Εδώ τα  $\pi_{i+j}$  αντιπροσωπεύουν τις διαγώνιες επιδράσεις και τα  $a_i$  και  $\beta_j$  έχουν τις ίδιες ερμηνείες όπως στο (2.4.1). Το μοντέλο (2.4.15) εξαρτάται από την παράμετρο

$$\theta_m = (\delta, a_1, \dots, a_m, \beta_0, \dots, \beta_{m-1}, \pi_1, \dots, \pi_{2m-1})$$

και (2.4.15) προφανώς δεν ταυτίζεται. Στην πραγματικότητα, εάν  $\mu = \{\mu_{ij}, (i, j) \in Am \cup Bm\}$  θεωρείται συνάρτηση του  $\theta_m$  τότε το  $\mu$  είναι αμετάβλητο υπό  $g$ , δηλαδή,  $\mu(g(\theta_m)) = \mu(\theta_m)$ , όπου

$$g : \begin{pmatrix} a_i, & 1 \leq i \leq m \\ \beta_j, & 0 \leq j \leq m-1 \\ \pi_{i+j}, & 1 \leq i+j \leq m \\ \delta \end{pmatrix} \mapsto \begin{pmatrix} a_i - a - d(i-1), & 1 \leq i \leq m \\ \beta_j - b - dj, & 0 \leq j \leq m-1 \\ \pi_{i+j} - c + d(i+j-1), & 1 \leq i+j \leq 2m-1 \\ \delta + a + b + c \end{pmatrix}$$

και  $a, b, c$  και  $d$  είναι αυθαίρετες σταθερές. Σε σύγκριση με την (2.4.1), στην (2.4.15) μπορεί κανείς να προσθέσει και να αφαιρέσει μια γραμμική τάση σε αντίθεση με το (2.4.1) όπου κάποιος θα μπορούσε μόνο να αφαιρέσει / να προσθέσει μια σταθερά. Μια γραμμική τάση μπορεί να αφαιρεθεί χρησιμοποιώντας διπλές διαφορές, όπως  $\Delta^2 a_3 = \Delta(a_3 - a_2) = a_3 - 2a_2 + a_1$ . Για παράδειγμα

$$\Delta^2(a_i - a - d(i-1)) = \Delta(a_i - d) = \Delta^2 a_i.$$

Μια κανονική παραμετροποίηση,  $\theta_m^*$ , μπορεί έτσι να δημιουργηθεί χρησιμοποιώντας διπλές διαφορές και τρία αρχικά σημεία,

$$\theta_m^* = (\Delta^2 a_3, \dots, \Delta^2 a_m, \Delta^2 \beta_2, \dots, \Delta^2 \beta_{m-1}, \Delta^2 \pi_3, \dots, \Delta^2 \pi_{2m-1}, \mu_{10}, \mu_{11}, \mu_{20}).$$

Το Kuang (2008a, 2008b) δείχνει ότι  $\theta_m^*$  είναι η μέγιστη αναλλοίωτη συνάρτηση του  $\theta_m$  κάτω από. Ειδικότερα, αν μπει

$$a_i = (i-1)(\mu_{20} - \mu_{10}) + \sum_{k=3}^i \sum_{l=3}^k \Delta^2 a_l,$$

$$\beta_j = j(\mu_{11} - \mu_{10}) + \sum_{k=3}^j \sum_{l=3}^k \Delta^2 \beta_l,$$

$$\pi_{i+j} = \sum_{k=3}^{i+j} \sum_{l=3}^k \Delta^2 \pi_l,$$

$$\delta = \mu_{10},$$

τότε λαμβάνεται το σχήμα ταυτοποίησης  $\alpha_1 = \beta_0 = \pi_1 = \pi_2 = 0$ . Με αυτήν την επιλογή του συστήματος αναγνώρισης, ο Venter (2007) δίνει εξισώσεις πιθανοφάνειας που μπορούν να εκτελεστούν αναδρομικά έως ότου συγκλίνουν οι εκτιμήσεις παραμέτρων,

$$e^{a_i} = \frac{\sum_{j=0}^{m-i} X_{ij}}{\sum_{j=0}^{m-i} e^{\beta_j + \pi_{i+j}}}, \quad 2 \leq i \leq m$$

$$e^{\beta_j} = \frac{\sum_{i=1}^{m-j} X_{ij}}{\sum_{i=1}^{m-i} e^{a_i + \pi_{i+j}}}, \quad 1 \leq j \leq m-1$$

$$e^{\pi_k} = \frac{\sum_{j=0}^{m-k} X_{i, m-k-i}}{\sum_{i=1}^{m-k} e^{a_i + b_{m-k-i}}}, \quad 3 \leq k \leq m.$$

Ο Venter (2007) αναφέρει ότι περίπου 50 βρόχοι είναι ικανοποιητικοί όταν χρησιμοποιούνται οι συνηθισμένες εκτιμήσεις Chain Ladder για  $e^{a_i}$  και  $e^{\beta_j}$  και  $e^{\pi_k} = 1$  ως σημεία εκκίνησης για την αναδρομή. Στη συνέχεια, η κανονική παράμετρος εκτιμάται από, για παράδειγμα,

$$e^{\widehat{\Delta^2 a_i}} = e^{a_i - 2a_{i-1} + a_{i-2}}$$

για  $3 \leq i \leq m$  και αντίστοιχα για  $e^{\Delta^2 \pi_{i+j}}$  και  $e^{\Delta^2 \beta_j}$ .

Για να εκτιμηθεί το  $EN_{ij}$ ,  $(i, j) \in B_m$ , είναι απαραίτητο να προβλεφθεί η μορφή  $\pi_{i+j}$  για  $m+1 \leq i+j \leq 2m-1$ . Η πιο συνηθισμένη προσέγγιση είναι να υποθέσουμε ότι οι εκτιμήσεις πιθανοφάνειας είναι παρατηρήσεις από ορισμένες χρονοσειρές από τις οποίες μπορούμε να προβλέψουμε μελλοντικές επιδράσεις ημερολογίου. Κατά την εφαρμογή μιας μεθόδου για να γίνει αυτό, πρέπει να ληφθεί υπόψη το ακόλουθο αποτέλεσμα.

### Πρόταση 5

Ένας προγνωστικός παράγοντας του  $\pi_{i+j}$ ,  $m+1 \leq i+j \leq 2m-1$ , που είναι συνάρτηση του  $\theta_m^*$ , δεν εξαρτάται από την επιλογή του σχήματος ταυτοποίησης.  $\square$

Παρόμοιο αποτέλεσμα ισχύει για  $\beta_j$ ,  $j \geq m-1$ .

Αυτό ολοκληρώνει τις μελέτες μας για τη μέθοδο Chain Ladder σε σχέση με τα δεδομένα μετρήσεων. Συνεχίζουμε μελετώντας τη μέθοδο Chain Ladder που χρησιμοποιείται σε ένα λεγόμενο τρίγωνο εξέλιξης αποζημιώσεων στην επόμενη παράγραφο.

## 2.5 Ποσά Αποζημιώσεων

Στις περισσότερες περιπτώσεις, τα πραγματικά αντικείμενα ενδιαφέροντος είναι οι αποζημιώσεις που προκαλούνται από τον αναφερόμενο αριθμό απαιτήσεων. Στην αποθεματοποίηση απαιτήσεων, κάποιος συνήθως θέλει να εκτιμήσει το τελικό ποσό που θα καταβληθεί για κάθε ασφαλιστική περίοδο μαζί με τις σχετικές ταμειακές ροές. Όπως περιγράφεται στην παράγραφο 3, είναι ένα εμπειρικά αποδεδειγμένο γεγονός ότι η μέθοδος Chain Ladder λειτουργεί καλά στις αποζημιώσεις. Σε αυτήν την Παράγραφο θα εξηγήσουμε γιατί η μέθοδος Chain Ladder είναι ένας λογικός τρόπος εκτίμησης των παραμέτρων που περιγράφουν τις αποζημιώσεις που ενεργοποιούνται από την  $N_{ij}$ ,  $(i, j) \in A_m$ . Ως εκ τούτου, θα προτείνουμε ένα μοντέλο για τις αποζημιώσεις που κληρονομούν τις σημαντικές ιδιότητες του μοντέλου (2.4.1). Στην πραγματικότητα, τα περισσότερα αποτελέσματα από την Παράγραφο 4 μεταφέρονται στο μοντέλο για ποσά αποζημιώσεων.

### 2.5.1 Μοντέλο για Ποσά Αποζημιώσεων

Ας υποθέσουμε ότι ο αριθμός των αναφερόμενων απαιτήσεων,  $N_{ij}$ , ενεργοποιεί έναν διάνυσμα από αμοιβαίων ανεξάρτητων αριθμών αποζημιώσεων

$$N_{ijl} \sim \text{Poisson}(e^{\delta + \alpha_i + \beta_j + \rho_l}), \quad l \geq 0$$

Για τις αποζημιώσεις  $i, j$  και  $l$ , τα  $Z_{ijl}^{(k)}, k \geq 1$ , θεωρείται ότι υπολογίζονται με την περασμένη δεύτερη ροπή. Επιπλέον, τα  $Z_{ijl}^{(k)}, k \geq 1$  και  $N_{ijl}$  θεωρούνται όλα αμοιβαία ανεξάρτητα για  $(i, j) \in A_m \cup B_m$  και  $0 \leq l \leq d$ . Έχουμε τη σχετική διακύμανση κάθε αποζημίωσης, που καθορίζεται από

$$\varphi = \frac{E(Z_{ijl}^{(k)})^2}{E Z_{ijl}^{(k)}},$$

να είναι σταθερές. Ορισμένες λογιστικές διευθύνσεις μας λένε ότι οι συνολικές αποζημιώσεις για την περίοδο  $i$  που πληρώνονται με καθυστέρηση περιόδων  $j$  δίνονται από

$$X_{ij} = \sum_{l=0}^j \sum_{k=1}^{N_{i,j-l,l}} Z_{i,j-l,l}^{(k)}$$

όπου  $N_{ijl} = 0$  για  $l < 0$  κατά ορισμό. Επακόλουθο είναι ότι

$$EX_{ij} = e^{\delta+\alpha_i} \sum_{l=0}^j e^{\beta_{j-l}+\rho_l} EZ_{i,j-l,l}^{(1)} = e^{\delta+\alpha_i+\zeta_j} \quad (2.5.1)$$

όπου

$$e^{\zeta_j} = \sum_{l=0}^j e^{\beta_{j-l}+\rho_l} EZ_{i,j-l,l}^{(1)}.$$

Επιπλέον, καθώς το  $X_{ij}$ ,  $(i, j) \in Am \cup Bm$ , είναι αμοιβαία ανεξάρτητα, λαμβάνουμε ότι

$$VX_{ij} = e^{\delta+\alpha_i} \sum_{l=0}^j e^{\beta_{j-l}+\rho_l} E\left(Z_{i,j-l,l}^{(1)}\right)^2 = \varphi e^{\delta_0+\alpha_i+\zeta_j}$$

που αποδίδει τη σχέση

$$VX_{ij} = \varphi EX_{ij}, \quad \forall (i, j) \in Am \cup Bm. \quad (2.5.2)$$

Οι δομές μέσης και διακύμανσης (2.5.1) και (2.5.2) σχετίζονται στενά με το μοντέλο (2.4.1) και αποδεικνύεται επαρκής για να εξηγήσει γιατί η μέθοδος Chain Ladder λειτουργεί σε δεδομένα όπως περιγράφεται παραπάνω.

Για απλότητα θα χρησιμοποιήσουμε την ίδια ορολογία με τις τελευταίες παραγράφους και θα έχουμε

$$\mu_{ij} = \delta + \alpha_i + \zeta_j.$$

## 2.5.2 Εκτίμηση

Σε αυτήν την υποπαράγραφο θεωρείται ότι δεν παρατηρούνται  $N_{ij}$ ,  $(i, j) \in Am$ . Οι εκτιμήσεις βασίζονται αποκλειστικά στις παρατηρήσεις  $X_{ij}$ ,  $(i, j) \in Am$ .

Τα μόνα χαρακτηριστικά κατανομών που καθορίζονται για τα  $X_{ij}$ ,  $(i, j) \in Am \cup Bm$  είναι η μέση και η διακύμανση μαζί με αμοιβαία ανεξαρτησία. Για δεδομένα με τέτοια προδιαγραφή McCullagh και Nelder (1989), σ. 326, προτείνουν μια συνάρτηση πιθανοφάνειας Poisson. Η ιδέα υφίσταται στον δανεισμό της συνάρτησης πιθανοφάνειας από μια γνωστή κατανομή με παρόμοια σχέση μέσου και διακύμανσης που ανήκει σε μια εκθετική οικογένεια κατανομών. Όπως στην υποπαράγραφο 2.4.4, αυτή η προσέγγιση αποδίδει τις εκτιμήσεις της Chain Ladder των  $e^\delta$ ,  $e^{\alpha_i}$  και  $e^{\zeta_j}$  με βάση τα  $X_{ij}$ ,  $(i, j) \in Am$ , που είναι παρόμοια με τα (2.4.4) - (2.4.9).

Απομένει να εκτιμηθεί το  $\varphi$ . Με τον ίδιο τρόπο όπως στην υποπαράγραφο 2.4.4, προτείνεται ο ακόλουθος εκτιμητής ρολής:

$$\hat{\phi} = \frac{2}{m(m-1)} \sum_{(i,j) \in Am} \frac{(X_{ij} - e^{\rho_{ij}})^2}{e^{\rho_{ij}}}.$$

Συνεχίζουμε εξετάζοντας την ασυμπτωτική συμπεριφορά των τελευταίων εκτιμήσεων που είναι παρόμοια με την ασυμπτωτική όπως αναφέρεται στην Πρόταση 1.

### 2.5.3 Ασυμπτωτικά Αποτελέσματα

Τα ασυμπτωτικά του μοντέλου για τα ποσά αποζημιώσεων σχετίζονται με τα ασυμπτωτικά των δεδομένων μετρήσεων. Στην πραγματικότητα από τον Mikosch (2006), σελ. 118-119, η αποζημίωση  $X_{ij}$  παρουσιάζεται ως

$$X_{ij} = \sum_{k=1}^{N_{ij}^*} Z_{ijk}^*$$

όπου  $N_{ij}^*$  είναι ανεξάρτητα των  $Z_{ijk}^*$ ,  $k \geq 1$ ,

$$N_{ij}^* \sim \text{Poisson} \left( \sum_{l=0}^j e^{\delta + a_i + \zeta_{j-l}} \right)$$

και  $Z_{ijk}^*$ ,  $k \geq 1$  είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες με

$$Z_{ijk}^* \triangleq \sum_{l=0}^j I\{J = l\} Z_{ijk}^{(1)}, \quad P(J = l) = \frac{e^{\rho_l}}{\sum_{l=0}^j e^{\rho_l}},$$

όπου  $J$  και  $Z_{ijk}^{(k)}$  είναι αμοιβαία ανεξάρτητα. Δεδομένου ότι ο  $X_{ij}$  μπορεί να ερμηνευθεί ως σύνθετη διαδικασία Poisson, ακολουθεί από Mikosch (2006), σ. 81, ότι ως  $\delta \rightarrow \infty$ ,

$$\left( \sum_{l=0}^j e^{\delta + a_i + \zeta_{j-l}} \right)^{-1} X_{ij} \xrightarrow{\text{asympt}} \mathcal{N} \left( EZ_{ij1}^*, \varphi EZ_{ij1}^* \left( \sum_{l=0}^j e^{\delta + a_i + \zeta_{j-l}} \right)^{-1} \right).$$

Παρόμοια επιχειρήματα όπως στην υποπαράγραφο 2.4.2 αποδίδουν το ακόλουθο αποτέλεσμα.

#### Πρόταση 2.4

Έχουμε τα  $X_{ij}$ ,  $(i, j) \in Am$ , ακολουθούν την υπόθεση (2.5.1) με  $\alpha_1 = \zeta_0 = 0$ . Στη συνέχεια ως  $\delta \rightarrow \infty$ ,

$$\begin{bmatrix} \hat{f}_0 \\ \vdots \\ \hat{f}_{m-2} \end{bmatrix} \underset{\text{asympt}}{\sim} \mathcal{N} \left( \begin{bmatrix} f_0 \\ \vdots \\ f_{m-2} \end{bmatrix}, \text{diag} \begin{pmatrix} \varphi \left( \sum_{i=1}^{m-1} e^{\delta+a_i+\zeta_0} \right)^{-1} f_0(f_0-1) \\ \vdots \\ \varphi \left( \sum_{j=0}^{m-2} e^{\delta+a_1+\zeta_j} \right)^{-1} f_{m-2}(f_{m-2}-1) \end{pmatrix} \right)$$

όπου  $\hat{f}_j$ ,  $0 \leq j \leq m-2$ , δίνονται στην (2.3.2).

Με τον ίδιο τρόπο όπως στην υποπαράγραφο 2.4.2, προκύπτει απευθείας από την πρόταση 4 ότι

$$\sup_{x \in \mathbb{R}} |P(Y_{i,m-i} \prod_{j=m-i}^{m-2} \hat{f}_j \leq x) - P(U \prod_{j=m-i}^{m-2} U_j \leq x)| \rightarrow 0, \quad \delta \rightarrow \infty,$$

όπου  $U, U_{m-i}, \dots, U_{m-2}$  είναι αμοιβαία ανεξάρτητα

$$U \sim \mathcal{N} \left( \sum_{j=0}^{m-i} e^{\delta+a_i+\zeta_j}, \varphi \sum_{j=0}^{m-i} e^{\delta+a_i+\zeta_j} \right)$$

και

$$U_j \sim \mathcal{N} \left( f_j, \varphi \left( \sum_{i=1}^{m-j-1} e^{\delta+a_i+\zeta_j} \right)^{-1} f_j(f_j-1) \right), \quad m-i \leq j \leq m-2.$$

Ανακαλούμε τον ορισμό του  $\hat{f}_j$  στην (2.3.2). Στη συνέχεια, από τον Mikosch (2006), σ. 81, ακολουθεί ως  $\delta \rightarrow \infty$ ,

$$\hat{f}_j = \frac{\sum_{i=1}^{m-j-1} Y_{i,j+i}}{\sum_{i=1}^{m-j-1} M_{ij}} = f_j \frac{\left( \sum_{i=1}^{m-j-1} \sum_{k=0}^{j+1} e^{\mu_{ik}} \right)^{-1} \sum_{i=1}^{m-j-1} Y_{i,j+1}}{\left( \sum_{i=1}^{m-j-1} \sum_{k=0}^j e^{\mu_{ik}} \right)^{-1} \sum_{i=1}^{m-j-1} Y_{ij}} \xrightarrow{a.s.} f_j.$$

Ο Taylor (2003) δείχνει ότι οι εκτιμήσεις της Chain Ladder είναι προς τα πάνω μεροληπτικές στο μέτρο πιθανότητας υπό όρους όπου το  $\hat{f}_j$  είναι πεπερασμένο.

## 2.5.4 Η Μέθοδος bootstrap

Η μέθοδος Bootstrap ως γενική μέθοδο δειγματοληψίας εισήχθη από τον Efron (1979). Στο πλαίσιο της αποθεματοποίησης των απαιτήσεων η μέθοδος χρησιμοποιήθηκε από τους England και Verrall (1999) και England (2002). βλ. επίσης Cairns (2000) για τη συζήτηση σχετικά με τους τρόπους που πρέπει να ποσοτικοποιηθεί η αβεβαιότητα του



αποθέματος. Στο πλαίσιο των αποθεμάτων απαιτήσεων, είναι σύνηθες να βασιζόμαστε τη Bootstrap στη δειγματοληψία στις ποσότητες

$$\hat{\epsilon}_{ij} = \frac{X_{ij} - e^{\mu_{ij}}}{e^{\frac{1}{2}\mu_{ij}}}, \quad (i, j) \in Am.$$

Η βασική παραδοχή για να έχει νόημα η μέθοδος Bootstrap είναι αυτή

$$\epsilon_{ij} = \frac{X_{ij} - e^{\mu_{ij}}}{e^{\frac{1}{2}\mu_{ij}}}, \quad (i, j) \in Am,$$

είναι ανεξάρτητες και ομοιόμορφα κατανομημένες. Είναι λοιπόν η ελπίδα μας ότι η ιδιότητα των ανεξάρτητων και ομοιόμορφα κατανομημένων μεταφέρεται ασυμπτωτικά στα  $\hat{\epsilon}_{ij}$ ,  $(i, j) \in Am$  καθώς  $\delta \rightarrow \infty$ . Για απλότητα θα αποκαλούμε τα  $\hat{\epsilon}_{ij}$  και  $\epsilon_{ij}$  σαν σφάλματα (του Pearson).

Η απλοϊκή μέθοδος Bootstrap αντιστοιχεί στο σχέδιο ανεξάρτητα από την εμπειρική συνάρτηση κατανομής του  $\hat{\epsilon}_{ij}$ , δίνεται από

$$F_m(x) = \frac{2}{m(m+1)} \sum_{(i,j) \in Am} I\{\hat{\epsilon}_{ij} \leq x\}, \quad x \in \mathbb{R}$$

Αν υποθέσουμε ρεαλιστικά ότι  $\epsilon_{ij} \sim F_m$  τότε ανεξάρτητες εκδόσεις των υπολειμμάτων

$$\epsilon_{ij}^*(k), \quad (i, j) \in Am,$$

για  $1 \leq k \leq B$  μπορεί να προσομοιωθεί. Στη συνέχεια μπορούν να δημιουργηθούν ανεξάρτητες εκδόσεις του  $X_{ij}$  από

$$X_{ij}^*(k) = e^{\frac{1}{2}\mu_{ij}} \epsilon_{ij}^*(k) + e^{\mu_{ij}}.$$

Με τη δημιουργία μεγάλου αριθμού,  $B$ , των ανεξάρτητων και ομοιόμορφα κατανομημένων πραγματοποιήσεων, τα  $X_{ij}^*(k)$ ,  $1 \leq k \leq B$ , επιτρέπει σε κάποιον να εκτιμήσει τις ιδιότητες κατανομής των σχετικών τύπων των δεδομένων. Εάν, στην πραγματικότητα, η σχέση

$$X_{ij}^*(k) = X_{ij}, \quad (i, j) \in Am,$$

ισχύει ότι

$$\hat{f}_j^* = \frac{\sum_{i=1}^{m-j-1} \sum_{l=0}^{j+1} X_{ij}^*(k)}{\sum_{i=1}^{m-j-1} \sum_{l=0}^j X_{ij}^*(k)}$$

με τον νόμο των μεγάλων αριθμών θα τείνει

$$\frac{1}{B} \sum_{(i,j) \in Am}^B I\{\hat{f}_j^* \leq x\} \xrightarrow{P} P(\hat{f}_j \leq x)$$

για  $B \rightarrow \infty$ . Ομοίως, ακολουθεί αυτό

$$\frac{1}{B} \sum_{(i,j) \in Am}^B I\left\{\sum_{l=0}^{m-i} X_{ij}^*(k) \leq x\right\} \xrightarrow{P} P\left(\sum_{l=0}^{m-i} X_{ij} \leq x\right) = P(Y_{i,m-i} \leq x).$$

### Προσαρμογές για τη λήψη των ανεξάρτητων και ομοιόμορφα καταναμημένων υπολειμμάτων

Σε ορισμένα σύνολα δεδομένων,  $X_{ij}$ ,  $(i, j) \in Am$ , τα υπολείμματα δεν είναι ομοιογενή ως συνάρτηση του  $j$  (και πιθανώς  $i$ ). Σε αυτήν την ενότητα, προτείνουμε μία πιθανή επίλυση σε αυτό το πρόβλημα. Υποθέτουμε ότι τα  $\epsilon_{ij}$ ,  $(i, j) \in Am$ , είναι ανεξάρτητα έτσι ώστε

$$E\epsilon_{ij} = 0, \quad E\epsilon_{ij}^2 = V\epsilon_{ij} = \varphi_j,$$

όπου  $\varphi_j$  είναι κάποια συνάρτηση του  $j$ . Για απλότητα, έχουμε  $\varphi_j = a + b_j$  αλλά το  $\varphi_j$  μπορεί να οριστεί και με πολλούς άλλους τρόπους. Οι παράμετροι  $a, b \in \mathbb{R}$  υπολογίζονται ελαχιστοποιώντας το άθροισμα

$$\sum_{i=1}^m \sum_{j=0}^{m-i} (\epsilon_{ij}^2 - \varphi_j)^2.$$

Στην περίπτωση που  $\varphi_j = a + b_j$  το ελάχιστο επιτυγχάνεται με

$$\hat{b} = \frac{2(\sum_{i=1}^m \sum_{j=0}^{m-i} j \epsilon_{ij}^2)k_m - m(m+1)(\sum_{i=1}^m \sum_{j=0}^{m-i} \epsilon_{ij}^2)}{2k_m - m(m+1)(\sum_{i=1}^m \sum_{j=0}^{m-i} j^2)},$$

$$\hat{a} = \frac{2}{m(m+1)} \left( \sum_{i=1}^m \sum_{j=0}^{m-i} \epsilon_{ij}^2 - \hat{b}_{k_m} \right).$$

όπου

$$k_m = \sum_{i=1}^m \sum_{j=0}^{m-i} j = \sum_{i=1}^m \frac{(m-i)^2 + m-i}{2} = \frac{(m-1)m}{4} \left( \frac{2m-1}{3} + 1 \right)$$

Τα νέα υπολείμματα μπορούν συνεπώς να οριστούν ως

$$\tilde{\epsilon}_{ij} = \varphi_j^{-1/2} \epsilon_{ij},$$

όπου  $E\tilde{\epsilon}_{ij} = 0$  και  $V\tilde{\epsilon}_{ij} = 1$ . Για να επαναδιατυπώσουμε την αβεβαιότητα του αποθέματος σε αυτό το πλαίσιο, θεωρείται ότι τα  $\epsilon_{ij}^2$  είναι τέτοια ώστε να μπορούν να εμφανιστούν ξανά στο δείγμα. Ο τύπος

$$\tilde{\epsilon}_{ij} \varphi_j^{1/2} = \epsilon_{ij}$$

τείνει σε ένα δείγμα καταλοίπων όπως ορίσθηκαν στην υποπαράγραφο 2.5.4. Μια παρόμοια προσέγγιση όπως υπήρχε στην υποπαράγραφο 2.5.4 μπορεί επομένως να χρησιμοποιηθεί.

Όπως στην Παράγραφο 4 συνεχίζουμε να μελετάμε τη μέθοδο Bornheutter-Ferguson σε σχέση με τα δεδομένα πληρωμών.

### **2.5.5 Η μέθοδος Bornheutter-Ferguson και ένα μοντέλο αξιοπιστίας που συνδυάζει τις μεθόδους Chain Ladder και Bornheutter-Ferguson**

Η μέθοδος Bornheutter-Ferguson λειτουργεί σε δεδομένα αποζημιώσεων σχεδόν με τον ίδιο τρόπο όπως με τους αριθμούς απαιτήσεων και τα περισσότερα αποτελέσματα από τις υποπαραγράφους 2.4.5 και 2.4.6 μεταφέρονται σε αυτήν την υποπαράγραφο.

Όπως στις υποπαραγράφους 2.4.5 και 2.4.6, το σχήμα ταυτοποίησης

$$\left( \delta, \sum_{j=0}^{m-1} e^{\beta_j} \right) = (0,1)$$

χρησιμοποιείται σε αυτήν την υποπαράγραφο. Και πάλι, έχουμε τις  $T_i$  να είναι οι προηγούμενες αναμενόμενες απώλειες που σχετίζονται με το ασφαλιστικό έτος  $i$ . Με τον ίδιο τρόπο όπως στην υποπαράγραφο 2.4.5, η εκτίμηση του  $e^{\alpha_i}$  αντικαθίσταται από την προηγούμενη εκτίμησή του,  $T_i$ , για την απόκτηση του Bornheutter-Ferguson αποθέματος

$$\hat{R}_m = \sum_{i=2}^m T_i \sum_{j=m-i+1}^{m-1} e^{\beta_j}$$

Αυτό το απόθεμα βασίζεται αποκλειστικά στην προηγούμενη εκτίμηση  $T_i$  των τελικών αποζημιώσεων,  $Y_{i,m-1}$ . Η μέθοδος Chain Ladder βασίζεται αποκλειστικά σε δεδομένα με την έννοια ότι

$$\hat{Y}_{i,m-1} = \frac{Y_{i,m-i}}{\sum_{j=0}^{m-i} e^{\beta_j}}$$

όπως φαίνεται στο (2.4.12). Έτσι, φαίνεται απλό να προτείνουμε ένα μοντέλο αξιοπιστίας που συνδυάζει αυτές τις δύο θεωρίες.

Όπως στην υποπαράγραφο 2.4.6 υποθέτουμε ότι το  $\theta_i$  είναι ανεξάρτητο και έχουμε

$$E\theta_i = T_i, \quad \text{και} \quad V\theta_i = \frac{T_i}{\alpha}.$$

Περαιτέρω υποθέτουμε ότι τα  $X_{ij}$ ,  $(i, j) \in Am \cup Bm$ , είναι υπό όρους ανεξάρτητα δεδομένου  $\theta_i$ ,  $1 \leq i \leq m$ , και ότι

$$E(X_{ij}|\theta_i) = e^{\beta_j}\theta_i \quad \text{και} \quad V(X_{ij}|\theta_i) = \varphi e^{\beta_j}\theta_i$$

για  $(i, j) \in Am \cup Bm$ . Στη συνέχεια, ο γραμμικός εκτιμητής,  $\hat{\mu}_i^L$ , ελαχιστοποιώντας το  $E(\hat{\mu}_i^L - \theta_i)^2$  επιπλέον όλη η γραμμική συνάρτηση του  $X_{ij}$ ,  $(i, j) \in Am$ , δίνεται από

$$\hat{\mu}_i^L = \omega_i T_i + (1 - \omega_i) \frac{Y_{i,m-i}}{\sum_{j=0}^{m-i} e^{\beta_j}}, \quad (2.5.3)$$

όπου  $\omega_i = ((\varphi\alpha)^{-1} \sum_{j=0}^{m-i} e^{\beta_j} + 1)^{-1}$ . Αυτό είναι παρόμοιο με την υποπαράγραφο 2.4.7, ωστόσο, με διαφορετικά βάρη,  $\omega_i$ , όπου το  $\alpha^{-1}$  αντικαθίσταται με  $\alpha^{-1}\varphi$ . Μια απόδειξη της (2.5.3) ακολουθεί απευθείας από το μοντέλο Bühlmann-Straub που χρησιμοποιείται στο  $X_{ij}/e^{\beta_j}$ . βλ. Mikosch σελ. 213-214.

Για να εκτιμήσουμε τις παραμέτρους σε αυτό το μοντέλο, παρατηρούμε πρώτα αυτό

$$V\left(\sum_{j=0}^{m-i} X_{ij}\right) = \left(\varphi + \frac{\sum_{j=0}^{m-i} e^{\beta_j}}{\alpha}\right) E\left(\sum_{j=0}^{m-i} X_{ij}\right).$$

Έτσι, αν βάλουμε

$$b_i = \sum_{j=0}^{m-i} e^{\beta_j} \quad \text{και} \quad \tilde{X}_i = \frac{(\sum_{j=0}^{m-i} X_{ij} - T_i \sum_{j=0}^{m-i} e^{\beta_j})^2}{T_i \sum_{j=0}^{m-i} e^{\beta_j}}$$

τότε  $E\tilde{X}_i = \varphi + \alpha^{-1}b_i$  που υποδηλώνουν μια γραμμική παλινδρόμηση του  $\tilde{X}_i$  στο  $b_i$ . Μια τυπική γραμμική παλινδρόμηση αποδίδει εκτιμητές  $\alpha^{-1}$  και  $\varphi$  που δίδονται από

$$\hat{\alpha}^{-1} = \frac{(\sum_{i=1}^m b_i)(\sum_{i=1}^m X_i) - m \sum_{i=1}^m b_i X_i}{(\sum_{i=1}^m b_i)^2 - m \sum_{i=1}^m b_i^2}, \quad \hat{\varphi} = \frac{\sum_{i=1}^m X_i - \hat{\alpha}^{-1} \sum_{i=1}^m b_i}{m}$$

Οι εκτιμήσεις της Chain Ladder χρησιμοποιούνται για την εκτίμηση του  $e^{\beta_j}$ ,  $0 \leq j \leq m - 1$ . Με παρόμοια υπόθεση όπως στην υποπαράγραφο 2.4.3 μπορεί να αποδειχθεί ότι οι εκτιμήσεις της Chain Ladder συγκλίνουν πιθανότατα στις πραγματικές τους τιμές.

## 2.5.6 Διαγώνιες Επιδράσεις

Τα αποτελέσματα σε αυτήν την ενότητα σχετίζονται με τα δεδομένα μετρήσεων,  $N_{ij}$ ,  $(i, j) \in Am$ . Οι μέθοδοι που προτείνονται για μετρήσεις μπορούν εύκολα να εφαρμοστούν σε δεδομένα αποζημιώσεων,  $X_{ij}$ ,  $(i, j) \in Am$ , λαμβάνοντας υπόψη το μοντέλο Γενικευμένο Γραμμικό Μοντέλο.

$$EX_{ij} = e^{\mu_{ij}} \quad \text{και} \quad VX_{ij} = \varphi EX_{ij},$$

για ποσά αποζημιώσεων. Όλα τα αποτελέσματα από την υποπαράγραφο 2.4.7 μεταφέρονται απευθείας.

## Κεφάλαιο Τρίτο

### 3. Διαγώνιες Επιδράσεις Στην Αποθεματοποίηση Απαιτήσεων

#### 3.1 Περίληψη

Σε αυτό το άρθρο παρουσιάζουμε δύο διαφορετικές προσεγγίσεις για το πώς μπορεί κανείς να συμπεριλάβει διαγώνιες επιδράσεις σε ασφαλίσεις ζημιών που βασίζονται σε τρίγωνα εξέλιξης. Οι εμπειρικές αναλύσεις υποδηλώνουν ότι οι προσεγγίσεις του Zehnwirth (2003) και στους Kuang, Nielsen και Nielsen (2008a, 2008b) δεν λειτουργούν καλά με χαμηλών διαστάσεων τρίγωνα εξέλιξης επειδή η αβεβαιότητα εκτίμησης είναι πολύ μεγάλη. Για να ξεπεράσουμε αυτό το πρόβλημα θεωρούμε παρόμοια μοντέλα με μικρότερο αριθμό παραμέτρων. Αυτά σχετίζονται στενά με το πλαίσιο που εξετάστηκε από τον Taylor (1977, 2000) και τον Verbeek (1972). τη μέθοδο διαχωρισμού. Εξηγούμε ότι αυτά τα μοντέλα μπορούν να ερμηνευθούν ως επεκτάσεις των πολλαπλασιαστικών μοντέλων Poisson που εισήγαγαν οι Hachemeister και Stanard (1975) και Mack (1991).

#### 3.2 Εισαγωγή

Πρόσφατα, δόθηκε κάποια προσοχή στις διαγώνιες επιδράσεις στα τρίγωνα εξέλιξης  $\{N_{ij}\}$  του αριθμού των απαιτήσεων. Για παράδειγμα, η προσέγγιση της εκτίμησης και της πρόβλεψης όπως υιοθετήθηκε από τον Zehnwirth (2003) αντιμετωπίστηκε στα Kuang, Nielsen και Nielsen (2008a, 2008b), όπου θεωρείται επίσης ότι

$$N_{ij} \sim Pois(\mu_{ij}) \tag{3.2.1}$$

με

$$\mu_{ij} = \alpha_i \beta_j \delta_{i+j}. \tag{3.2.2}$$

Οι Kuang, Nielsen και Nielsen (2008a, 2008b) επισημαίνουν ότι οι παράμετροι (3.2.2) δεν αναγνωρίζονται με μοναδικό τρόπο και, ως εκ τούτου, πρέπει να είμαστε

προσεκτικοί όταν πρόκειται για τον προσδιορισμό του μοντέλου. Πράγματι, η επιλογή του σχήματος ταυτοποίησης στον Zehnwirth (2003) έχει επίδραση στην τελική πρόβλεψη.

Το μοντέλο (3.2.1) είναι μια γενίκευση των πολλαπλασιαστικών μοντέλων Poisson που προτείνονται από τους Hachemeister και Stanard (1975) και Mack (1991). Εκεί φαίνεται ότι (3.2.1) με  $\delta_j = 1, j \geq 1$ , υponοεί τη μέθοδο Chain Ladder. Δεδομένου ότι η μέθοδος Chain Ladder είναι κοντά στο εταιρικό πρότυπο, φαίνεται λογικό να ληφθεί ως σημείο εκκίνησης. Ωστόσο, δεν μας ενδιαφέρει ο αριθμός των απαιτήσεων αλλά τα ποσά που αποζημιώθηκαν. Στο πνεύμα του Venter (2007) και των England και Verrall (1999), λοιπόν, εξετάζουμε το πλαίσιο Γενικευμένων Γραμμικών Μοντέλων, όπου υποτίθεται απλώς για ένα τρίγωνο εξέλιξης αποζημιώσεων  $\{X_{ij}\}$  που

$$EX_{ij} = \mu_{ij} \quad \text{και} \quad VX_{ij} = \varphi EX_{ij}. \quad (3.2.3)$$

Οι McCullagh και Nelder (1989) θεωρούν μοντέλα τύπου (3.2.3). Προτείνουν να χρησιμοποιηθεί ένα Poisson πιθανοτικά ισοδύναμο με την πιθανότητα στο (3.2.1) για να εκτιμηθούν οι παράμετροι στο (3.2.3). Αυτή η προσέγγιση συνδέεται πάλι με τη μέθοδο Chain Ladder (με  $\delta_j = 1, j \geq 1$ ) για ποσά αποζημιώσεων, επειδή η εκτίμηση του αποθέματος παραμένει ίδια με αυτή των Hachemeister και Stanard (1975) και Mack (1991) για τα ποσά αποζημιώσεων.

Ο Rietdorf (2008) επισημαίνει ότι οι διαγώνιες επιδράσεις,  $\delta_j$ , συνήθως προέρχονται από δύο πηγές: 1) οικονομικός πληθωρισμός, δηλαδή, οι απαιτήσεις πληρωμών ακολουθούν έναν σχετικό δείκτη τιμών που ακολουθεί και πάλι τον ημερολογιακό χρόνο και 2) νομικά ζητήματα, αλλαγές στον τρόπο χειρισμού των απαιτήσεων ή παρόμοια δίνουν διαγώνιες επιδράσεις τον αριθμό των αποζημιώσεων. 2) αναφέρεται ως πληθωρισμός απαιτήσεων. Ο οικονομικός πληθωρισμός θεωρείται συνήθως ότι λειτουργεί με πολλαπλασιαστικό τρόπο, γεγονός που υποδηλώνει ότι

$$EX_{ij} = \alpha_i \beta_j \delta_{i+j}, \quad VX_{ij} = \varphi \alpha_i \beta_j \delta_{i+j}^2,$$

όπου ο οικονομικός πληθωρισμός (βλέπε Kuang et al. (2008a, 2008b)) θεωρείται ότι ικανοποιεί

$$EX_{ij} = \alpha_i \beta_j \delta_{i+j}, \quad VX_{ij} = \varphi \alpha_i \beta_j \delta_{i+j}.$$

Με βάση αποκλειστικά δεδομένα από ένα run-off τρίγωνο πληρωμών, δεν μπορούμε να διακρίνουμε εάν οι αλλαγές κατά μήκος μιας διαγώνιας οφείλονται στο 1) ή 2). Για να ληφθεί υπόψη αυτή η πτυχή, ο Rietdorf (2008) εξετάζει μοντέλα

$$EX_{ij} = \mu_{ij} \quad \text{και} \quad VX_{ij} = \varphi \delta_{i+j} EX_{ij}. \quad (3.2.4)$$

βελτιστοποιώντας τις επεκτάσεις των συναρτήσεων πιθανοφάνειας Poisson που προτείνονται από Nelder και Pregibon (1987). Με παρόμοιο τρόπο διάφορες τροποποιήσεις της (3.2.4), συμπεριλαμβάνουν δομών όπως

$$VX_{ij} = \varphi \delta_{i+j}^c EX_{ij} \quad \text{και} \quad VX_{ij} = \varphi \delta_{i+j}^c (1 + kj) EX_{ij} \quad (3.2.5)$$

για  $c, k > 0$ , στην μελέτη μας. Οι διαφορετικές τροποποιήσεις του (3.2.4) στο (3.2.5) οδηγούν σε διαφορετικούς εκτιμητές των διαγώνιων επιδράσεων προκαλώντας σημαντικές τροποποιήσεις στην πρόβλεψη τελικού αποθέματος.

Το κίνητρο για την εργασία που παρουσιάζεται σε αυτή την εργασία προέρχεται από τις μελέτες δεδομένων των τριγώνων εξέλιξης αποζημιώσεων. βλ. Rietdorf (2008). Το συμπέρασμα που βασίζεται ο Rietdorf (2008) είναι ότι είναι διαισθητικά δύσκολο να κατανοήσουμε και να εξηγήσουμε τις εκτιμώμενες διαγώνιες επιδράσεις που προέρχονται από το (3.2.4). Όταν προστίθενται διαγώνιες επιδράσεις στη λειτουργία του Γενικευμένου Γραμμικού Μοντέλου έχουμε ότι

$$EX_{ij} = \alpha_i \beta_j, \quad VX_{ij} = \varphi \alpha_i \beta_j,$$

κάποιος υπερβολικά υπερπαραμετροποιεί αυτό το μοντέλο, το οποίο μπορεί να προκαλέσει αναξιόπιστες εκτιμήσεις παραμέτρων στο πλαίσιο των διαθέσιμων δεδομένων σε ένα τρίγωνο εξέλιξης αποζημιώσεων. Η αβεβαιότητα εκτίμησης οδηγεί σε αναξιόπιστη πρόβλεψη αποθεμάτων που στην πράξη αποκλείει αυτά τα μοντέλα για αναλογιστικούς σκοπούς.

Επομένως, είναι φυσικό να ψάξουμε σε αυτά τα μοντέλα συμπεριλαμβανομένων των διαγώνιων επιδράσεων με λιγότερες παραμέτρους. Οι Taylor (1977, 2000) και Verbeek (1972) θεωρούν την πρώτη κατηγορία δομής ενός τέτοιου μοντέλου. Αυτό το άρθρο προτείνει δύο διαφορετικές παραλλαγές των μοντέλων που προτείνονται στο Taylor (1977, 2000) και στο Verbeek (1972) όπου έχουν προστεθεί δομές δεύτερης κατηγορίας, μια παραλλαγή οδηγεί στις ίδιες εκτιμήσεις αποθεμάτων με τους Taylor και Verbeek, η άλλη δίνει διαφορετικές εκτιμήσεις αποθεμάτων μέσω ενός μοντέλου σταθμισμένης αξιοπιστίας. Στο πλαίσιο μας ο υπολογισμός των σταθμών αξιοπιστίας καθίσταται δυνατός επειδή μπορούμε να αναλύσουμε τη δομή δεύτερης κατηγορίας. Όπως φαίνεται στην Παράγραφο 3, τα αποτελέσματά μας μπορούν να ερμηνευθούν ως την μέθοδο Benktander με διαγώνιες επιδράσεις. Η δομή δεύτερης κατηγορίας επιτρέπει περαιτέρω μελέτες προσομοίωσης της αβεβαιότητας των προβλέψεων αποθεμάτων. βλ. Παράγραφο 4.

Επιμένουμε ότι οι τύποι των μοντέλων παραμένουν τόσο απλοϊκοί όσο (4) και οι διαδικασίες εκτίμησης που εφαρμόζονται εύκολα στα δεδομένα. Μια μελέτη δεδομένων δείχνει ότι οι εκτιμώμενες διαγώνιες επιδράσεις είναι αρκετά ελκυστικές.

Το κεφάλαιο είναι οργανωμένο ως εξής. Στην Παράγραφο 2 παρουσιάζεται μια μέθοδος τύπου Bornhuetter-Ferguson, συμπεριλαμβανομένων διαγώνιων επιδράσεων. Στην Παράγραφο 3 παρουσιάζουμε ένα μοντέλο αξιοπιστίας με διαγώνιες επιδράσεις

που τελικά δίνει μια πρόβλεψη βάσει ενός σταθμισμένου μέσου όρου ενός εκτιμητή τύπου Chain Ladder και της μεθόδου Bornhuetter-Ferguson. Τέλος, στην Παράγραφο 4 δίνονται μελέτες δεδομένων.

### 3.3 Μέθοδος Bornhuetter – Ferguson συμπεριλαμβανομένων των διαγώνιων επιδράσεων

Σε αυτήν την παράγραφο θα εξετάσουμε ένα μοντέλο βασισμένο στη μέθοδο Bornhuetter-Ferguson. Βλ. Bornhuetter and Ferguson (1972). Ωστόσο, θα χρησιμοποιήσουμε το πλαίσιο των Schmidt και Zocher (2008) καθώς θεωρούν τις επιδράσεις των γραμμών ως γνωστές, οδηγώντας σε μείωση του αριθμού των παραμέτρων. Αυτή η προσέγγιση είναι προτιμότερη επειδή σχετίζεται στενά με την Chain Ladder του Mack (1991) ή Hachemeister και Standard (1975) όπως αναφέρεται στην εισαγωγή.

Έστω το  $X_{ij}$  να είναι οι καταχωρήσεις σε ένα τρίγωνο εξέλιξης αποζημιώσεων με διάσταση  $m$ :

$$\Delta_m = \{X_{ij} : (i, j) \in Am\}$$

όπου

$$Am = \{(i, j) \in \mathbb{N} \times \mathbb{N}_0 : 1 \leq i + j \leq m\}.$$

Γράφουμε

$$Bm = \{(i, j) \in \mathbb{N} \times \mathbb{N}_0 : 1 \leq m, j \leq m - 1\}.$$

Ενδιαφερόμαστε για την πρόβλεψη του

$$X_{ij} : (i, j) \in Bm \setminus Am$$

που είναι μη παρατηρούμενες τυχαίες μεταβλητές σε χρόνο  $m$ . Ως τεχνική βάση για την πρόβλεψη, θεωρούμε ένα μοντέλο για όλες τις τυχαίες μεταβλητές,  $X_{ij}, (i, j) \in Bm$ , που δίδονται από τις ακόλουθες απαιτήσεις.

- i. τα  $X_{ij}, (i, j) \in Bm$  είναι αμοιβαία εξαρτημένα
- ii. Για  $(i, j) \in Bm$

$$EX_{ij} = T_i \beta_j \delta_{i+j} \quad \text{και} \quad VX_{ij} = \varphi \delta_{i+j}^c EX_{ij}.$$

όπου τα καθαρά ασφάλιστρα  $T_i > 0, 1 \leq i \leq m$ , θεωρείται ότι είναι γνωστά,  $c \in [0, 1]$  και

$$\varphi, \beta_0, \dots, \beta_{m-1} \quad \text{και} \quad \delta_1, \dots, \delta_{2m-1}$$



είναι θετικές άγνωστες σταθερές. Για τους σκοπούς της ταυτοποίησης υποθέτουμε επίσης ότι

$$\sum_{j=0}^{m-1} \beta_j = 1.$$

Η παράμετρος  $c \in [0,1]$  θεωρείται γνωστή. Συγκεκριμένα, θα επιλέξουμε  $c \in \{0,0.5,1\}$  για τους ακόλουθους λόγους:

- $c = 0$  αντιστοιχεί σε πληθωρισμό απαιτήσεων βλ. 2) σ. 2. Αυτό σημαίνει ότι οι διαγώνιες επιδράσεις είναι αθροιστικές που υποδηλώνουν ότι οι αλλαγές στις διαγώνιες προκαλούνται από αυξημένο αριθμό ή ανακοινωμένες απαιτήσεις (λόγω, για παράδειγμα, αλλαγών στη νομική πρακτική).
- $c = 1$  αντιστοιχεί στον οικονομικό πληθωρισμό. βλ. 1) σ. 2. Εάν δεν υπήρξαν μεγάλες αλλαγές στα μοτίβα αναφοράς και οι μόνες διαγώνιες επιδράσεις στα δεδομένα οφείλονται στον οικονομικό πληθωρισμό (ο οποίος συνήθως λειτουργεί με πολλαπλασιαστικό τρόπο).
- $c = 1/2$  επιλέγεται εάν βρισκόμαστε σε μια κατάσταση όπου και οι δύο επιδράσεις που περιγράφονται παραπάνω έχουν αντίκτυπο στα δεδομένα.

Η συγκεκριμένη επιλογή του  $c$  μπορεί να βασιστεί στη διαίσθηση καθώς και σε γραφήματα υπολειμμάτων. δείτε τον ορισμό του  $\hat{e}_{ij}$  στο (3.3.5).

Παρατηρήστε ότι αν υποθέσουμε ότι τα  $T_i, 1 \leq i \leq m$ , είναι άγνωστα και  $\delta_j = 1$ , τότε το μοντέλο i), ii) θα ήταν πανομοιότυπο με τη ρύθμιση του Γενικευμένου Γραμμικού Μοντέλου στη Chain Ladder, δείτε για παράδειγμα England και Verrall (1999). Υποθέτοντας ότι τα  $T_i, 1 \leq i \leq m$ , γνωστά και  $\delta_j = 1$ , i), ii) είναι πανομοιότυπο με την κατάσταση από τους Schmidt and Zocher (2008). Εάν τα  $T_i, 1 \leq i \leq m$ , είναι προηγούμενες αναμενόμενες τιμές των προεξοφλημένων συσσωρευμένων απαιτήσεων, τότε το  $\delta_j, 1 \leq j \leq m$ , μπορεί να συγκριθεί άμεσα με εξωτερικά δεδομένα σε δείκτες πληθωρισμού. Εάν τα  $T_i, 1 \leq i \leq m$ , είναι προσαρμοσμένα στον πληθωρισμό, τότε  $\delta_j, 1 \leq j \leq m$ , μπορεί να θεωρηθεί ως πληθωρισμός σε αυτό το χαρτοφυλάκιο μείον τον εκτιμώμενο εκ των προτέρων πληθωρισμό.

Εκτιμούμε τις παραμέτρους στο μοντέλο i), ii) σε μια διαδικασία δύο βημάτων.

Πρώτον, χρησιμοποιούμε τη μέθοδο διαχωρισμού, βλέπε Taylor (1977, 2000) και Verbeek (1972). Αυτή η μέθοδος βασίζεται στην ακόλουθη παρατήρηση όπου ένας μετασχηματισμός κάνει τις διαγώνιες να γίνουν επιδράσεις γραμμών :

$$\tilde{X}_{ij} = \frac{X_{i+jj}}{T_{i+j}} \quad \text{τότε} \quad E\tilde{X}_{ij} = \delta_i \beta_j \quad (3.3.1)$$

για  $1 \leq i \leq m$  και  $0 \leq j \leq i - 1$ . Η γνωστή πολλαπλασιαστική δομή (3.3.1) υποδηλώνει ότι η εκτίμηση μπορεί να βασιστεί στο συνολικό περιθωριακό κεφάλαιο (βλέπε για παράδειγμα Schmidt και Wünsch (1998)).

$$\hat{\delta}_i \sum_{j=0}^{i-1} \hat{\beta}_j = \sum_{j=0}^{i-1} \bar{X}_{ij}, \quad 1 \leq i \leq m, \quad \text{και} \quad \hat{\beta}_j \sum_{i=j+1}^m \hat{\delta}_i = \sum_{i=j+1}^m \bar{X}_{ij}, \quad 0 \leq j \leq m - 1.$$

Οι Schmidt και Wünsch επισημαίνουν ότι το συνολικό βασικό περιθώριο αποδίδει μια μέθοδο Chain Ladder - εδώ υπό τον περιορισμό  $\sum_{j=0}^{m-1} \beta_j = 1$ . Επομένως, η διαδικασία εκτίμησης μπορεί να πραγματοποιηθεί χρησιμοποιώντας τον ακόλουθο υπολογισμό. Πρώτο

$$\sum_{j=0}^{m-1} \hat{\beta}_j = 1. \quad (3.3.2)$$

Στη συνέχεια, χρησιμοποιούμε τους λόγους

$$\frac{\sum_{j=0}^k \hat{\beta}_j}{\sum_{j=0}^{k-1} \hat{\beta}_j} = \frac{\sum_{j=0}^k \sum_{i=k+1}^m \bar{X}_{ij}}{\sum_{j=0}^{k-1} \sum_{i=k+1}^m \bar{X}_{ij}}, \quad 1 \leq k \leq m - 1, \quad (3.3.3)$$

να υπολογίσουμε αναδρομικά τους εκτιμητές  $\hat{\beta}_0, \dots, \hat{\beta}_{m-1}$  ξεκινώντας από (3.3.2). Τέλος, υπολογίζουμε

$$\hat{\delta}_i = \frac{\sum_{j=0}^{i-1} \bar{X}_{ij}}{\sum_{j=0}^{i-1} \hat{\beta}_j}, \quad 1 \leq i \leq m. \quad (3.3.4)$$

Το δεύτερο μέρος της διαδικασίας εκτίμησης σχετίζεται με τη δομή διακύμανσης. Ορίζεται

$$\hat{\epsilon}_{ij} = \frac{\bar{X}_{ij} - T_i \hat{\beta}_j \hat{\delta}_{i+j-1}}{(T_i \hat{\beta}_j \hat{\delta}_{i+j-1}^{1+c})^{1/2}}, \quad (i, j) \in Am. \quad (3.3.5)$$

Η εκτίμηση του  $\phi$  πραγματοποιείται χρησιμοποιώντας τη μέθοδο των ρολών:

$$\hat{\phi} = \frac{2}{m(m+1)} \sum_{i=1}^m \sum_{j=0}^{m-i} \hat{\epsilon}_{ij}^2.$$

Ένα άλλο κίνητρο για τη μορφή των εκτιμητών (3.3.2) - (3.3.4) είναι ότι οι εκτιμητές είναι στην πραγματικότητα συνεπείς με την ακόλουθη έννοια.

### Πρόταση 3.1

Ας υποθέσουμε ότι το μοντέλο i), ii). Εάν  $T_i \rightarrow \infty$  για όλα τα  $1 \leq i \leq m$  τότε

$$\hat{\delta}_{j+1} \xrightarrow{P} \delta_{j+1} \quad \text{και} \quad \hat{\beta}_j \xrightarrow{P} \beta_j$$

για  $0 \leq j \leq m - 1$ .

### Απόδειξη:

Από τις παραδοχές i) και ii) και από την ανισότητα του Chebyshev είναι απόρροια

$$\frac{X_{ij}}{T_i} \xrightarrow{P} \beta_j \delta_{i+j}$$

ως  $T_i \rightarrow \infty$ . Το αναδρομικό σχήμα (3.3.3) και (3.3.2) μαζί με το θεώρημα συνεχούς χαρτογράφησης δίνουν το επιθυμητό αποτέλεσμα.  $\square$

Καθώς τα  $X_{ij}$ ,  $(i, j) \in Bm$ , θεωρούνται ανεξάρτητα, είναι φυσικό να προβλέψουμε τα  $X_{ij}$ ,

$(i, j) \in Bm \setminus Am$ , από την αναμενόμενη τιμή:

$$\mu(X_{ij}) = EX_{ij} = T_i \beta_j \delta_{i+j}.$$

Οι παράμετροι,  $\delta_j$ ,  $0 \leq j \leq m - 1$ , υπολογίζονται απευθείας από δεδομένα χρησιμοποιώντας αναδρομή (3.3.4). Οι διαγώνιες παράμετροι,  $\delta_j$ ,  $m + 1 \leq j \leq 2m - 1$ , σχετίζονται με μη παρατηρήσιμα δεδομένα,  $X_{ij}$ ,  $(i, j) \in Bm \setminus Am$  και επομένως αυτές δεν μπορούν να εκτιμηθούν απευθείας χρησιμοποιώντας (3.3.4). Δεδομένου ότι ο τελευταίος εκτιμητής,  $\mu(X_{ij})$ , εξαρτάται αποκλειστικά από διαγώνιες παραμέτρους  $\delta_j$ ,  $m + 1 \leq j \leq 2m - 1$ , πρέπει ωστόσο να είμαστε σε θέση να εκτιμήσουμε ή να τις προβλέψουμε. Η ιδέα είναι να κατασκευαστεί ένας εκτιμητής των  $\delta_j$ ,  $m + 1 \leq j \leq 2m - 1$ , με βάση τις διαγώνιες του δείγματος,  $\delta_1, \dots, \delta_m$ . Στη συνέχεια, προτείνεται ένας τρόπος δημιουργίας μιας πρόβλεψης και προσθέτουμε επίσης μια παρατήρηση για έναν πιο θεωρητικό χαρακτήρα που δικαιολογεί αυτήν την προσέγγιση.

### Παράδειγμα 3.1

Ας υποθέσουμε ότι  $\hat{\delta}_j$ ,  $j \geq 1$ , ακολουθήστε μια διαδικασία  $AR(1)$ , δηλαδή, ότι

$$\hat{\delta}_j \Gamma + \epsilon_j = \hat{\delta}_{j+1},$$

όπου  $\epsilon_j$ ,  $1 \leq j \leq 2m - 1$ , είναι μια μέση διαδικασία μηδέν ένα (μηδενικού λευκού θορύβου). Με βάση τους εκτιμητές διαγώνιας παραμέτρου στο δείγμα εκτιμούμε ότι  $\Gamma$  με

$$\hat{\Gamma} = \frac{\sum_{j=1}^{m-1} \hat{\delta}_{j+1} \hat{\delta}_j}{\sum_{j=1}^{m-1} \hat{\delta}_j^2}.$$

Η υπόθεση  $AR(1)$  υποδηλώνει μια πρόβλεψη του  $\hat{\delta}_j$ ,  $j \geq m + 1$ , που δίνεται από

$$\tilde{\delta}_j = \hat{\delta}_m \hat{\Gamma}^{j-m}.$$

Κατά συνέπεια, η πρόβλεψη των  $X_{ij}$ ,  $(i, j) \in Bm \setminus Am$ , δίνεται από

$$\tilde{\mu}(X_{ij}) = T_i \hat{\beta}_j \tilde{\delta}_{i+j}.$$

□

### Παρατήρηση 3.1

Η προσέγγιση της πρόβλεψης που ακολουθείται στο Παράδειγμα 3.1 υποδεικνύει ότι το  $\delta_j$ ,  $j \geq 1$ , θα πρέπει να θεωρηθεί ως χρονοσειρά και παρά σαν καθορισμένες παραμέτρους. Τυπικά, αυτό υποδηλώνει τις ακόλουθες εναλλακτικές παραδοχές μοντέλου:

- i. υποθέτουμε ότι τα  $X_{ij}$ ,  $(i, j) \in Am$ , είναι υπό όρους ανεξάρτητα δεδομένου  $\delta_j$ ,  $j \geq 1$ , και
- ii. το εξής

$$E(X_{ij} | \delta_{i+j-1}) = T_i \beta_j \delta_{i+j}$$

$$V(X_{ij} | \delta_{i+j-1}) = \varphi \delta_{i+j}^c E X_{ij}$$

για  $(i, j) \in Bm$ .

Σε αυτήν την μορφή η πρόταση 1 παραμένει έγκυρη. Στην πραγματικότητα από την πρόταση 3.1 προκύπτει ότι

$$P\left(\left|\frac{X_{ij}}{T_i \beta_j \delta_{i+j}} - 1\right| > \epsilon \mid \delta_{i+j}\right) \rightarrow 0, \quad T_i \rightarrow \infty,$$

και η σύγκλιση τείνει σε αυτό

$$\frac{X_{ij}}{T_i} \xrightarrow{P} \beta_j \delta_{i+j}, \quad T_i \rightarrow \infty.$$

Με τον ίδιο τρόπο όπως στην απόδειξη της Πρότασης 3.1 προκύπτει τώρα ότι εάν  $T_i \rightarrow \infty$  για όλα τα  $1 \leq i \leq m$  τότε

$$\hat{\delta}_{j+1} \xrightarrow{P} \delta_{i+j}, \quad \hat{\beta}_j \xrightarrow{P} \beta_j, \quad 0 \leq j \leq m-1,$$

όπου οι τελευταίοι εκτιμητές δίδονται από το (3.3.2).

Επιπλέον, εάν θέλουμε να προβλέψουμε ελαχιστοποιώντας το μέσο τετράγωνο σφάλμα,

$$E[X_{ij} - \mu(X_{ij})]^2$$

πάνω από όλες τις μετρήσιμες συναρτήσεις διακύμανσης,  $\mu(X_{ij})$ , με  $\Delta_m, \delta_1, \dots, \delta_m$ , τότε λαμβάνουμε την πρόβλεψη

$$\mu(X_{ij}) = T_i \beta_j E(\delta_{i+j} | \delta_1, \dots, \delta_m).$$

Δεδομένου ότι τα  $\delta_1, \dots, \delta_m$  δεν είναι παρατηρήσιμα, πρέπει να τα αντικαταστήσουμε με τις ψευδο-παρατηρήσεις  $\hat{\delta}_1, \dots, \hat{\delta}_m$ , οδηγώντας τελικά στην ίδια πρόβλεψη όπως προτείνεται στο Παράδειγμα 3.1.  $\square$

Για να εκτιμήσουμε την αβεβαιότητα στο μοντέλο i), ii) μπορούμε να εφαρμόσουμε μια μέθοδο bootstrap παρόμοια με αυτήν από τους England και Verrall (1999). Η μέθοδος bootstrap θα πρέπει να βασίζεται στην υπόθεση ότι τα υπόλοιπα  $\hat{e}_{ij}$ ,  $(i, j) \in Am$ , είναι τέτοια ώστε να μπορούν να αναπαράγονται στο δείγμα  $B$  φορές,  $e_{ij}^*(k)$ ,  $1 \leq k \leq B$ , για τη δημιουργία ανεξάρτητων και ομοιόμορφα κατανεμημένων μορφών του  $\Delta_m$  με

$$X_{ij}(k) = T_i \hat{\beta}_j \hat{\delta}_{i+j} + e_{ij}^*(k) [\hat{\phi} T_i \hat{\beta}_j \hat{\delta}_{i+j}^{1+c}]^{\frac{1}{2}}, \quad 1 \leq k \leq B. \quad (3.3.6)$$

Σε αυτήν την ενότητα, οι επιδράσεις των γραμμών,  $T_i$ ,  $1 \leq i \leq m$ , θεωρείται ότι δίνουν μεγάλη σημασία στις εκ των προτέρων εκτιμήσεις  $T_i$ ,  $1 \leq i \leq m$ , των συνολικών αποζημιώσεων των ασφαλιστικών ετών. Στην επόμενη ενότητα εξετάζουμε ένα μοντέλο αξιοπιστίας που οδηγεί σε εκτιμητές της επιδράσεις των γραμμών βάσει ενός σταθμισμένου μέσου όρου των  $T_i$  και ενός εκτιμητή τύπου Chain Ladder των συνολικών αποζημιώσεων.

### 3.4 Ένα μοντέλο αξιοπιστίας που περιλαμβάνει διαγώνιες επιδράσεις

Σε αυτήν την ενότητα θεωρούμε ένα μοντέλο αξιοπιστίας σε ένα παρόμοιο πλαίσιο όπως στην Παράγραφο 3 αλλά με τυχαίες επιδράσεις γραμμών. Σε μια ειδική περίπτωση ( $c = 1$ ) αυτή η προσέγγιση αποδίδει προβλέψεις βάσει ενός σταθμισμένου μέσου όρου των εκτιμήσεων Chain Ladder (βλ. Για παράδειγμα Mack (1991,1993,1994)) και εκτιμήσεις Bornhuetter Ferguson που δίνονται στην Παράγραφο 3. Αυτός ο τύπος αποτελεσμάτων βρίσκεται επίσης στους Verrall (2004) και Wüthrich (2007) όπου δεν περιλαμβάνονται διαγώνιες επιδράσεις.

Χρησιμοποιώντας την ίδια σημείωση όπως στην Παράγραφο 3, υποθέτουμε

- i.  $\theta_i$ ,  $1 \leq i \leq m$ , αμοιβαία ανεξάρτητοι με  $E\theta_i = T_i$  και  $V\theta_i = \xi^{-1}T_i$ ,
- ii. δεδομένου  $\theta_i$ ,  $1 \leq i \leq m$ ,  $X_{ij}$ ,  $(i, j) \in Bm$ , είναι αμοιβαία ανεξάρτητα και
- iii. για  $(i, j) \in Bm$

$$E(X_{ij}|\theta_i) = \theta_i \beta_j \delta_{i+j}, \quad V(X_{ij}|\theta_i) = \varphi \delta_{i+j}^c E(X_{ij}|\theta_i)$$

όπου τα καθαρά ασφάλιστρα  $T_i > 0, 1 \leq i \leq m$ , θεωρούνται γνωστά και

$$\varphi, \xi, \beta_0, \dots, \beta_{m-1}, \delta_1, \dots, \delta_{2m-1}$$

είναι θετικές άγνωστες σταθερές. Για να αναγνωρισθεί το μοντέλο έχουμε

$$\sum_{j=0}^{m-1} \beta_j = 1.$$

Αυτό οδηγεί απευθείας στο ακόλουθο αποτέλεσμα.

### Πρόταση 3.2

Ας υποθέσουμε ότι τα i) -iii) ισχύουν. Στη συνέχεια, ο καλύτερος γραμμικός εκτιμητής Bayes του  $\theta_i$  δίνεται από

$$\mu^{LB}(\theta_i) = T_i \omega_i + (1 - \omega_i) \bar{X}_i$$

όπου

$$\bar{X}_i = \frac{\sum_{j=0}^{m-i} \delta_{i+j}^{-c} X_{ij}}{\sum_{j=0}^{m-i} \beta_j \delta_{i+j}^{1-c}} \quad \text{και} \quad \omega_i = (1 + (\varphi \xi)^{-1} \sum_{j=0}^{m-i} \beta_j \delta_{i+j}^{1-c})^{-1}.$$

Η καλύτερη γραμμική πρόβλεψη Bayes του  $X_{ij}$  δίνεται επιπλέον από

$$\mu(X_{ij}) = \mu^{LB}(\theta_i) \beta_j \delta_{i+j}.$$

### Απόδειξη

Η απόδειξη της Πρότασης 3.2 προκύπτει απευθείας από μια εφαρμογή του μοντέλου Bühlmann Straub (βλέπε για παράδειγμα Mikosch (2006) σελ. 213-215)

$$Z_{ij} = \frac{X_{ij}}{\beta_j \delta_{i+j}}, \quad 1 \leq j \leq m - i.$$

Το μοντέλο Bühlmann-Straub προϋποθέτει σταθερές και γνωστές παραμέτρους που δεν συμβαίνουν στα παραπάνω, αλλά αν οι παράμετροι έχουν ρυθμιστεί ή όχι δεν επηρεάζει το πρόβλημα βελτιστοποίησης που εξετάζεται. Το αποτέλεσμα ισχύει άμεσα.  $\square$

### Παρατήρηση 3.2

Η ειδική περίπτωση  $c = 1$  είναι ιδιαίτερα ενδιαφέρουσα. Σε αυτήν την περίπτωση

$$\bar{X}_i = \frac{\sum_{j=0}^{m-i} \delta_{i+j}^{-1} X_{ij}}{1 - \sum_{j=0}^{m-i} \beta_j},$$

όπου μια ερμηνεία είναι ότι το  $\sum_{j=0}^{m-i} \delta_{i+j}^{-1} X_{ij}$  είναι ένα άθροισμα των προεξοφλημένων αποζημιώσεων πολλαπλασιαζόμενο με έναν παράγοντα της Chain Ladder. Στην πραγματικότητα, εάν κάποιος υπολογίσει τους παράγοντες της Chain Ladder,  $\hat{f}_0, \dots, \hat{f}_{m-2}$ , βλ. Mack (1991), με βάση τις συσσωρευμένες προεξοφλημένες αποζημιώσεις

$$C_{ij} = \sum_{j=0}^{m-i} \delta_{i+j}^{-1} X_{ij}, (i, j) \in Am,$$

τότε  $(1 - \sum_{j=m-i+1}^{m-1} \hat{\beta}_j)^{-1}$  και  $\prod_{j=m-i}^{m-2} \hat{f}_j$  είναι δύο εκτιμητές της σχετικής αύξησης από  $j = m - i + 1$  σε  $j = m - 1$ . Δεδομένου ότι ο  $C_{ij}$  καθορίζει μια ρύθμιση Γενικευμένου Γραμμικού Μοντέλου όπως στον England κ.ά. (1999), μια πιθανοφάνεια Poisson σχεδόν ίδια με αυτή του Mack (1991) και του Hachemeister (1975) μπορεί να χρησιμοποιηθεί για την εκτίμηση των παραμέτρων. Αυτό τελικά οδηγεί σε πρόβλεψη με τη μέθοδο Chain Ladder,

$$C_{ij} \prod_{j=m-i}^{m-2} \hat{f}_j = \left[ \sum_{j=0}^{m-i} \delta_{i+j}^{-1} X_{ij} \right] \prod_{j=m-i}^{m-2} \hat{f}_j.$$

Μια ματιά στην πρόταση 3.2 εξηγεί τώρα πώς μπορεί κανείς να ερμηνεύσει τον εκτιμητή Bayes,  $\mu^{LB}(\theta_i)$ , ως σταθμισμένο μέσο όρο μεταξύ των εκτιμητών της Chain Ladder και των εκτιμητών Bornhuetter-Ferguson.  $\square$

Με τον ίδιο τρόπο όπως στην Παράγραφο 3, η εκτίμηση πραγματοποιείται σε διαδικασία δύο βημάτων. Στο πρώτο βήμα οι παράμετροι  $\delta_{i+j}$ ,  $\beta_j$ ,  $0 \leq j \leq m - 1$ , υπολογίζονται με τις αναδρομικές διαδικασίες (3.3.2) και (3.3.4). Ένα κίνητρο για επαναχρησιμοποίηση της διαδικασίας εκτίμησης από την Παράγραφο 3 είναι ότι η Πρόταση 3.1 εξακολουθεί να ισχύει υπό τις προϋποθέσεις i) -iii).

Ο εκτιμητής Bayes που δίνεται στην πρόταση 3.2 εξαρτάται από παραμέτρους που είναι άγνωστες. Αυτές οι άγνωστες παράμετροι πρέπει να εκτιμηθούν για να παράγουν τους πραγματικούς εκτιμητές σημείων. Η σύνδεση σε εκτιμητές για τις παραμέτρους στην πρόταση 3.2 αποδίδει μια εμπειρική διαδικασία Γραμμικής Bayes.

**Πρόταση 3.3**

Ας υποθέσουμε ότι το i) -iii) κρατά. Εάν  $T_i \rightarrow \infty$  για όλα τα  $1 \leq i \leq m$  τότε

$$\hat{\delta}_{j+1} \xrightarrow{P} \delta_{i+j} \quad \text{και} \quad \hat{\beta}_j \xrightarrow{P} \beta_j \quad \text{για} \quad 0 \leq j \leq m - 1.$$

**Απόδειξη :**

Βάζουμε  $Y_{ij} = X_{ij}/T_i$ . Από την ανισότητα του Chebyshev ισχύει

$$P(|Y_{ij} - \beta_j \delta_{i+j}| > \epsilon | \theta_i) \leq \frac{E\left([X_{ij} - T_i \beta_j \delta_{i+j}]^2 | \theta_i\right)}{T_i^2 \epsilon^2}$$

$$\frac{E\left([X_{ij} - T_i \beta_j \delta_{i+j}]^2 | \theta_i\right)}{T_i^2 \epsilon^2} \leq \frac{E(X_{ij} | \theta_i) + (T_i \beta_j \delta_{i+j} - \theta_i \beta_j \delta_{i+j})^2}{T_i^2 \epsilon^2}.$$

Λαμβάνοντας προσδοκίες και από τις δύο πλευρές, επιτυγχάνουμε την ανισότητα

$$P(|Y_{ij} - \beta_j \delta_{i+j}| > \epsilon) \leq \frac{\varphi \beta_j \delta_{i+j}^{1+c} T_i + (\beta_j \delta_{i+j})^2 \frac{T_i}{\xi}}{T_i^2 \epsilon^2},$$

όπου η δεξιά πλευρά πηγαίνει στο μηδέν όταν  $T_i \rightarrow \infty$ . Το αναδρομικό σχήμα (3.3.4) μαζί με το θεώρημα συνεχούς χαρτογράφησης αποδίδει τώρα το επιθυμητό αποτέλεσμα.  $\square$

Το δεύτερο βήμα στη διαδικασία εκτίμησης βασίζεται στον ορισμό των υπολειμμάτων της σειράς

$$e_i = \frac{\sum_{j=0}^{m-i} X_{ij} - T_i \sum_{j=0}^{m-i} \beta_j \delta_{i+j}}{\left(T_i \sum_{j=0}^{m-i} \beta_j \delta_{i+j}\right)^{1/2}}, \quad 1 \leq i \leq m,$$

όπου

$$E e_i^2 = \frac{1}{\xi} \left( \sum_{j=0}^{m-i} \beta_j \delta_{i+j} \right) + \varphi \frac{\sum_{j=0}^{m-i} \beta_j \delta_{i+j}^{1+c}}{\sum_{j=0}^{m-i} \beta_j \delta_{i+j}}.$$

Αντικαθιστώντας τις παραμέτρους  $\beta_j$ ,  $\delta_{j+1}$ ,  $0 \leq j \leq m-1$ , από τους εκτιμητές τους μπορεί να χρησιμοποιηθεί μια μέθοδος ελάχιστων τετραγώνων βάσει της τελευταίας προσδοκίας για την εκτίμηση των  $c, \varphi$  και  $\xi$ . Στην περίπτωση  $c = 0$  η τελευταία διαδικασία μπορεί να μειωθεί σε μια συνηθισμένη γραμμική παλινδρόμηση  $e_i^2$  έναντι  $\sum_{j=0}^{m-i} \beta_j \delta_{i+j}$ ,  $1 \leq i \leq m$ . Εάν  $c = 1$  τότε  $\xi$  και  $\varphi$  μπορούν επίσης να εκτιμηθούν με μια γραμμική παλινδρόμηση με βάση τη συσχέτιση

$$\frac{e_i^2}{\sum_{j=0}^{m-i} \beta_j \delta_{i+j}} \cong \frac{1}{\xi} + \varphi \frac{\sum_{j=0}^{m-i} \beta_j \delta_{i+j}^2}{\left(\sum_{j=0}^{m-i} \beta_j \delta_{i+j}\right)^2}, \quad 1 \leq i \leq m.$$



Δεδομένου ότι υπάρχουν μόνο  $m$  γραμμών υπολείμματα,  $e_i$ , μία ή δύο ακραίες τιμές στο δείγμα μπορούν να επηρεάσουν πολύ την παραπάνω παλινδρόμηση. Συνιστάται επομένως να σχεδιαστεί η αναμενόμενη τιμή έναντι της παρατηρούμενης και, σε ορισμένες περιπτώσεις, χρησιμοποιούμε υποκειμενική κρίση για να διορθώσουμε την εκτίμηση.

Για να προσομοιώσουμε τις ανεξάρτητες και ομοιόμορφα κατανεμημένες μορφές των  $\Delta_m$  στο μοντέλο i) -iii) μπορούμε επίσης να εφαρμόσουμε μια μέθοδο bootstrap παρόμοια με αυτή των England και Verrall (1999). Η μέθοδος bootstrap πρέπει να βασίζεται στην υπόθεση ότι τα υπολείμματα

$$e_{ij}^* = \frac{x_{ij} - \mu^{LB}(\theta_i)\beta_j\delta_{i+j}}{(\varphi\mu^{LB}(\theta_i)\beta_j\delta_{i+j}^{1+c})^{1/2}}, \quad (i, j) \in Am, \quad (3.4.1)$$

είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα ώστε να μπορούν να επιλεγθούν ξανά στο δείγμα.

Στην επόμενη Παράγραφο τα μοντέλα που προτείνονται στις παραγράφους 3 και 4 εφαρμόζονται σε ένα σύνολο δεδομένων.

### 3.5 Έρευνα σε δεδομένα

Σε αυτήν την Παράγραφο εξετάζεται ένα σύνολο δεδομένων από την Codan Insurance. Τα δεδομένα περιέχουν 13 χρόνια λειτουργίας για ένα χαρτοφυλάκιο τρίτου προσώπου για ασφάλιση αυτοκινήτων. Αυτό το τρίγωνο εξέλιξης δίνεται παρακάτω στον Πίνακα 3.1.

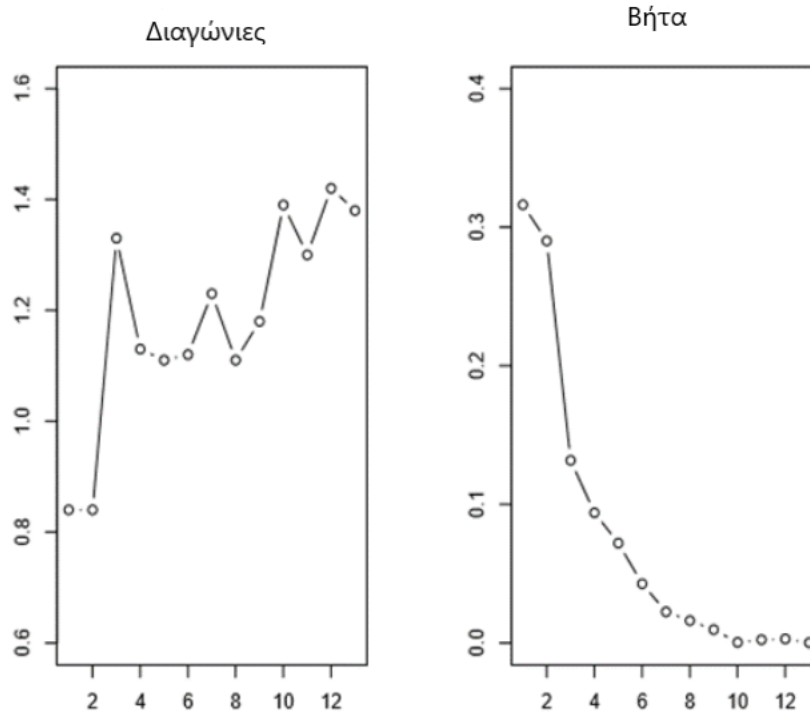
$i \setminus j$	0	1	2	3	4	5	6	7	8	9	10	11	12
1	22564	17331	17377	7723	5058	2530	1443	1195	1889	106	33	139	14
2	22901	26734	8974	7089	3116	1911	3284	1591	879	21	575	476	
3	36152	26513	10973	6714	7155	2176	1656	1094	-89	8	115		
4	34733	29642	13593	11496	6256	6404	3900	2157	1133	25			
5	30709	28020	12465	8504	9929	5592	910	3413	1428				
6	33727	32190	13318	9211	8129	5225	2149	773					
7	30727	27677	9251	9221	6169	7492	2952						
8	32498	35446	18432	15110	13990	4986							
9	32228	42937	16231	12942	11078								
10	41947	41634	21056	15442									
11	37247	34135	19061										
12	32891	29719											
13	35993												

Πίνακας 3.1. Τρίγωνο εξέλιξης

Στα μοντέλα που εξετάζονται στις Παραγράφους 3 και 4 θεωρείται ότι είναι γνωστά οι αναμενόμενες επιδράσεις κατά σειρά. Αυτά δίνονται παρακάτω στον Πίνακα 3.2.

$i$	1	2	3	4	5	6	7
$T_i$	85047	74409	86077	83082	83427	81557	79495
$i$	8	9	10	11	12	13	
$T_i$	101564	95482	107062	90091	85413	81995	

Πίνακας 3.2. Επιδράσεις γραμμών  $T_i, 1 \leq i \leq 13$ .



Σχήμα 3.1. Αριστερά Διαγώνιες Επιδράσεις  $\delta_j, 1 \leq j \leq 13$ , Δεξιά Το μοτίβο αποζημιώσεων  $\beta_j, 0 \leq j \leq 12$ .

Για την εφαρμογή μοντέλων i), ii) στην Παράγραφο 3 και i) –iii) στην Παράγραφο 4 στα δεδομένα οι Πίνακες 3.1 και 3.2 είναι όλα τα δεδομένα που χρειαζόμαστε. Η εκτίμηση των παραμέτρων  $\beta_j, \delta_j, 0 \leq j \leq 12$ , είναι ίδια στις Παράγραφο 3 και 4. Οι εκτιμητές φαίνονται στον Πίνακα 3.3 και στην Σχήμα 3.1.

$j$	0	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{\beta}_j$	0.32	0.29	0.13	0.09	0.07	0.04	0.02	0.02	0.01	0.00	0.00	0.00	0.00
$\hat{\delta}_{j+1}$	0.84	0.84	1.33	1.13	1.11	1.12	1.23	1.11	1.18	1.39	1.30	1.41	1.38

Πίνακας 3.3. Εκτιμητές  $\hat{\beta}_j, \hat{\delta}_{j+i}, 0 \leq j \leq 12$  όπως δίνονται στην (3.3.2).

Για να υπολογίσουμε την εκτίμηση του τελικού αποθέματος πρέπει να προβλέψουμε τις διαγώνιες επιδράσεις,  $\delta_{14}, \dots, \delta_{25}$ . Ακολουθούμε τη διαδικασία που προτείνεται στο Παράδειγμα 3.1. Η αναμενόμενη ετήσια αύξηση,  $\Gamma$ , εκτιμάται κατά

$$\hat{\Gamma} = 1.0245$$

καταλήγουμε στην πρόβλεψη των  $\delta_{14}, \dots, \delta_{25}$  που δίνονται από

$$\hat{\delta}_{13+j} = \hat{\delta}_{13} \hat{\Gamma}^j = 1.38 \times 1.0245^j \quad j \geq 1.$$

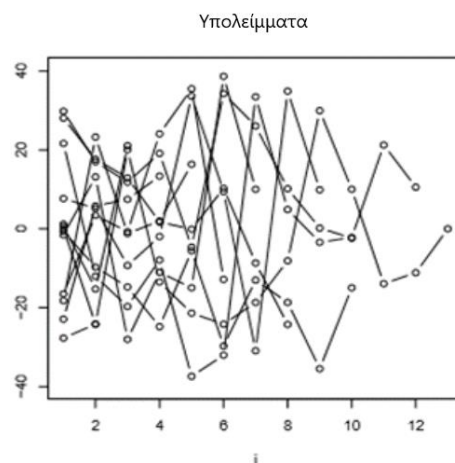
Επίσης, αν υποθέσουμε ότι  $\epsilon_j \sim \mathcal{N}(0, \sigma^2), j \geq 1$ , τότε

$$\hat{\sigma}^2 = (m-1)^{-1} \sum_{j=1}^{m-1} (\hat{\delta}_{j+1} - \hat{\Gamma} \hat{\delta}_j)^2 = 0.03037.$$

Συνεχίζουμε να εξετάσουμε λεπτομερώς το μοντέλο από την Παράγραφο 3.

### 3.5.1 Έρευνα σε δεδομένα : Μέθοδος Bornhuetter - Ferguson που περιλαμβάνει διαγώνιες επιδράσεις

Το πρώτο πράγμα που πρέπει να καθορίσουμε είναι η παράμετρος  $c \in \{0, 1/2, 1\}$ . Ένας τρόπος για να το κάνουμε αυτό είναι να εξετάσουμε τα υπολείμματα,  $\hat{\epsilon}_{ij}, (i, j) \in A_{13}$ , όπως δίδεται στην Παράγραφο 3. Για κάθε  $c = 0, 1/2, 1$  έχουμε σχεδιάσει σημεία  $(i, \hat{\epsilon}_{ij}), (i, j) \in Am$ , για να ελέγξουμε αν αυτά τα υπολείμματα "φαίνονται ανεξάρτητα και ομοιόμορφα καταναμημένα". Εάν έτσι φέρονται, μπορεί να ληφθούν ως δείκτης ότι το υπολογιζόμενο μοντέλο i), ii) μπορεί να πάρει δεδομένα. Κρίνοντας από τέτοια υπολείμματα σχεδιάζουμε  $c = 0$  δίνει το καλύτερο πρόβλεψη του μοντέλου, Σχήμα 3.2 παρακάτω για την περίπτωση  $c = 0$ . Με  $c = 0$  ο εκτιμητής ρολής του  $\varphi$  είναι  $\hat{\varphi} = 345.08$ . Προσδιορίζουμε τις εκτιμήσεις της σειράς αποθεμάτων κατά



Σχήμα 3.2. Γραμμικά υπολείμματα:  $\{j, \hat{e}_{ij}, 0 \leq j \leq 13 - i\}$  για  $1 \leq i \leq 12$  στην περίπτωση που  $c = 0$ .

$$\hat{R}_i^{(1)} = \sum_{j=m-i+1}^{m-1} \hat{\mu}(X_{ij}), \quad 2 \leq i \leq 13.$$

Χρησιμοποιώντας αυτήν την σημείωση, οι εκτιμητές αποθέματος δίνονται παρακάτω στον Πίνακα 3.4.

$i$	2	3	4	5	6	7	8	9	10	11	12	13
$\hat{R}_i$	12	364	631	690	1793	3589	7911	13377	26264	34599	49503	82308

Πίνακας 3.4. Εκτιμητές αποθέματος  $\hat{R}_i^{(1)}$ ,  $2 \leq i \leq 13$ .

### Μελέτη προσομοίωσης

Ο ορισμός των υπολειμμάτων μας επιτρέπει να ποσοτικοποιήσουμε την αβεβαιότητα που σχετίζεται με το μοντέλο i) ii) από την Παράγραφο 3 χρησιμοποιώντας μια μέθοδο bootstrap. Εξετάσαμε το σφάλμα διαδικασίας που καθορίστηκε με την ακόλουθη διαδικασία προσομοίωσης. Με τον ίδιο τρόπο όπως στην (3.3.6) δημιουργούμε ανεξάρτητες και ομοιόμορφα κατανομημένες πραγματοποιήσεις  $X_{ij}$ ,  $(i, j) \in Bm \setminus Am$ , με

$$X_{ij}(k) = T_i \hat{\beta}_j \tilde{\delta}_{i+j}(k) + e_{ij}^*(k) [\hat{\varphi} T_i \hat{\beta}_j \tilde{\delta}_{i+j}^{1+c}(k)]^{\frac{1}{2}}, \quad 1 \leq k \leq B,$$

όπου  $\tilde{\delta}_j(k), j \geq 1$ , προσομοιώνεται ανεξάρτητα από  $e_{ij}^*(k)$ ,  $(i, j) \in Bm \setminus Am$ , χρησιμοποιώντας την αναδρομή που δίνεται στο Παράδειγμα 3.1 με  $\epsilon_j(k)$ ,  $j \geq 1$ , ανεξάρτητα και ομοιόμορφα κατανομημένα με μέση μηδέν και διακύμανση  $\sigma^2$ .

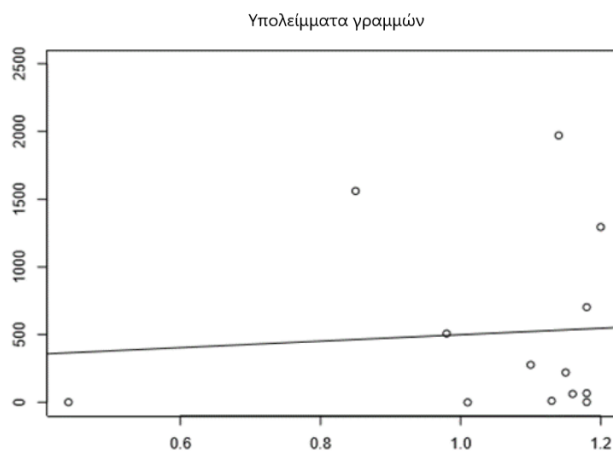
Ο Πίνακας 3.5 παρακάτω δείχνει ορισμένα σχετικά στατιστικά στοιχεία των 50000 προσομοιώσεων.

$i$	$SD$	$Qt(95\%)$	$ES(95\%)$
2	66	130	12
3	358	991	99
4	476	1446	194
5	505	1558	218
6	829	3216	355
7	1225	5697	554
8	2023	11399	928
9	2916	18384	1328
10	4932	34698	2231
11	6371	45349	2868
12	8896	64496	3941
13	13671	105310	5993
Σύνολο	34890	279392	14916

Πίνακας 3.5. Αποτελέσματα που σχετίζονται με τη μελέτη προσομοίωσης του σφάλματος της διαδικασίας που σχετίζεται με τους εκτιμητές αποθέματος  $\hat{R}_i^{(1)}$ ,  $2 \leq i \leq 13$ . Η στήλη SD δίνει την τυπική απόκλιση, η στήλη Qt (95%) δίνει την 95% ποσότητα και η στήλη ES (95%) δίνει το αναμενόμενο έλλειμμα στο επίπεδο 95%. βλ. Mikosch (2006) σ. 94 για έναν επίσημο ορισμό.

### 3.5.2 Έρευνα σε δεδομένα : Ένα μοντέλο αξιοπιστίας που περιλαμβάνει διαγώνιες επιδράσεις

Σε αυτήν την ενότητα εφαρμόζουμε το μοντέλο i) -iii) σε δεδομένα του Πίνακα 3.1. Οι παραπάνω εκτιμητές των  $\beta_j, \delta_{i+j}$ ,  $0 \leq j \leq 12$ , χρησιμοποιούνται επίσης σε αυτήν την ενότητα. Η εκτίμηση  $\xi$  και  $\varphi$  λαμβάνονται με τη συνηθισμένη γραμμική παλινδρόμηση των υπολειμμάτων γραμμών,  $e_i$  έναντι  $\sum_{j=0}^{m-i} \beta_j \delta_{i+j}$ ,  $1 \leq i \leq 13$ , βλ. Σχήμα 3.3.



Εικόνα 3. Σφάλματα γραμμών,  $\hat{e}_i^2$ ,  $1 \leq i \leq 13$ , σχεδιασμένα  $\sum_{j=0}^{m-i} \beta_j \delta_{i+j}$ ,  $1 \leq i \leq 13$ . Η σταθερή γραμμή προσαρμόστηκε στα υπολείμματα γραμμών με τη μέθοδο των ελαχίστων τετραγώνων.

Αυτή η διαδικασία αποδίδει τις εκτιμήσεις

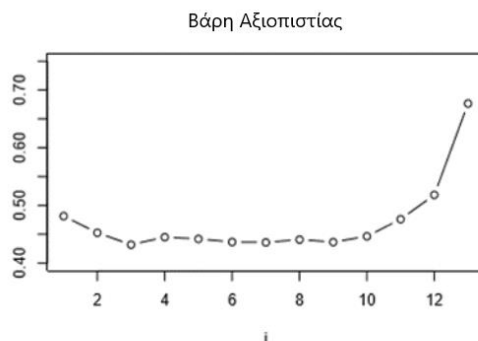
$$\hat{\xi} = 0.00383, \hat{\varphi} = 238.50.$$

Χρησιμοποιώντας τη συνθήκη ότι  $\hat{R}_i^{(2)} = \sum_{j=m-i+1}^{m-1} \hat{\mu}(X_{ij})$  και η πρόβλεψη των  $\delta_{14}, \dots, \delta_{25}$  που δίνονται στο Παράδειγμα 3.1 μπορούμε τώρα να υπολογίσουμε τους εκτιμητές αξιοπιστίας των αποθεμάτων γραμμών.

$i$	2	3	4	5	6	7	8	9	10	11	12	13
$\hat{R}_i$	12	342	685	709	1883	3578	8015	13564	26124	34576	46251	82476

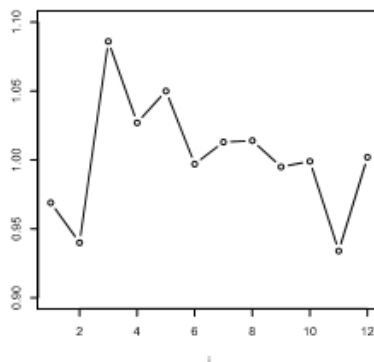
Πίνακας 3.6. Εκτιμητές αποθεμάτων  $\hat{R}_i^{(2)}$ ,  $2 \leq i \leq 13$ .

Το Σχήμα 3.4 δείχνει τα βάρη αξιοπιστίας  $\omega_i$  με τον καθορισμό της αξιοπιστίας που πρέπει να δοθεί στην εκ των προτέρων προσδοκία  $T_i$ .



Σχήμα 3.4. Βάρη αξιοπιστίας,  $\omega_i$ ,  $1 \leq i \leq 13$ , όπως δίδεται στην πρόταση 3.2.

Συγκρίναμε τελικά τις εκτιμήσεις του αποθέματος γραμμών  $\hat{R}_i^{(1)}$  και  $\hat{R}_i^{(2)}$  για να δούμε πόσο το μοντέλο αξιοπιστίας επηρεάζει τον τρόπο πρόβλεψης Bornhuetter-Ferguson. Ο λόγος των εκτιμήσεων αποθέματος απεικονίζεται στο Σχήμα 3.5.



Σχήμα 3.5. Δείκτες,  $\hat{R}_i^{(2)}/\hat{R}_i^{(1)}$ ,  $2 \leq i \leq 13$ , μεταξύ των εκτιμήσεων αποθέματος που ελήφθησαν για το μοντέλο που προτάθηκε στις παραγράφους 4 και 3.

### Μελέτη Προσομοίωσης

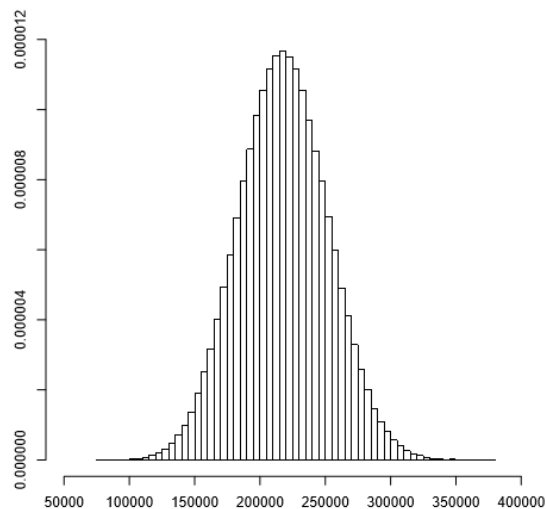
Ακολουθούμε την ίδια διαδικασία προσομοίωσης όπως περιγράφεται στην υποπαράγραφο 3.5.1, ωστόσο, με βάση τον ορισμό του υπολείμματος (3.4.1). Όπως στην υποπαράγραφο 3.5.1 έχουμε δημιουργήσει 500.000 ανεξάρτητες προσομοιώσεις. Τα

αποτελέσματα δεν είναι εκπληκτικά παρόμοια με την υποπαράγραφο 3.5.1 και δίνονται παρακάτω στον Πίνακα 3.7.

$i$	$SD$	$Qt(95\%)$	$ES(95\%)$
2	59	121	15
3	319	917	102
4	457	1461	203
5	471	1515	217
6	786	3225	360
7	1135	5506	528
8	1926	11289	904
9	2815	18340	1302
10	4750	34134	2154
11	6209	44936	2800
12	8221	60003	3652
13	13491	104857	5895
Σύνολο	34188	274513	14566

Πίνακας 7. Αποτελέσματα που σχετίζονται με τη μελέτη προσομοίωσης του σφάλματος της διαδικασίας που σχετίζεται με τους εκτιμητές αποθεματικών  $\hat{R}_i^{(2)}$ ,  $2 \leq i \leq 13$ . Η στήλη  $SD$  δίνει την τυπική απόκλιση, η στήλη  $Qt(95\%)$  δίνει την 95% ποσότητα και η στήλη  $ES(95\%)$  δίνει το αναμενόμενο έλλειμμα στο επίπεδο 95%.

Το Σχήμα 3.6 δείχνει το ιστόγραμμα των 500.000 προσομοιώσεων. Αυτό ολοκληρώνει επίσης την ανάλυση των δεδομένων μας.



Σχήμα 3.6. Ιστόγραμμα για το συνολικό απόθεμα.

## Κεφάλαιο Τέταρτο

### 4. Πρόβλεψη RBNS και IBNR απαιτήσεων χρησιμοποιώντας ποσά απαιτήσεων και μετρήσεις απαιτήσεων

#### 4.1 Περίληψη

Το τρίγωνο εξέλιξης αποζημιώσεων λαμβάνει υπόψη και θεωρείται ότι επίσης είναι παρατηρούμενοι οι αριθμοί των αναφερόμενων απαιτήσεων (σε μια παρόμοια τριγωνική συστοιχία). Σε αυτό το έγγραφο μόνο αυτές οι δύο τριγωνικές συστοιχίες χρησιμοποιούνται ως δεδομένα για το μοντέλο που συνεχίζουμε να ρυθμίζουμε. Οι περιορισμοί δεδομένων, σε κάποιο βαθμό, περιπλέκουν τη στατιστική ανάλυση, αλλά επιτρέπουν το μοντέλο που προτείνεται να εφαρμοστεί σε μια μεγάλη ποικιλία συνόλων δεδομένων. Με βάση αυτά τα δεδομένα προτείνουμε ένα στοχαστικό μοντέλο που βασίζεται σε ένα σύνθετο πλαίσιο Poisson. Το μοντέλο λαμβάνει ρητά υπόψη την καθυστέρηση από τη στιγμή που προκύπτει μια απαίτηση και μέχρι την αναφορά της (καθυστέρηση IBNR) και την καθυστέρηση από τη στιγμή που αναφέρεται μια απαίτηση και μέχρι την πλήρη εξόφληση (η καθυστέρηση RBNS). Υποτίθεται ότι οι μεμονωμένοι (μη παρατηρούμενοι) ισχυρισμοί είναι ανεξάρτητοι και ομοιόμορφα κατανομημένοι και ακολουθούν την Γάμμα κατανομή.

#### 4.2 Εισαγωγή

Προτείνουμε ένα διμετάβλητο στοχαστικό μοντέλο αποθεματικών ζημιών που βασίζεται σε παρατηρήσιμους αυξημένους αριθμούς απαιτήσεων  $N_{ij}$  και ποσά αποζημιώσεων  $X_{ij}$  και το οποίο χρησιμεύει στην πρόβλεψη RBNS και IBNR απαιτήσεων.

Ξεκινάμε με ένα μικρο-μοντέλο που υποδηλώνει ορισμένες σαφώς καθορισμένες ιδιότητες των αναφερόμενων συσσωρευμένων αριθμών απαιτήσεων και των συσσωρευμένων ποσών αποζημιώσεων. Στη συνέχεια, μελετάμε το θέμα τον εκτιμητή μέγιστης πιθανοφάνειας των παραμέτρων βάσει των τριγώνων  $\{N_{ij}\}$  και  $\{X_{ij}\}$ . Αυτό είναι σχετικά απλό, καθώς η πολλαπλασιαστική δομή της συνάρτησης πιθανοφάνειας επιτρέπει ξεχωριστή εκτίμηση των εισαγμένων παραμέτρων που ονομάζουμε  $\{\alpha_i, \beta_j\}$  και των  $p_0, \dots, p_d, \delta, \nu$ .

Το μοντέλο έχει τρία βασικά στοιχεία, πρώτον οι τυχαίες μεταβλητές των διαφόρων ετών ατυχημάτων θεωρούνται ανεξάρτητες για κάθε έτος ατυχήματος  $i$  και τους αριθμούς των αναφερόμενων απαιτήσεων

$$N_{i0}, \dots, N_{i,m-1}$$

θεωρούνται ότι έχουν πολλαπλασιαστική κατανομή Poisson με παραμέτρους  $e^{\alpha_i} e^{\beta_0}, \dots, e^{\alpha_i} e^{\beta_{m-1}}$ .



Δεύτερον, όταν εξαρτάται

$$N_{ij} := \sum_{k=0}^d N_{ijk} + N_{ij}^{zero},$$

οι αριθμοί  $N_{ijk}$  με καθυστέρηση  $k$  και ο αριθμός μηδενικών απαιτήσεων  $N_{ij}^{zero}$  έχουν μια πολυωνυμική κατανομή με την παράμετρο  $N_{ij}$  και τις παραμέτρους  $p_0, \dots, p_d$  και  $p^{zero}$  (ανεξάρτητα από το έτος ατυχήματος  $i$  και το έτος εξέλιξης  $j$ ) με

$$\sum_{k=0}^d p_k + p^{zero} = 1.$$

Τρίτον και τελευταίο, όταν εξαρτάται

$$N_{ij}^{paid} := \sum_{k=0}^{\min\{j,d\}} N_{i,j-k,k}$$

το ποσό απαίτησης αποζημιώσεων  $X_{ij}$  ακολουθεί κατανομή Γάμμα με παραμέτρους  $\delta N_{ij}^{paid}$  και  $\nu$  (επίσης ανεξάρτητο από το έτος ατυχήματος  $i$  και το έτος εξέλιξης  $j$ ).

Στο τελικό βήμα πρόβλεψης χρησιμοποιούμε εκτιμητές μέλη της οικογένειας  $\{N_{ijk}\}$  για την κατασκευή αριθμών RBNS και IBNR απαιτήσεων για κάθε κελί του κάτω τριγώνου. Υποτίθεται τότε ότι τα αντίστοιχα ποσά απαιτήσεων έχουν κατανομή Γάμμα όπως προηγουμένως, αλλά με  $N_{ij}^{rbns}$  και  $N_{ij}^{ibnr}$ , αντίστοιχα, στη θέση του  $N_{ij}^{paid}$ . Στη συνέχεια υπολογίζουμε τον εκτιμητή Bayes των  $X_{ij}^{ibnr}$  και  $X_{ij}^{rbns}$ , δεδομένου του τριγώνου των ανακοινωμένων αριθμών απαιτήσεων.

Σε αυτήν την παράγραφο εξετάζουμε μερικά από τα επιχειρήματά μας για το γιατί επιλέξαμε να εργαστούμε με τα παραπάνω συγκεντρωτικά δεδομένα και γιατί επιλέξαμε το προαναφερθέν μοντέλο. Υπάρχουν ορισμένα στοχαστικά μοντέλα που μπορούν να χρησιμοποιηθούν για τον υπολογισμό των αποθεμάτων στα μαθηματικά ασφάλισης ζημιών. Τα περισσότερα από αυτά τα μοντέλα έχουν σχεδιαστεί για να αντιμετωπίζουν δεδομένα που έχουν συγκεντρωθεί κατά κάποιο τρόπο, καθώς αυτό γίνεται σχετικά εύκολα από τον εξειδικευμένο αναλογιστή. Η συσσώρευση δεδομένων οδηγεί σε απώλεια πληροφοριών που σε ορισμένες περιπτώσεις μπορούν να παρέχουν σχετικά κακή εκτίμηση και πρόβλεψη των εκκρεμών υποχρεώσεων. Αυτό αποτέλεσε το αντικείμενο ορισμένων πρόσφατων εργασιών σχετικά με την αποθεματοποίηση : για παράδειγμα, χρησιμοποιούμε ένα γενικευμένο γραμμικό μοντέλο για να μοντελοποιήσουμε τα χαρακτηριστικά των μεμονωμένων απαιτήσεων. Πιο πίσω, ένα αξιοσημείωτο σύνολο εγγράφων καθορίζει ένα πλαίσιο για την εμφάνιση των απαιτήσεων, την αναφορά και τη διαδικασία αποζημίωσης, σε επίπεδο μεμονωμένων

απαιτήσεων. Αυτοί οι τύποι μοντέλων είναι πολύ λεπτομερείς, συχνά μάλλον περίπλοκοι και χρησιμοποιούν εκτεταμένα δεδομένα για την εκτίμηση των παραμέτρων. Ωστόσο, για τον εξειδικευμένο αναλογιστή, έχουν ορισμένους περιορισμούς: συγκεκριμένα, είναι δύσκολο να εφαρμοστούν επειδή η χρήση δεδομένων σε ατομικό επίπεδο είναι ιδιαίτερα ένας δύσκολος υπολογισμός. Επιπλέον, πολύ μεγάλα και περίπλοκα σύνολα δεδομένων είναι συχνά δύσκολο να βρεθούν σε ασφαλιστικές εταιρείες, και συχνά συμβαίνει ότι ένα μοντέλο θα χρησιμοποιηθεί μόνο σε μια πρακτική κατάσταση, εάν μπορεί να εφαρμοστεί σε μια μεγάλη ποικιλία συνόλων δεδομένων σε μια μεγάλη ποικιλία επιχειρηματικών τομέων.

Μένουμε λοιπόν με το δίλημμα του κατά πόσον θα χρησιμοποιήσουμε μεμονωμένα δεδομένα, τα οποία θεωρητικά είναι ελκυστικά, αλλά με δύσκολους υπολογισμούς, ή αν θα χρησιμοποιήσουμε συγκεντρωτικά δεδομένα, τα οποία είναι πολύ πιο εύκολο να αντιμετωπιστούν αλλά από τα οποία έχουν χαθεί ορισμένες (πιθανώς σημαντικές) πληροφορίες. Σε αυτό το κεφάλαιο, ακολουθούμε μια παρόμοια προσέγγιση, καθώς χτίζουμε το μοντέλο από βασικές αρχές στο επίπεδο των μεμονωμένων δεδομένων και βελτιώνουμε την ακρίβεια της αποθεματοποίησης προσθέτοντας συγκεντρωτικά δεδομένα μετρήσεων που είναι άμεσα διαθέσιμα στις περισσότερες αναλογιστικές υπηρεσίες. Ένας άλλος ενδιαφέρων τρόπος προσθήκης περισσότερων πληροφοριών στην προσέγγιση της κλασικής Chain Ladder, εμφανίζεται στην ανάγκη αυτή η τεχνική της Chain Ladder του Munich, όπου οι αθροιστικές αποζημιώσεις και τα συγκεντρωτικά πραγματοποιημένα αναμιγνύονται σε ένα ενωμένο μοντέλο. Άλλες ενδιαφέρουσες δυνατότητες προσθήκης επιπλέον δεδομένων σε σύγκριση με τη μέθοδο της απλής Chain Ladder είναι δοσμένες στην βιβλιογραφία. Αυτές προτείνουν διαφορετικούς τρόπους χειρισμού των RBNS και IBNR απαιτήσεων. Οι επιπλέον πληροφορίες των συγκεντρωτικών δεδομένων μετρήσεων μας επιτρέπουν να μοντελοποιούμε μοτίβα αποζημιώσεων για RBNS απαιτήσεις. Πιστεύουμε ότι αυτό παρέχει καλύτερη εκτίμηση και πρόβλεψη των εκκρεμών υποχρεώσεων. Ένα από τα ισχυρά επιχειρήματά μας είναι ότι το μοντέλο μας βασίζεται σε ένα υποκείμενο ρεαλιστικό μαθηματικό στατιστικό μοντέλο. Οι παραδοχές κατανομής ορίζονται καλά ως απλές απόρροιες αυτού του υποκείμενου μοντέλου.

Η μέθοδος Chain Ladder, η οποία σε πολλές περιπτώσεις είναι ενταγμένη στον χώρο των επιχειρήσεων, εισήχθη αρχικά χωρίς ένα στοχαστικό μοντέλο που προσδιορίστηκε χρησιμοποιώντας ευρετική συλλογιστική σκέψη για να εκτιμήσει το άθροισμα των απαιτήσεων που πραγματοποιήθηκαν αλλά δεν έχουν αναφερθεί (IBNR) και των απαιτήσεων RBNS. Σε αρκετά στοχαστικά μοντέλα από τότε έχουν διατυπωθεί οδηγούν στις ίδιες εκτιμήσεις με τη μέθοδο Chain Ladder. Και στις τρεις περιπτώσεις, αυτά τα μοντέλα λαμβάνουν τα δεδομένα όπως δίνονται και δεν επιχειρούν να δημιουργήσουν ένα μοντέλο βασισμένο στο κοινώς αποδεκτό σύνθετο πλαίσιο Poisson, που χρησιμοποιείται αλλού στη θεωρία κινδύνου. Θα μπορούσε να υποστηριχθεί ότι το υπερ-διασκορπισμένο μοντέλο Poisson θα μπορούσε να ερμηνευτεί με αυτόν τον τρόπο, αλλά αυτή δεν ήταν η αρχική προσέγγιση που είχε παρθεί. Σε αυτό το κεφάλαιο

παίρνουμε ένα σημείο εκκίνησης σε ένα σύνθετο μοντέλο Poisson. Αυτό είναι δυνατό επειδή παρατηρούμε και μοντελοποιούμε τον αριθμό των ανακοινωμένων απαιτήσεων.

Μια αυστηρή διαμόρφωση μοντέλου για IBNR και RBNS απαιτήσεις διατυπώνεται και επιτρέπει την πρόβλεψη IBNR και RBNS απαιτήσεων ξεχωριστά. Αυτό είναι δυνατό επειδή έχουμε προσθέσει συσσωρευμένους αριθμούς στο σύνολο δεδομένων. Στην Παράγραφο 3 καθορίζουμε τη σημειογραφία και περιγράφουμε τα δεδομένα που θα υποθέσουμε ότι είναι διαθέσιμα. Στην Παράγραφο 4, δίνεται η θεωρητική εξέλιξη, βασισμένη σε υποθέσεις στο επίπεδο των μεμονωμένων δεδομένων. Επίσης στην Παράγραφο 4, δίνεται το μοντέλο που πραγματικά θα εφαρμόσουμε, ως προσέγγιση των πιο λεπτομερών μοντέλων για μεμονωμένα δεδομένα. Η Παράγραφος 5 εξετάζει την εκτίμηση, η Παράγραφος 6 την πρόβλεψη, η Παράγραφος 7 εξετάζει μερικά αποτελέσματα με βάση το μοντέλο και στην Παράγραφο 8 συλλέγουμε τα πιο σημαντικά συμπεράσματα.

### 4.3 Δεδομένα και Σημειογραφία

Τα δεδομένα με τα οποία εργαζόμαστε επιλέγονται προσεκτικά. Τα δύο κύρια σημεία της επιλογής μας δεδομένων είναι ότι τα επιπλέον συγκεντρωτικά δεδομένα μετρήσεων είναι άμεσα διαθέσιμα για τους περισσότερους πρακτικούς αναλογιστές και ότι αυτά τα επιπλέον δεδομένα παρέχουν πολύ καλύτερο χειρισμό των RBNS και IBNR αποθεμάτων. Δίνει επίσης μια καλά καθορισμένη εκτιμώμενη ταμειακή ροή που προσφέρεται για πολλές αναλογιστικές αιτήσεις, συμπεριλαμβανομένων των κεφαλαιακών εκτιμήσεων.

#### 4.3.1 Δεδομένα και Σημειογραφία

Ξεκινάμε με τον καθορισμό του συμβολισμού. Ας πούμε τώρα ότι ένα τρίγωνο εξέλιξης αποτελείται από τις τυχαίες μεταβλητές  $\Delta_m = \{X_{ij}; (i, j) \in Am\}$  όπου  $Am = \{(i, j) \in \mathbb{N} \times \mathbb{N}_0; 1 \leq i + j \leq m\}$ .

$$\begin{array}{cccccc} X_{10} & X_{11} & \cdots & X_{1,m-2} & X_{1,m-1} & \\ X_{20} & X_{21} & \cdots & X_{2,m-2} & & \\ \vdots & \ddots & & & & \\ X_{m0} & & & & & \end{array}$$

Η πρώτη κατάληξη είναι ένα σήμα που υποδηλώνει την περίοδο προέλευσης της απαίτησης, η περίοδος κατά την οποία πραγματοποιήθηκε. Η δεύτερη κατάληξη υποδηλώνει την καθυστέρηση από την πραγματοποιηθείσα περίοδο έως την εξόφληση της αποζημίωσης (ή αποθέματος απαιτήσεων).

Υπάρχουν διαφορετικές δυνατότητες για το τι θα μπορούσαν να αντιπροσωπεύουν οι τυχαίες μεταβλητές στα  $\Delta_m$ . Τα  $X_{ij}$ ,  $(i, j) \in Am$ , θα μπορούσαν να είναι οι συνολικές απαιτήσεις που έγιναν στην περίοδο  $i$  και αποζημιώθηκαν με καθυστερήσεις περιόδων  $j$  από όταν πραγματοποιήθηκαν. Αυτό είναι τότε το τρίγωνο των αποζημιωμένων

απαιτήσεων. Μια άλλη δυνατότητα προσδιορίζει το  $X_{ij}$  ως συνολικές απαιτήσεις που έγιναν στην περίοδο  $i$  με καθυστέρηση  $j$ . Στις περιπτώσεις όπου οι απαιτήσεις αναφέρονται αλλά δεν έχουν εξοφληθεί, περιλαμβάνεται μια εκτίμηση απαίτησης, αντί της πραγματικής αποζημίωσης. Σε αυτήν την περίπτωση, το τρίγωνο αντιπροσωπεύει τις πραγματοποιηθείσες απαιτήσεις. Η χρήση είτε των εξοφλημένων είτε των πραγματοποιούμενων απαιτήσεων έχει διαφορετικά πλεονεκτήματα και μειονεκτήματα. Όταν περιλαμβάνει μόνο δεδομένα αποζημιώσεων, τότε το  $\Delta_m$  δεν περιέχει ανθρώπινη κρίση και συνεπώς ασχολούμαστε με "πραγματικά δεδομένα". Ωστόσο, υπάρχουν πληροφορίες σχετικά με μελλοντικές αποζημιώσεις, RBNS απαιτήσεις, οι οποίες στη συνέχεια αγνοούνται. Από την άλλη πλευρά, η συμπερίληψη των εκτιμήσεων απαιτήσεων είναι συζητήσιμη, καθώς κανείς δεν θεωρεί ότι είναι πλέον "πραγματικά δεδομένα". Μπορεί να υπάρχουν πολιτικά ή επιχειρηματικά ζητήματα που καθιστούν τις εκτιμήσεις μεμονωμένων απαιτήσεων αναξιόπιστες. Υπάρχει επιπλέον κάποια μεταβλητότητα που δεν λαμβάνεται υπόψη, καθώς οι εκτιμητές των απαιτήσεων και το πραγματικό καταβληθέν ποσό συχνά διαφέρουν. Τέλος, οι εκτιμήσεις απαιτήσεων εμφανίζονται ως εξοφλημένες σε λάθος χρονική στιγμή που διαταράσσει τη μοντελοποίηση των ταμειακών ροών.

Για τους λόγους αυτούς, υιοθετούμε την πρώτη προσέγγιση σε αυτό το κεφάλαιο και χρησιμοποιούμε το τρίγωνο των απαιτήσεων αποζημιώσεων. Έτσι, το  $X_{ij}$  είναι το σύνολο των απαιτήσεων που πραγματοποιήθηκαν στην περίοδο  $i$  και εξοφλήθηκαν με καθυστέρηση περιόδων  $j$ . Τα καταβληθέντα ποσά είναι πραγματικά δεδομένα και είναι εύκολα προσβάσιμα στις περισσότερες εταιρείες, καθώς είναι οι αριθμοί που λαμβάνονται απευθείας από τα λογιστικά έγγραφα. Συνδυάζουμε αυτό το τρίγωνο αποζημιώσεων με τις πραγματοποιημένες πληροφορίες. Χρησιμοποιούμε επίσης ένα δεύτερο τρίγωνο, με την ίδια ακριβώς μορφή με το παραπάνω τρίγωνο αποζημιώσεων. Προκειμένου να ενσωματώσουμε τις διαθέσιμες πληροφορίες μας σχετικά με τον αριθμό των πραγματοποιημένων RBNS απαιτήσεων, θεωρούμε τυχαίες μεταβλητές  $\mathfrak{N}_m = \{N_{ij} : (i, j) \in Am\}$  όπου  $N_{ij}$  αντιπροσωπεύει τον αριθμό των απαιτήσεων που πραγματοποιήθηκαν στην περίοδο  $i$  και αναφέρθηκαν με καθυστέρηση περιόδων  $j$  (στην περίοδο  $i + j$ ) για  $(i, j) \in Am$ . Λάβετε υπόψη ότι αυτά τα δεδομένα αποτελούνται από τις πραγματοποιημένες απαιτήσεις και, ως εκ τούτου, χρησιμοποιούν ορισμένες από τις πληροφορίες που δεν χρησιμοποιούνται όταν χρησιμοποιούνται μόνο οι συνολικές απαιτήσεις αποζημιώσεων. Θα ήταν επίσης δυνατό να εξεταστεί ο αριθμός των αποζημιώσεων, οι οποίες θα αφαιρούσαν τις απαιτήσεις που καταλήγουν χωρίς να πραγματοποιηθεί αποζημίωση - οι μηδενικές απαιτήσεις - από τα δεδομένα. Ωστόσο, αυτό μπορεί να οδηγήσει σε ορισμένες δυσκολίες. Για παράδειγμα, ο αριθμός των αποζημιώσεων σπάνια είναι εύκολα προσβάσιμος στις ασφαλιστικές εταιρείες. Ο αριθμός των αναφερόμενων απαιτήσεων από την άλλη πλευρά είναι συνήθως σχετικά εύκολο να ληφθεί.

Επομένως, θα υποθέσουμε ότι μπορούμε να παρατηρήσουμε  $\{(N_{ij}, X_{ij}): (i, j) \in Am\}$ . Στη συνέχεια ακολουθεί η σημειογραφία  $(\mathfrak{K}_m, \Delta_m) := \{(N_{ij}, X_{ij}): (i, j) \in Am\}$  που χρησιμοποιείται. Αυτά τα δεδομένα είναι εύκολο να ληφθούν στις περισσότερες περιπτώσεις. Σημειώνουμε εδώ ότι γνωρίζουμε ότι η χρήση δεδομένων στα οποία έχουμε περιοριστεί σε κάποιο βαθμό περιπλέκει τη στατιστική ανάλυση. Θα ήταν καλύτερα, από αυτήν την άποψη τουλάχιστον, να υποθέσουμε ότι διαθέτουμε δεδομένα σε οποιοδήποτε επίπεδο λεπτομέρειας χρειαζόμαστε. Ένα μειονέκτημα αυτού θα ήταν ότι η εκτίμηση των μοντέλων έγινε πολύ πιο υπολογιστικά δύσκολη και τα μοντέλα δεν μπορούσαν να χρησιμοποιηθούν όταν δεν ικανοποιήθηκαν οι απαιτήσεις των δεδομένων.

Έτσι, έχουμε κάνει έναν συμβιβασμό σχετικά με τα δεδομένα που χρησιμοποιούμε, αλλά θα φανεί στην Παράγραφο 6 ότι με την απλή συμπερίληψη των δεδομένων μετρήσεων για τις πραγματοποιημένες απαιτήσεις, μπορούμε να βελτιώσουμε σημαντικά την Chain Ladder μέθοδο χωρίς να εγκαταλείψουμε πλήρως την γνωστή Chain Ladder μέθοδο. Η μέθοδος Chain Ladder είναι στην πραγματικότητα μια ειδική περίπτωση του μοντέλου μας.

Στην επόμενη παράγραφο συνδέουμε  $\Delta_m$  και  $\mathfrak{K}_m$  χρησιμοποιώντας μερικές μη παρατηρήσιμες τυχαίες μεταβλητές. Η δομή των μη παρατηρήσιμων μεταβλητών προορίζεται να μιμείται τα μοντέλα του Norberg αλλά χρησιμοποιούμε ένα διακριτό χρονικό πλαίσιο.

### 4.3.2 IBNR και RBNS Απαιτήσεις

Σε αυτήν την υποπαράγραφο παρουσιάζουμε το μοντέλο μας που περιλαμβάνει έναν αριθμό (στην πράξη συχνά) μη παρατηρήσιμων τυχαίων μεταβλητών. Με βάση αυτό το μοντέλο είμαστε σε θέση να διατυπώσουμε μια ερμηνεία σύνθετης Poisson των IBNR και RBNS απαιτήσεων (σε ατομικό επίπεδο). Στην Υποπαράγραφο 4.3.1, περιεγράφηκαν τα διαθέσιμα στοιχεία για τη στατιστική ανάλυση, σε αυτήν την υποπαράγραφο σημειώθηκε σχετική με ορισμένες βασικές μεταβλητές που μιμούνται τη δομή των υποκείμενων δεδομένων που σχετίζονται με την αποζημίωση των απαιτήσεων.

Το λεπτομερές μοντέλο αυτού του κεφαλαίου με παραδοχές για την ανάπτυξη μεμονωμένων απαιτήσεων ακολουθεί σε κάποιο βαθμό τις θεωρητικές έρευνες των Mack (1991) και Norberg (1993,1999) Αυτή η προσέγγιση λειτουργεί μέσω περιορισμών στα διαθέσιμα δεδομένα. Ενώ οι υποθέσεις γίνονται σε επίπεδο μεμονωμένων δεδομένων, μόνο τα συγκεντρωτικά δεδομένα θεωρείται ότι είναι διαθέσιμα για τη διαδικασία στατιστικής εκτίμησης.

Εξετάζουμε την  $k$ -οστή αιτία των  $N_{ij}$  απαιτήσεων που πραγματοποιήθηκαν στην περίοδο  $i$  και αναφέρθηκαν με καθυστέρηση περιόδων  $j$ . Συνήθως μια αιτία δεν καταβάλλεται αμέσως μετά από την ανακοίνωση στην ασφαλιστική εταιρεία. Το τελικό ποσό της αιτίας καταβάλλεται γενικά με κάποιο χρόνο αναμονής από την ανακοίνωση, συχνά λόγω γενικής εξέτασης της υπόθεσης, νομικών ζητημάτων,

συλλογής περαιτέρω πληροφοριών σχετικά με την υπόθεση κ.λπ. Με άλλα λόγια, υπάρχει καθυστέρηση από την υποβολή απαίτησης και έως ότου εξοφληθεί πλήρως. Οι απαιτήσεις που έχουν αναφερθεί αλλά δεν έχουν ακόμη αποζημιωθεί είναι οι λεγόμενες RBNS απαιτήσεις (ή απαιτήσεις που έχουν πραγματοποιηθεί αλλά δεν έχουν αναφερθεί αρκετά (IBNER)). Η σχετική καθυστέρηση αποζημίωσης αναφέρεται ως καθυστέρηση RBNS.

Προκειμένου να διατυπωθεί μαθηματικά πώς αποζημιώνονται οι απαιτήσεις, εισάγουμε μερικές στοχαστικές μεταβλητές που σε ορισμένες περιπτώσεις μπορεί να είναι παρατηρήσιμες, αλλά υποθέτουμε γενικά ότι δεν είναι, και το μοντέλο δεν βασίζεται στο να έχει παρατηρήσεις για αυτές. Δείχνουμε από τα  $N_{ijk}$  το μέρος των απαιτήσεων  $N_{ij}$  που αποζημιώνονται (πλήρως) με καθυστέρηση περιόδων  $k$ ,  $k = 0, \dots, d$ . Εδώ το  $k = 0$  αντιστοιχεί σε μια απαίτηση που αποζημιώθηκε την ίδια περίοδο με την οποία αναφέρθηκε, ενώ το  $k = d$  είναι η μέγιστη δυνατή RBNS καθυστέρηση στο μοντέλο,  $d$  θα μπορούσε να επιλεγεί χρησιμοποιώντας πληροφορίες από τα υποκείμενα δεδομένα ή την κρίση από ένα χειριστή απαιτήσεων.

Μερικές φορές συμβαίνει ότι οι αναφερόμενες απαιτήσεις διευθετούνται χωρίς αποζημίωση, για παράδειγμα, εάν υπάρχει προβληματισμός σχετικά με το ποιος φέρει την ευθύνη για μια απαίτηση, απάτη ή παρόμοιο. Αυτά αναφέρονται ως μηδενικές απαιτήσεις. Δηλώστε τον αριθμό των μηδενικών απαιτήσεων που έχουν προκύψει (ή προφανώς έχουν προκύψει σε περιπτώσεις απάτης κ.λπ.) στην περίοδο  $(i, j)$  και διακανονίζονται στην τιμή μηδέν κάποια στιγμή από το  $N_{ij}^{zero}$ . Στη συνέχεια πρέπει να έχουμε  $N_{ij0} + \dots + N_{ijd} \leq N_{ij}$  και  $N_{ij}^{zero} + N_{ij0} + \dots + N_{ijd} = N_{ij}$  για  $(i, j) \in Am$ . Ο αριθμός των απαιτήσεων που προέκυψαν κατά την περίοδο  $i$  και εξοφλήθηκαν (πλήρως) με καθυστερήσεις περιόδων  $j$  μετά την ανακοίνωση, προσδιορίζεται ως

$$N_{ij0} + N_{i,j-1,1} + \dots + N_{i,j-\min\{j,d\},d} = \sum_{k=0}^{\min\{j,d\}} N_{i,j-k,k} =: N_{ij}^{paid}$$

για  $(i, j) \in Am$  όπου έχουμε  $N_{ijk} = 0$  για  $j < 0, i, k \geq 1$ , από συνθήκη. Η μη αρνητική τυχαία μεταβλητή  $Y_{ijk}$  είναι η  $k$ -οστή απαίτηση που αποζημιώθηκε στην περίοδο  $i + j$  η οποία πραγματοποιήθηκε στην περίοδο  $i$  για  $(i, j) \in Am$  και  $k = 1, \dots, N_{ij}^{paid}$ .

Έτσι, διαιρούμε τη διάρκεια ζωής μιας απαίτησης σε δύο: την IBNR καθυστέρηση και την RBNS καθυστέρηση (και IBNER).

Ο συμβολισμός έχει πλέον καθοριστεί και η επόμενη παράγραφος αφορά τη διαμόρφωση ενός μοντέλου και στη συζήτηση των δυνατοτήτων και των περιορισμών του.

#### 4.4 Το Στοχαστικό Μοντέλο

Σε αυτήν την ενότητα, βασικές παραδοχές σχετικά με τον τρόπο με τον οποίο οι απαιτήσεις αναφέρονται και αποζημιώνονται θεωρούνται ότι οδηγούν σε ένα μοντέλο για τα καταβληθέντα ποσά και τον αναφερόμενο αριθμό απαιτήσεων. Το μοντέλο για τον αναφερόμενο αριθμό απαιτήσεων παρατίθεται στην Υποπαράγραφο 4.4.1 και τα μοντέλα για ποσά απαιτήσεων προέρχονται στην Υποπαράγραφο 4.4.2. Αυτά τα μοντέλα είναι κατασκευασμένα από αρχές που σχετίζονται στενά με τη μορφή στα μοντέλα του Norberg (1993, 1999).

#### 4.4.1 Μετρήσεις Απαιτήσεων

Σε αυτήν την υποπαράγραφο, προτείνεται ένα μοντέλο για τον αριθμό των αναφερόμενων απαιτήσεων,  $\mathfrak{N}_m$ . Υποτίθεται ότι τα  $N_{ij}$  είναι ανεξάρτητες τυχαίες μεταβλητές, οι οποίες ακολουθούν κατανομή Poisson με μέση  $\mu_{ij}$  όπου

$$\log \mu_{ij} = a_i + \beta_j \quad (4.4.1)$$

και  $\sum_{j=0}^{m-1} e^{\beta_j} = 1$ . Οι εκτιμητές μέγιστης πιθανοφάνειας των  $a_i$ ,  $1 \leq i \leq m$ , και  $\beta^j$ ,  $0 \leq j \leq m-1$ , είναι τέτοιοι που η εκτίμηση του αριθμού των IBNR απαιτήσεων αντιστοιχεί στη μέθοδο Chain Ladder που εφαρμόζεται στο  $\mathfrak{N}_m$ . Δεδομένου ότι η μέθοδος της Chain Ladder είναι στάνταρ μεταξύ των επαγγελματιών αναλογιστών και επειδή η μοντελοποίηση του  $\mathfrak{N}_m$  δεν είναι ο πρωταρχικός στόχος αυτής της υποπαραγράφου, θα υποθέσουμε απλώς το παραπάνω μοντέλο για το  $\mathfrak{N}_m$ .

#### 4.4.2 Ποσά Αποζημιώσεων : Μία Αποζημίωση ανά Απαίτηση

Το μοντέλο αυτής της υποπαραγράφου προϋποθέτει μόνο μία αποζημίωση ανά απαίτηση. Αυτός ο περιορισμός απλοποιεί τη θεωρία, την εκτίμηση και τις ερωτήσεις δεδομένων σημαντικά. Για παράδειγμα, στην πραγματική ζωή υπάρχουν συχνά περισσότερες από μία αποζημιώσεις ανά απαίτηση, ωστόσο, η μοντελοποίηση αυτού μέσω στοχαστικών προσεγγίσεων χρονοσειρών είναι ασήμαντη. Επίσης, συχνά δεν είναι διαθέσιμα δεδομένα για την εξέλιξη των τρόπων αποζημίωσης και οι ορισμοί για τις αποζημιώσεις ενδέχεται να διαφέρουν από τη μία ασφαλιστική εταιρεία στην άλλη, ή ακόμη και εντός της ίδιας ασφαλιστικής εταιρείας. Υποθέτοντας ότι έχουμε μια αποζημίωση για κάθε απαίτηση, πρέπει να θεωρήσουμε ένα καλά επιλεγμένο μοντέλο απλοποίησης που οδηγεί σε περισσότερη δυναμική, λιγότερες παραμέτρους για εκτίμηση και ευρύτερη εφαρμογή της μεθόδου.

Η αναλογία των απαιτήσεων διευθετήθηκε στο μηδέν,  $Q \in [0,1)$ , υπολογίζεται ως γνωστή σταθερά για  $(i, j) \in Am$ . Δεδομένου  $N_{ij}$  έχουμε

$$(N_{ij0}, \dots, N_{ija}, N_{ij}^{zero}) \sim Multi(N_{ij}; p_0, \dots, p_a, Q)$$

για  $(i, j) \in Am$  και  $p_0 + \dots + p_d = 1 - Q$  όπου  $p_i \in (0, 1), 0 \leq i \leq d$ . Με άλλα λόγια, υποτίθεται ότι η υπό όρους πυκνότητα (το μέτρο μέτρησης σε  $\mathbb{N}_0^{d+2}$ ) του  $(N_{ij0}, \dots, N_{ijd}, N_{ij}^{zero})$  δίνεται από

$$f_{N_{ij0}, \dots, N_{ijd}, N_{ij}^{zero} | N_{ij}}(n_0, \dots, n_{d+1}) = \binom{N_{ij}}{n_0, \dots, n_{d+1}} p_0^{n_0}, \dots, p_d^{n_d} Q^{n_{d+1}}$$

για  $n_l \in \mathbb{N}_0, 0 \leq l \leq d + 1$  έτσι ώστε  $n_0 + \dots + n_{d+1} = N_{ij}$ , βλέπε Υποπαράγραφο 4.3.2. Οι συνολικές αποζημιώσεις που γίνονται με καθυστέρηση περιόδων  $l$  ορίζονται ως

$$S_{ijl} = \sum_{k=N_{ij0}+\dots+N_{ij,l-1}+1}^{N_{ij0}+\dots+N_{ijl}} Y_{ij}^{(k)} \quad (4.4.2)$$

για  $(i, j) \in Am$ . Όπου  $Y_{ij}^{(k)}, (i, j) \in Am, k \geq 1$ , όπως ορίζεται στην Υποπαράγραφο 4.4.2. Εδώ η δομή δεύτερης τάξης είναι

$$E(S_{ijl} | N_{ij} = n_{ij}) = n_{ij} EY_{11}^{(1)} p_l,$$

$$Cov(S_{ijl}; S_{ijl'} | N_{ij} = n_{ij}) = n_{ij} \left[ V(Y_{11}^{(1)}) p_l I\{l = l'\} - (EY_{11}^{(1)})^2 p_l p_{l'} \right]$$

για  $(i, j) \in Am$  και  $l, l' = 0, \dots, d$ . Στην πραγματικότητα, υπάρχουν διάφοροι αριθμοί αποζημιώσεων που σχετίζονται με διαφορετικές απαιτήσεις: μερικές φορές υπάρχει μία αποζημίωση και μερικές φορές υπάρχουν περισσότερες. Από ρεαλιστική άποψη, αυτό σημαίνει ότι σε ορισμένες περιπτώσεις οι εκτιμήσεις παραμέτρων ενδέχεται να μην έχουν άμεσες ερμηνείες. Ενώ υπάρχουν φυσικά απαιτήσεις που αποζημιώνονται σε μία αποζημίωση, αναγνωρίζουμε ότι για πολλές απαιτήσεις αυτό δεν ισχύει. Ωστόσο, όταν υπάρχουν περισσότερες από μία αποζημιώσεις, τότε συχνά υπάρχει μια κύρια αποζημίωση που είναι μεγάλη σε σύγκριση με τις άλλες (που είναι προσαρμογές, πρόσθετο κόστος ή παρόμοια). Έτσι, το μοντέλο (4.4.2) έχει τον τύπο ιδιοτήτων στο συνολικό επίπεδο που πιστεύουμε ότι είναι λογικό και μπορούμε να δεχτούμε κάποιες θεωρητικές αδυναμίες στο μοντέλο. Όπως δηλώθηκε στην αρχή του κεφαλαίου, η λογική διαδικασία που ακολουθήσαμε είναι να αντλήσουμε ένα μοντέλο όσο το δυνατόν περισσότερο, βασισμένο σε πολύ βασικές, μη παρατηρήσιμες τυχαίες μεταβλητές και, στη συνέχεια, να προσεγγίσουμε το μοντέλο όσο το δυνατόν πιο κοντά για να οργανώσουμε ένα μοντέλο για τα διαθέσιμα δεδομένα. Η αποδοχή της σύμβασης ότι υπάρχει μόνο μία αποζημίωση ανά απαίτηση, οδηγεί σε σχετικά απλά χαρακτηριστικά κατανομής. Η υπό όρους κατανομή του  $S_{ijl}$  δεδομένου του  $N_{ijl}$  είναι μια κατανομή  $\Gamma(\delta N_{ijl}, \nu)$  και έτσι έχουμε

$$X_{ij} := \sum_{l=0}^d \sum_{k=N_{ij0}+\dots+N_{ij,l-1}+1}^{N_{ij0}+\dots+N_{ijl}} Y_{ij}^{(k)} \triangleq \sum_{k=1}^{N_{ij0}+\dots+N_{ij,d}} Y_{11}^{(k)} \sim \Gamma(N_{ij}^{paid} \delta, \nu)$$



δεδομένου  $N_{ij}^{paid}$  χρησιμοποιώντας τη σημειογραφία από την Υποπαράγραφο 3.2.

Σε αυτήν την υποπαράγραφο, συζητήσαμε ένα μοντέλο για  $\Delta_m$  δεδομένου  $\mathfrak{x}_m$  και υποστηρίξαμε ότι αυτό το μοντέλο (4.4.1) και (4.4.2) είναι κατάλληλο για τη διαθέσιμη μορφή δεδομένων. Οι επόμενες παράγραφοι συζητούν την εκτίμηση και την πρόβλεψη για αυτό το μοντέλο.

## 4.5 Εκτίμηση

Σε αυτήν την παράγραφο υπολογίζονται οι συναρτήσεις πιθανοφάνειας για το μοντέλο (4.4.1) και (4.4.2) που προτείνονται στις υποπαραγράφους 4.4.1 και 4.4.2, δεδομένου ότι τα δεδομένα  $(\mathfrak{x}_m, \Delta_m)$  υπολογίζονται. Η συνάρτηση πιθανότητας για  $(\mathfrak{x}_m, \Delta_m)$  μπορεί να γραφτεί ως

$$\begin{aligned} & \mathcal{L}_{\mathfrak{x}_m, \Delta_m}(\{\alpha_i, 1 \leq i \leq m\}, \{\beta_j, 0 \leq j \leq m-1\}, \delta, \nu, \{p_l, 0 \leq l \leq d\}) \quad (4.5.1) \\ & \mathcal{L}_{\mathfrak{x}_m}(\{\alpha_i, 1 \leq i \leq m\}, \{\beta_j, 0 \leq j \leq m-1\}) \times \mathcal{L}_{\Delta_m | \mathfrak{x}_m}(\delta, \nu, p_0, \dots, p_d) = \\ & = \left( \prod_{i=1}^m \prod_{j=0}^{m-i} P(N_{ij} = n_{ij}) \right) \times \left( \prod_{i=1}^m f_{X_{i0}, \dots, X_{i,m-i} | N_{i0}, \dots, N_{i,m-i}}(X_{i0}, \dots, X_{i,m-i} | n_{i0}, \dots, n_{i,m-i}) \right). \end{aligned}$$

Δεδομένου ότι τα  $\mathcal{L}_{\mathfrak{x}_m}$  και  $\mathcal{L}_{\Delta_m | \mathfrak{x}_m}$  δεν είναι συναρτήσεις των ίδιων παραμέτρων, αρκεί να μεγιστοποιήσουμε τα  $\mathcal{L}_{\mathfrak{x}_m}$  και  $\mathcal{L}_{\Delta_m | \mathfrak{x}_m}$  ξεχωριστά για να μεγιστοποιηθεί το  $\mathcal{L}_{\mathfrak{x}_m, \Delta_m}$  στην (4.5.1).

Η συνάρτηση πιθανοφάνειας του  $\mathfrak{x}_m$  μπορεί να βελτιστοποιηθεί με τον ακόλουθο τρόπο χρησιμοποιώντας τη μέθοδο Chain Ladder. Προσδιορίζουμε το συσσωρευτικό τρίγωνο εξέλιξης ως

$$C_{ij} = \sum_{k=0}^j N_{ik}, \quad \text{για } (i, j) \in AM,$$

και έχουμε τα  $\hat{C}_{i,m-1}$  να είναι οι προγνωστικοί δείκτες της κλίμακας αλυσίδας του  $C_{i,m-1}$ . Στη συνέχεια δίνονται οι εκτιμήσεις μέγιστης πιθανοφάνειας  $e^{\alpha_i}$  και  $e^{\beta_j}$  από

$$\begin{aligned} e^{\alpha_i} &= \hat{C}_{i,m-1}, & 1 \leq i \leq m \\ e^{\beta_j} &= \frac{\hat{C}_{mj} - \hat{C}_{i,m-1}}{e^{\alpha_i}}, & 0 \leq j \leq m-1. \end{aligned}$$

Παράρτημα Α για την απόδειξη.

Η συνάρτηση λογαριθμικής - πιθανοφάνειας για  $\Delta_m$  δεδομένου  $\mathfrak{x}_m$ , βλ. Παράρτημα Α, δίνεται από την (4.5.2)

$$\log(\mathcal{L}_{\Delta_m | \mathfrak{x}_m}(\delta, \nu, p_0, \dots, p_d)) = C' - \nu x_{..} +$$

$$+ \sum_{i=1}^m \log \left( \sum_{j=0}^{m-i} Q^{n_{ij}} n_{ij}! \left[ \sum_{n_{ij0}=0}^{n_{ij}} \left(\frac{p_0}{Q}\right)^{n_{ij0}} \frac{1}{n_{ij0}!} \sum_{n_{ij1}=0}^{n_{ij}-n_{ij0}} \left(\frac{p_1}{Q}\right)^{n_{ij1}} \frac{1}{n_{ij1}!} \cdots \sum_{n_{ijd}=0}^{n_{ij}+n_{ijd}-n_{ij}} \left(\frac{p_d}{Q}\right)^{n_{ijd}} \frac{1}{n_{ijd}!} \frac{1}{(n_{ij}-n_{ij\cdot})!} \prod_{k=0}^{m-i} \frac{(v x_{ik})^{n_{ik}^{paid} \delta}}{\Gamma(n_{ik}^{paid} \delta)} \right] \right)$$

όπου  $n_{ik}^{paid} = n_{ij0} + \cdots + n_{i,-d,d}$ ,  $d$  με  $n_{ijl} := 0$  για  $l \leq -1$ ,  $(i, k) \in Am$  και

$$x_{..} = \sum_{i=1}^m \sum_{j=0}^{m-i} x_{ij}.$$

Εάν μερικά από τα  $n_{ij}$  είναι μεγάλα νούμερα, οι αριθμητικοί υπολογισμοί των αθροισμάτων στο (4.5.2) απαιτούν μεγάλο χρόνο. Για να ξεπεραστεί αυτό το πρόβλημα, ανατρέχουμε στο Παράρτημα Β για μια εναλλακτική μέθοδο εκτίμησης χρησιμοποιώντας γενικευμένα γραμμικά μοντέλα.

Τώρα έχουν προκύψει τύποι για την εκτίμηση των παραμέτρων του μοντέλου που καθορίζονται στους τύπους (4.4.1) και (4.4.2). Η επόμενη παράγραφος είναι αφοσιωμένη σε μια ανάλυση για το τι σημαίνει οι RBNS και IBNR απαιτήσεις σε συσσωρευτικό επίπεδο και πώς μπορούν να εκτιμηθούν.

## 4.6 Πρόβλεψη

Αυτή η παράγραφος περιγράφει τι είναι οι RBNS και IBNR απαιτήσεις (βάσει της υποπαραγράφου 4.3.2) και πώς μπορούν να προβλεφθούν. Πρώτον, αποσαφηνίζουμε τις έννοιες των RBNS και IBNR απαιτήσεις σε συσσωρευτικό επίπεδο στο πλαίσιο ενός κατάλληλου μοντέλου. Το μοντέλο για τις τριγωνικές συστοιχίες  $(\Delta_m, \mathfrak{X}_m) = \{(N_{ij}, X_{ij}) : (i, j) \in Am\}$  που ορίζεται στα (4.4.1) και (4.4.2) μπορεί να επεκταθεί με φυσικό τρόπο στις τυχαίες μεταβλητές  $N_{ij}, (i, j) \in Bm$ . Εδώ

$$Bm = \{(i, j) \in \mathbb{N}_0^2 : 1 \leq i \leq m, 0 \leq j \leq m-1\}$$

και  $X_{ij}, (i, j) \in Cm$ , όπου

$$Cm = \{(i, j) \in \mathbb{N}_0^2 : 1 \leq i \leq m, 0 \leq j \leq m+d-1\}.$$

Οι τυχαίες μεταβλητές εμφανίζονται έτσι σε αυτήν τη μορφή

$$\begin{array}{cccccc} N_{10} & \cdots & N_{1,m-1} & & X_{10} & \cdots & X_{1,m+d-1} \\ N_{20} & \cdots & N_{2,m-1} & & X_{20} & \cdots & X_{2,m+d-1} \\ \vdots & & \vdots & & \vdots & & \vdots \\ N_{m0} & \cdots & N_{m,m-1} & & X_{m0} & \cdots & X_{m,m+d-1} \end{array}.$$

Σύμφωνα με το (4.4.1), είναι φυσικό να υποθέσουμε ότι οι μεταβλητές  $N_{ij}$ ,  $m-i+1 \leq j \leq m-1$ , είναι ανεξάρτητες και

$$N_{ij} \sim \text{Poisson}(\mu_{ij}), \quad \log(\mu_{ij}) = a_i + \beta_j,$$

για  $1 \leq i \leq m$ . Περαιτέρω, κατ'αναλογία με το (4.4.2), θεωρείται ότι δεδομένου του  $N_{ij}$

$$(N_{ij0}, \dots, N_{ijd}, N_{ij}^{zero}) \sim \text{Multi}(N_{ij}; p_0, \dots, p_d, Q)$$

για  $(i, j) \in Bm$ . Ορίζουμε  $N_{ij}^{rep} := 0$  για  $j \geq m$  έτσι ώστε επίσης  $(N_{ij0}, \dots, N_{ijd}, N_{ij}^{zero}) = (0, \dots, 0)$  για  $j \geq m$  και  $1 \leq i \leq m$ . Με άλλα λόγια, θεωρείται ότι καμία απαίτηση δεν μπορεί να αναφερθεί με περισσότερες από  $m$  περιόδους μετά την πραγματοποίησή τους. Όπως στην υποπαράγραφο 4.4.2, τα συνολικά καταβληθέντα ποσά, δεδομένου του  $N_{ij}^{paid}$ ,  $(i, j) \in Cm$ , είναι ανεξάρτητα και

$$X_{ij} \sim \Gamma(N_{ij}^{paid} \delta, \nu)$$

όπου  $N_{ij}^{paid} := N_{ij0} + \dots + N_{i,j-d,d}$ ,  $(i, j) \in Cm$  και πάλι χρησιμοποιώντας τη σημείωση από την υποπαράγραφο 4.3.2.

Οι τριγωνικές συστοιχίες  $(\mathfrak{N}_m, \Delta_m)$  είναι υποσύνολα των  $N_{ij}^{rep}$ ,  $(i, j) \in Bm$  και  $X_{ij}$ ,  $(i, j) \in Cm$ . Οι μη παρατηρήσιμες μεταβλητές  $N_{ij}^{rep}$ ,  $(i, j) \in Bm \setminus Am$  και  $X_{ij}$ ,  $(i, j) \in Cm \setminus Am$  αντιπροσωπεύουν τον αριθμό των απαιτήσεων που θα αναφερθούν στο μέλλον και τις μελλοντικές αποζημιώσεις, αντίστοιχα. Είναι φυσικό να χωρίσουμε τις μελλοντικές αποζημιώσεις σε δύο μέρη: αυτές που σχετίζονται με απαιτήσεις που έχουν ήδη αναφερθεί και ως τέτοιες εμφανίζονται στο  $\mathfrak{N}_m$ , και αυτές που θα αναφερθούν στο μέλλον,  $N_{ij}^{rep}$ ,  $(i, j) \in Bm \setminus Am$ .

Επισημώς, αυτά σχετίζονται με τις πληρωμές (RBNS)

$$X_{ij}^{rbns} \sim \Gamma(N_{ij}^{rbns} \delta, \nu)$$

που είναι ανεξάρτητα δεδομένου  $N_{ij}^{rbns} := N_{i,m-i,i+j-m} + \dots + N_{i,\max\{j-d,0\},\min\{d,j\}}$  για  $m - i + 1 \leq j \leq m + i + d$  και  $1 \leq i \leq m$ . Για  $j > m - i + d$  και  $1 \leq i \leq m$ , ας  $X_{ij}^{rbns} := 0$ .

Οι απαιτήσεις που θα αναφερθούν και θα αποζημιωθούν στο μέλλον (IBNR απαιτήσεις) είναι ανεξάρτητες δεδομένου  $N_{ij}^{rbns}$ ,  $(i, j) \in Cm \setminus Am$  και

$$X_{ij}^{ibnr} \sim \Gamma(N_{ij}^{ibnr} \delta, \nu),$$

όπου  $N_{ij}^{ibnr} := N_{ij0} + \dots + N_{i,\max\{m-i+1,j-d\},\min\{j+i-m-1,d\}}$  για  $m - i + 1 \leq j \leq m + d - 1$  και  $2 \leq i \leq m$ . Έχουμε  $X_{ij}^{ibnr} := 0$  για  $j \geq m$ . Τέλος υποθέτουμε ότι οι τυχαίες μεταβλητές  $X_{ij}^{ibnr}$ ,  $m - i + 1 \leq j \leq m - i + d$  και  $X_{ij}^{rbns}$ ,  $m - i + 1 \leq j \leq m - i + d$  είναι αμοιβαία ανεξάρτητες έτσι ώστε

$$X_{ij}^{ibnr} + X_{ij}^{rbns} = X_{ij}$$

για  $(i, j) \in Cm \setminus Am$ .

### 4.6.1 Πρόβλεψη των συνολικών IBNR Απαιτήσεων

Αυτή η υποπαράγραφος εξετάζει την πρόβλεψη του  $X_{ij}^{ibnr}$  για  $m - i + 1 \leq j \leq m - i + d$  και  $2 \leq i \leq m$ . Σε αυτήν την περίπτωση, η πρόβλεψη των συνολικών IBNR απαιτήσεων είναι άμεση λειτουργία των παραμέτρων μας που εκτιμάται από τη μέγιστη πιθανοφάνεια. Η πρόβλεψη  $\hat{\mu}_{ij}^{rbns}$  του  $X_{ij}^{ibnr}$  ορίζεται ως ελαχιστοποιητής του

$$E(X_{ij}^{ibnr} - \mu_{ij}^{ibnr})^2$$

με  $\mu_{ij}^{ibnr}$  όπου το  $\mu_{ij}^{ibnr}$  είναι οποιαδήποτε μετρήσιμη συνάρτηση περιορισμένης διακύμανσης  $(\Delta_m, \mathfrak{X}_m)$ ,  $m - i + 1 \leq j \leq m - i + d$  και  $2 \leq i \leq m$ . Καθώς τα  $X_{ij}^{ibnr}$ ,  $m - i + 1 \leq j \leq m - i + d$  και  $2 \leq i \leq m$  είναι ανεξάρτητα από  $(\mathfrak{X}_m, \Delta_m)$ , ο ελαχιστοποιητής δίνεται από

$$\hat{\mu}_{ij}^{ibnr} = E(X_{ij}^{ibnr}) = \frac{e^{\alpha_i \delta}}{\nu} \sum_{k=0}^{\min\{d, j+i-m-1\}} p_k e^{\beta_{j-k}}. \quad (4.6.1)$$

### 4.6.2 Πρόβλεψη των συνολικών RBNS Απαιτήσεων

Η πρόβλεψη των RBNS απαιτήσεων είναι ασήμαντη, ακόμη και όταν όλες οι στατιστικές παράμετροι εισαγωγής έχουν εκτιμηθεί από τη διαδικασία μέγιστης πιθανοφάνειας. Ο λόγος είναι ότι η καλύτερη κατανομή των RBNS απαιτήσεων που εξαρτάται από τα διαθέσιμα δεδομένα είναι μια περίπλοκη έκφραση. Παρόλο που γνωρίζουμε τον αριθμό των απαιτήσεων πίσω από τις RBNS απαιτήσεις, δεν γνωρίζουμε τον ακριβή αριθμό των απαιτήσεων που έχουν εξοφληθεί αυτή τη στιγμή και δεν γνωρίζουμε το ακριβές ποσό που εξοφλήθηκε. Έχουμε μόνο συγκεντρωτικά δεδομένα που δείχνουν τάσεις, δεν διαθέτουμε τις ακριβείς υποκείμενες πληροφορίες. Στη συνέχεια προκύπτει ο προγνωστικός παράγοντας των RBNS απαιτήσεων ως μέσος όρος υπό όρους με βάση τις παρατηρούμενες πληροφορίες. Οι RBNS απαιτήσεις δεν είναι ανεξάρτητες από  $(\mathfrak{X}_m, \Delta_m)$  που περιπλέκουν ελαφρώς την πρόβλεψη. Για απλότητα αντλούμε τον προβλεπόμενο,  $\mu_{ij}^{rbns}$ , από τις σειρές πληρωμών RBNS,

$$X_{ij}^{rbns} := X_{i, m-i+1}^{rbns} + \dots + X_{i, m+d-i}^{rbns}, \quad 1 \leq i \leq m,$$

αλλά είναι επίσης δυνατό να αντληθούν προβλέψεις του  $X_{ij}^{rbns}$  με παρόμοιο τρόπο. Η πρόβλεψη,  $\hat{\mu}_{ij}^{rbns}$ , του  $X_{ij}^{rbns}$  και πάλι ορίζεται ως ελαχιστοποιητής του

$$E(X_{ij}^{rbns} - \mu_{ij}^{rbns})^2$$

σε όλες τις συναρτήσεις πεπερασμένης διακύμανσης  $(\mathfrak{N}_m, \Delta_m)$ . Καθώς τα  $Y_{ij}^{(k)}$  είναι ανεξάρτητα και ομοιόμορφα κατανομημένα,  $i, k \geq 1$  και  $j \geq 0$ , και ανεξάρτητα από  $\Delta_m$ , είναι ακόλουθο ότι

$$\hat{\mu}_{ij}^{rbns} = \frac{\delta E(N_i^{rbns} | \mathfrak{N}_m; \Delta_m)}{\nu}$$

για  $1 \leq i \leq m$ . Ένας τρόπος προσδιορισμού του  $E(N_i^{rbns} | \mathfrak{N}_m, \Delta_m)$  είναι να προκύψει ένας τύπος για  $P(N_i^{rbns} = k | \mathfrak{N}_m, \Delta_m)$ ,  $0 \leq k \leq N_i$ . όπου  $N_i := N_{i0} + \dots + N_{i,m-i}$  και  $1 \leq i \leq m$ . Ωστόσο, μια εφαρμογή του τύπου Bayes δείχνει ότι οι πιθανότητες είναι

$$P(X_{ij} \in (x_j, x_j + h_j], \xi_i \leq j \leq m - i | N_i^{rbns} = k, \mathfrak{N}_m)$$

εμφανίζονται στην έκφραση για  $E(N_i^{rbns} | \mathfrak{N}_m, \Delta_m)$ . Και αυτές οι πιθανότητες είναι σχετικά περίπλοκες για να προσδιοριστούν. Για να είναι εφαρμόσιμη η πρόβλεψη στην πράξη, προτείνεται ένας άλλος προγνωστικός παράγοντας. Για να λάβουμε σχετικά απλούς τύπους, παρουσιάζουμε τον προγνωστικό παράγοντα,  $\tilde{\mu}_i^{rbns}$ , του  $X_i^{rbns}$  ως ελαχιστοποιητής του

$$E(X_i^{rbns} - \mu_i^{rbns})^2$$

πάνω από όλες τις συναρτήσεις πεπερασμένης διακύμανσης  $\mathfrak{N}_m$  και  $X_i$ ,  $1 \leq i \leq m$ . Η ρύθμιση μόνο στο πόσα έχει καταβληθεί συνολικά,  $X_i$ ,  $1 \leq i \leq m$ , οδηγεί σε πολύ λιγότερες πολύπλοκες εκφράσεις για τους προγνωστικούς παράγοντες,  $\tilde{\mu}_i^{rbns}$ ,  $1 \leq i \leq m$ . Ένα μειονέκτημα είναι ότι οι προγνωστικοί παράγοντες,  $\tilde{\mu}_i^{rbns}$ , φυσικά δεν είναι βέλτιστοι με τη συνήθη έννοια του  $\mathbb{L}^2$ .

Σημειώνουμε με  $k \rightarrow p_{n,q}(k)$  τη μάζα σημείου σε  $k \in \{0, \dots, n\}$  μιας διωνυμικής κατανομής με μήκος  $n$  και πιθανότητα επιτυχίας  $1 - q \in (0,1)$ . Έχουμε το  $f_\Gamma(a,b)$  να είναι η πυκνότητα Lebesgue μιας κατανομής Γάμμα με τις παραμέτρους  $(a,b) \in (0, \infty)^2$ . Για  $0 \leq k \leq N_i$ . μια εφαρμογή του τύπου Bayes αποδίδει το εξής :

$$\begin{aligned} P(N_i^{rbns} = k | \mathfrak{N}_m, X_i \in (x_i, x_i + h]) &= \\ &= \frac{P(N_i^{rbns} = k | \mathfrak{N}_m) P(X_i \in (x_i, x_i + h] | N_i^{rbns} = k, \mathfrak{N}_m)}{P(X_i \in (x_i, x_i + h] | \mathfrak{N}_m)} \end{aligned}$$

για  $x_i, h > 0$  και  $1 \leq i \leq m$ . Με  $k_{m-i} := k - k_{\xi_i} - \dots - k_{m-i-1}$  προκύπτει απευθείας από την Υποπαράγραφο 4.4.2 ότι

$$P(N_i^{rbns} = k | \mathfrak{N}_m) = \sum_{k_{\xi_i}=0}^k \sum_{k_{\xi_i+1}=0}^{k-k_{\xi_i}} \dots \sum_{k_{m-i-1}=0}^{k-k_{\xi_i}-\dots-k_{m-i-2}} \prod_{j=\xi_i}^{m-i} p_{N_{ij}, Q} + \sum_{l=0}^{m-i-j} p_l(k_j).$$

για  $1 \leq i \leq m$ . Ιδιαίτερα με  $i = m$ , τότε  $P(N_i^{rbns} = k | \mathfrak{R}_m) = p_{N_{m0}, Q} + p_1(k)$ . Επιπλέον, για  $1 \leq i \leq m$  έχουμε

$$\begin{aligned} \frac{P(X_i \in (x_i, x_i + h] | N_m^{rbns} = k, \Delta_m)}{P(X_i \in (x_i, x_i + h] | \Delta_m)} &= \frac{P(X_i \in (x_i, x_i + h] | N_m^{rbns} = k, N_i)}{P(X_i \in (x_i, x_i + h] | N_i)} = \\ &= (1 + o(1)) \frac{h \sum_{l=1}^{N_i-k} p_{N_i-k, Q}(1) f_{\Gamma(l\delta, \nu)}(x_i)}{h \sum_{l=1}^{N_i} p_{N_i, Q}(1) f_{\Gamma(l\delta, \nu)}(x_i)} \end{aligned}$$

όταν  $h \rightarrow 0$  και ακολουθεί ότι  $C'_i := (\sum_{l=1}^{N_i} p_{N_i, Q}(1) f_{\Gamma(l\delta, \nu)}(x_i))^{-1}$  και  $\xi_i := \max\{0, m + 1 + d - i\}$  αυτό

$$\begin{aligned} P(N_i^{rbns} = k | \mathfrak{R}_m, X_i) &= \\ = C'_i &\left( \sum_{k_{\xi_i}=0}^k \sum_{k_{\xi_i+1}=0}^{k-k_{\xi_i}} \cdots \sum_{k_{m-i-1}=0}^{k-k_{\xi_i}-\cdots-k_{m-i-2}} \prod_{j=\xi_i}^{m-i} p_{N_{ij}, Q} + \sum_{l=0}^{m-i-j} p_l(k_j) \right) \\ &\times \left( \sum_{l=0}^{N_i-k} p_{N_i-k, Q}(1) f_{\Gamma(l\delta, \nu)}(X_i) \right). \end{aligned}$$

Η πρόβλεψη  $\tilde{\mu}_i^{rbns}$ , του  $X_i^{rbns}$ , δίνεται από

$$\begin{aligned} \tilde{\mu}_i^{rbns} &= \frac{\delta E(N_i^{rbns} | \mathfrak{R}_m, X_i)}{\nu} \tag{4.6.2} \\ &= \frac{\delta}{\nu} \sum_{k=0}^{N_i} k E(N_i^{rbns} | \mathfrak{R}_m, X_i). \end{aligned}$$

Παρατηρούμε ιδιαίτερα ότι  $\tilde{\mu}_i^{rbns} = \hat{\mu}_i^{rbns}$ .

Μια εναλλακτική και απλοϊκή πρόβλεψη των RBNS απαιτήσεων, το  $X_{ij}^{rbns}$ , το οποίο δεν λαμβάνει υπόψη τη δομή εξάρτησης των μοντέλων, δίνεται από

$$\tilde{\mu}_i^{rbns} := \frac{\delta}{\nu} \sum_{k=\xi_i}^{m-i} p_k N_{i, j-k}$$

για  $(i, j) \in Am$ . Θα μπορούσαμε να χρησιμοποιήσουμε το  $\tilde{\mu}_i^{rbns}$  ως πρόβλεψη αναφοράς.

Σε αυτήν την παράγραφο το μοντέλο (4.4.1) και (4.4.2) επεκτάθηκαν με φυσικό τρόπο και σε αυτό το έργο έχει διευκρινιστεί τι είναι οι RBNS και IBNR απαιτήσεις. Τέλος, έχουν προκύψει εκτιμητές που ελαχιστοποιούν τον κανόνα  $\mathbb{L}^2$ . Η πρόβλεψη των RBNS και IBNR απαιτήσεων είναι το τελευταίο βήμα στη θεωρητική ανάλυση του

μοντέλου (4.4.1) και (4.4.2) με βάση τα δεδομένα  $(\mathfrak{K}_m, \Delta_m)$ . Στην επόμενη Παράγραφο πραγματοποιείται εφαρμογή στα δεδομένα από τα αποτελέσματα στις Παραγράφους 5 και 6.

## 4.7 Έρευνα σε Δεδομένα

Σε αυτήν την παράγραφο, το μοντέλο (4.4.1) και (4.4.2) εφαρμόζεται σε ένα σύνολο δεδομένων από την Royal & Sun Alliance. Τα δεδομένα σχετίζονται με ένα χαρτοφυλάκιο ασφαλιστηρίων συμβολαίων για οχήματα, σε αυτό το παράδειγμα τα δεδομένα της ευθύνης τρίτου μέρους (Third Person Liability - TPL) λαμβάνονται υπόψη. Ο λόγος για την επιλογή αυτού του συνόλου δεδομένων είναι ότι αναμένουμε να υπάρχουν αρκετά μεγάλες καθυστερήσεις διακανονισμού (καθυστερήσεις RBNS). Αυτό θα μπορούσε να έχει ιδιαίτερο ενδιαφέρον, καθώς (4.4.1) και (4.4.2) υποδεικνύουν ρητά την καθυστέρηση RBNS. Τα δεδομένα που εμφανίζονται στον Πίνακα 4.1 διορθώνονται ως προς τα δεδομένα, έτσι

$$X_{ij} := \frac{Y_{ij}}{\delta_{i+j}}$$

όπου  $Y_{ij}, (i, j) \in A_{10}$ , είναι οι παρατηρούμενες πληρωμές και το  $\delta_i$  είναι ένας δείκτης πληθωρισμού,  $1 \leq i \leq 10$ . Σε μια πλήρη ανάλυση ενός συνόλου δεδομένων όπως αυτό, ο δείκτης πληθωρισμού θα μπορούσε να μοντελοποιηθεί ανεξάρτητα, για παράδειγμα από μια χρονοσειρά που θα έπρεπε στη συνέχεια να χρησιμοποιηθεί στην πρόβλεψη,  $X_{ij}\delta_j$  για  $j \geq 10 - i + 1$ . Για τους σκοπούς αυτού του κεφαλαίου, υποθέτουμε ότι οι ισχυρισμοί των πληροφοριών έχουν ήδη εκτιμηθεί και επικεντρωνόμαστε στο μοντελοποίηση των διορθωμένων πληροφοριών πληθωρισμού,  $\Delta_{10}$ , οι οποίες φαίνονται στον Πίνακα 1.

$i \setminus j$	0	1	2	3	4	5	6	7	8	9
1	451288	339519	333371	144988	93243	45511	25217	20406	31482	1729
2	448627	512882	168467	130674	56044	33397	56071	26522	14346	
3	693574	497737	202272	120753	125046	37154	27608	17864		
4	652043	546406	244474	200896	106802	106753	63688			
5	566082	503970	217838	145181	165519	91313				
6	606606	562543	227374	153551	132743					
7	536976	472525	154205	150564						
8	554833	590880	300964							
9	537238	701111								
10	684944									

Πίνακας 4.1. Τρίγωνο εξέλιξης πληρωμών,  $X_{ij}, (i, j) \in A_{10}$  για τα δεδομένα TPD

Οι μετρήσεις εμφανίζονται στον Πίνακα 4.2.

$i \setminus j$	0	1	2	3	4	5	6	7	8	9
1	6238	831	49	7	1	1	2	1	2	3
2	7773	1381	23	4	1	3	1	1	3	
3	10306	1093	17	5	2	0	2	2		
4	9639	995	17	6	1	5	4			
5	9511	1386	39	4	6	5				
6	10023	1342	31	16	9					
7	9834	1424	59	24						
8	10899	1503	84							
9	11954	1704								
10	10989									

Πίνακας 4.2. Αριθμοί των αναφερθέντων απαιτήσεων,  $N_{ij}, (i, j) \in A_{10}$  για τα δεδομένα TPD

Οι εξειδικευμένες συμβουλές από έναν χειριστή απαιτήσεων έχουν χρησιμοποιηθεί για τον προσδιορισμό του κλάσματος των αναφερθέντων μηδενικών απαιτήσεων,  $Q \in [0,1)$  και της μέγιστης πιθανής RBNS καθυστέρησης,  $d \leq 10$ . Σε αυτήν την περίπτωση  $(Q, d) = (0,2,7)$ .

Η εκτίμηση στο μοντέλο (4.4.1) και (4.4.2) έγινε με τη μέθοδο της Chain Ladder όπως περιγράφεται στην παράγραφο 5 και στο Παράρτημα Β αντίστοιχα. Η πρόβλεψη των IBNR και RBNS απαιτήσεων πραγματοποιήθηκε όπως προτάθηκε στις (4.6.1) και (4.6.2).

#### 4.7.1 Εκτίμηση στο μοντέλο (4.4.1) και (4.4.2)

Οι παράμετροι στο μοντέλο (4.4.1) υπολογίζονται μεγιστοποιώντας τη συνάρτηση πιθανότητας,  $L_{\mathbb{N}_m}(\gamma, \{\beta, 0 \leq j \leq 9\}, \{\alpha_i, 1 \leq i \leq 10\})$ , στο (4.5.1).

Ορισμένες από τις τιμές  $N_{i0}, 1 \leq i \leq 10$ , υπερβαίνουν τα 10000. Επομένως, η μεγιστοποίηση του (4.5.2) θα απαιτούσε μεγάλο χρόνο στον υπολογισμό. Αντίθετα, συνάρτηση πιθανότητας στο (4.9.1) στο Παράρτημα Β για την εκτίμηση των παραμέτρων  $(\delta, \nu, p_0, \dots, p_7)$ .

Οι βελτιστοποιήσεις πραγματοποιούνται στο λογισμικό R. Τα αποτελέσματα εμφανίζονται στον Πίνακα 4.3.

$j$	0	1	2	3	4	5	6	7	8	9
$e^{-b}$	0.875	0.118	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000
$p$	0.291	0.230	0.091	0.068	0.053	0.029	0.020	0.018		



Πίνακας 4.3. Η σειρά που δηλώνεται από το  $e^b$  είναι οι εκτιμήσεις μέγιστης πιθανοφάνειας των  $e^{\beta_j}$ ,  $0 \leq j \leq 9$  και η σειρά  $p$  είναι οι εκτιμήσεις μέγιστης πιθανοφάνειας  $p_l$ ,  $0 \leq l \leq 7$ .

Να σημειωθεί ότι το μέρος της απαίτησης που διακανονίζεται στην τιμή μηδέν έχει οριστεί σε  $Q = 0.2$  έτσι

$$p_0 + \dots + p_7 = 0.8$$

Οι εκτιμήσεις των  $(\nu, \delta)$  είναι

$$(\hat{\nu}, \hat{\delta}) = (0.0170, 9.395e - 5)$$

έτσι ώστε η κατανομή της απλής (μη παρατηρουμένης) απαίτησης,  $Y_{ijl}^{(k)}$ , εκτιμάται ότι είναι  $\Gamma(0, 0.0170, 8.395e - 5)$ .

Όσον αφορά την αρχή της παραγράφου 7, οι TPL απαιτήσεις αναμένεται να έχει σχετικά μεγάλες καθυστερήσεις διακανονισμού (RBNS καθυστερήσεις) καθώς οι απαιτήσεις σωματικού τραυματισμού χρειάζονται συχνά πολύ χρόνο για να διευθετηθούν. Όπως φαίνεται από την αριστερή γραφική παράσταση στο Σχήμα 4.1 υπάρχει εμπειρική απόδειξη αυτού: πολύ καιρό μετά την αναφορά των περισσότερων απαιτήσεων, εξακολουθούν να υπάρχουν σημαντικές αποζημιώσεις. Σημειώνουμε επίσης ότι το  $p_0$  μπορεί να είναι σχετικά μικρό. Καθώς συμβαίνουν απαιτήσεις, κατά μέσο όρο, στα μέσα του έτους υπάρχει κατά μέσο όρο μόνο μισό έτος για να λάβουμε την τελική αποζημίωση, προκειμένου να ολοκληρωθεί στην κατηγορία απαιτήσεων που σχετίζονται με το  $p_0$ . Όλες οι άλλες περιόδους καθυστέρησης είναι πλήρεις έτη.

Λαμβάνοντας υπόψη το  $N_{1,}$ , έχουμε

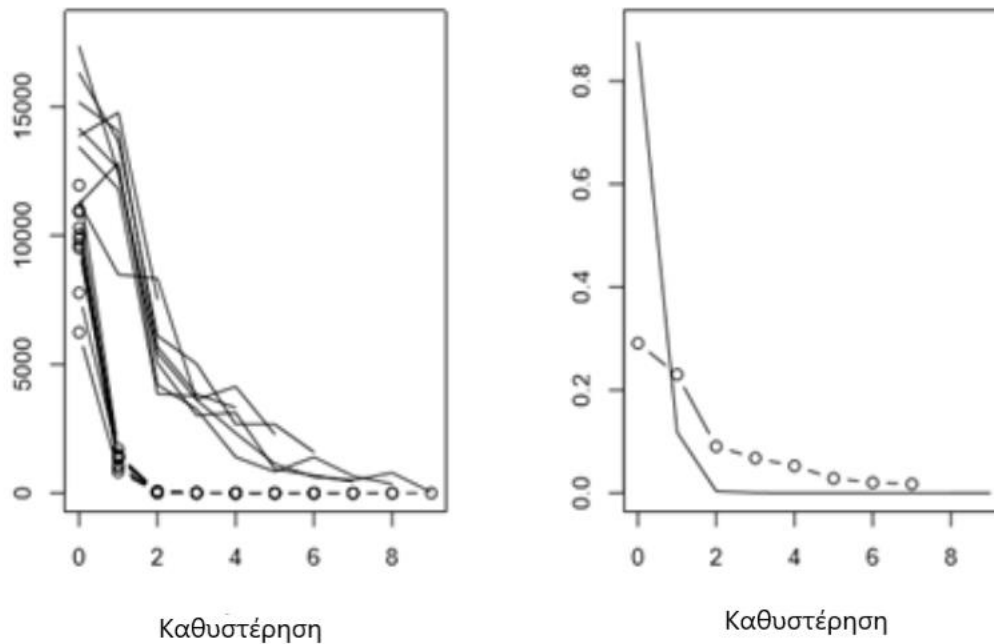
$$(N_{11}, \dots, N_{1m}) \sim \text{Multi}(N_{1,}, e^{\beta_0}, \dots, e^{\beta_9}).$$

Φαίνεται λοιπόν φυσικό να συγκρίνουμε τις εκτιμήσεις των IBNR και RBNS καθυστερήσεων λαμβάνοντας υπόψη τα  $p'_j := e^{\beta_j}$ ,  $0 \leq j \leq 9$  και  $p_l$ ,  $0 \leq l \leq 7$ . Αυτή η σύγκριση γίνεται στο δεξιό διάγραμμα στο Σχήμα 4.1. Με αυτόν τον τρόπο, η μέση IBNR καθυστέρηση που σχετίζεται με  $p'_j$ ,  $0 \leq j \leq 9$  είναι 0.129 έτη ενώ η μέση RBNS καθυστέρηση είναι 1.22. Ως εκ τούτου, το RBNS απόθεμα αναμένεται να είναι (περίπου δέκα φορές) μεγαλύτερο από το IBNR απόθεμα, επειδή οι μεμονωμένες απαιτήσεις υποθέτουμε ότι είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες.

### 4.7.2 Εγκυρότητα μοντέλου

Σε αυτήν την ενότητα επαληθεύουμε εάν το μοντέλο (4.4.1) και (4.4.2) φαίνεται να είναι συνεπές με το σύνολο δεδομένων στους Πίνακες 4.1 και 4.2. Καθώς οι τυχαίες μεταβλητές  $N_{ij}$ ,  $X_{ij}$ ,  $(i, j) \in A_m$ , έχουν όλες διαφορετικές κατανομές η ακόλουθη ιδέα χρησιμοποιείται για την επικύρωση του μοντέλου.

Ένα PP-plot κατασκευάζεται με τον ακόλουθο τρόπο. Υποδηλώνουμε τη συνάρτηση κατανομής των τυχαίων μεταβλητών,  $Z_j$ , με  $F_{Z_j}$ ,  $1 \leq j \leq N$ . Έπειτα  $U_j := F_{Z_j}(Z_j) \sim Unif(0,1)$  με την προϋπόθεση ότι τα  $F_{Z_j}$  είναι συνεχή και για το ζητούμενο δείγμα  $U_{(1)} \leq U_{(2)} \leq \dots \leq U_{(N)}$  υποστηρίζει ότι η  $EU_j = j/(N+1)$  για  $1 \leq j \leq N$ . Ένα PP-plot αποτελείται από τα σημεία  $((j/(N+1), EU_{(j)}))$  που προσεγγίζουν τη γραμμή  $\{(x, y) \in [0,1]^2: x = y\}$ . Στην αριστερή γραφική παράσταση στο Σχήμα 4.2 τα σημεία  $k/56$ ,  $1 \leq k \leq 55$  παριστάνονται γραφικά έναντι του ταξινομημένου δείγματος του  $\hat{F}_{N_{ij}}(N_{ij})$ ,  $(i, j) \in A_{10}$ , όπου  $\hat{F}_{N_{ij}}$  είναι η εκτιμώμενη συνάρτηση κατανομής του  $N_{ij}$  (από την υποπαράγραφο 4.7.1). Το δεξί PP-plot αντιστοιχεί στις μεταβλητές  $F_{(X_{ij}|\Delta_m)}(X_{ij})$ ,  $(i, j) \in A_{10}$ . Φαίνεται ότι υπάρχει κάποιο συστηματικό σφάλμα στο αριστερό διάγραμμα στο Σχήμα 4.2, αλλά το δεξί PP-plot φαίνεται πειστικό.



Σχήμα 4.1. Αριστερά: Οι διακεκομμένες γραμμές αντιπροσωπεύουν τα σημεία  $\{X_{ij}/40, 0 \leq j \leq 9\}$  για  $1 \leq i \leq 10$ . Τα γραφήματα είναι ο αριθμός των αναφερθέντων απαιτήσεων,  $\{N_{ij}, 0 \leq j \leq 9\}$  για  $1 \leq i \leq 9$ . Δεξιά: Οι εκτιμήσεις του  $(p_0, \dots, p_7)$  που αντιπροσωπεύουν την RBNS καθυστέρηση είναι τα διακεκομμένα γραφήματα και τα συμπαγή γραφήματα είναι η IBNR καθυστέρηση,  $p'_j, 0 \leq j \leq 9$ .

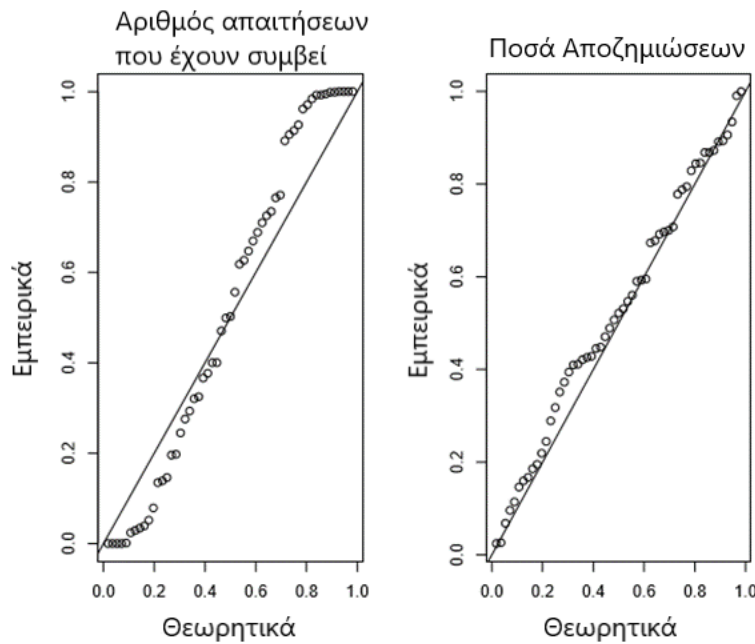
Για να απεικονίσουμε πιθανές αλλαγές στη δομή των δεδομένων με την πάροδο του χρόνου, έχουμε επίσης σχεδιάσει το  $F_{N_{ij}}(N_{ij})$  και το  $F_{X_{ij}}(X_{ij})$  έναντι των δεικτών γραμμής και στήλης αντίστοιχα, Σχήμα 4.3. Σημειώνουμε συγκεκριμένα ότι η κάτω

δεξιά γραφική παράσταση δείχνει ότι είτε  $p_l$ ,  $0 \leq l \leq d$ , ή η μέση απαίτηση,  $EY_{1j}^{(1)}$ , αλλάζει ως συνάρτηση της στήλης,  $0 \leq j \leq m - 1$ .

Τα γραφήματα δείχνουν μία λογική προσαρμογή του μοντέλου (4.4.1) και (4.4.2) στα δεδομένα στους Πίνακες 4.1 και 4.2. Στην επόμενη ενότητα συνεχίζουμε λοιπόν να προβλέψουμε τις IBNR και RBNS απαιτήσεις.

### 4.7.3 Εκτίμηση των IBNR και RBNS απαιτήσεων

Σε αυτήν την ενότητα προβλέπουμε τις IBNR και RBNS απαιτήσεις όπως προτείνονται στα (4.6.1) και (4.6.2) αντίστοιχα. Επιλέξαμε να λάβουμε υπόψη μόνο τις σειρές  $i = 4, \dots, 10$  επειδή η πλειονότητα του συνολικού αποθέματος απαιτήσεων σχετίζεται με αυτές. Λαμβάνοντας υπόψη τις εκτιμήσεις παραμέτρων στην υποπαράγραφο 4.7.1, το εκτιμώμενο αποθεματικό δίνεται στον Πίνακα 4.4.



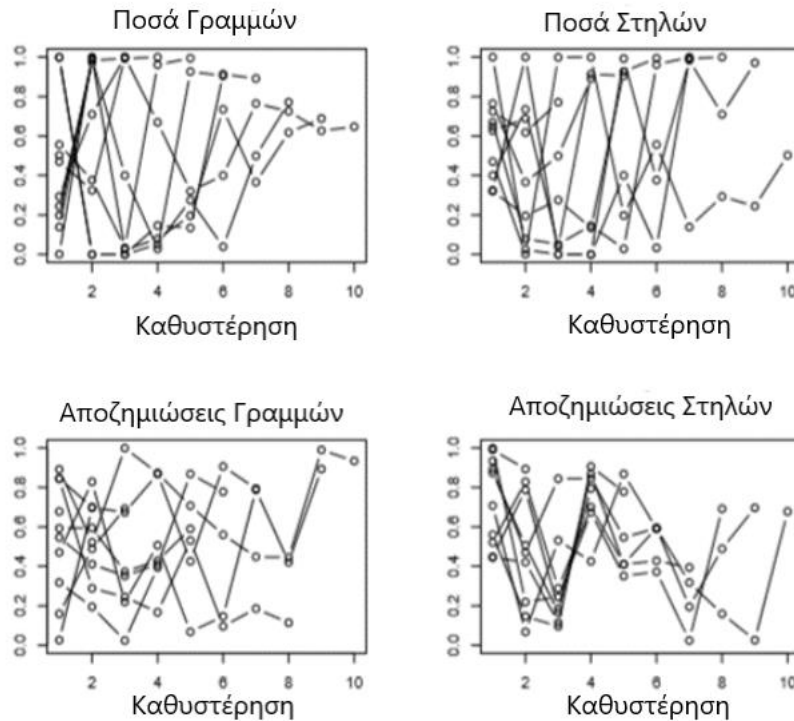
Σχήμα 4.2. Αριστερά: Ένα PP-plot για  $\alpha_m$ . Δεξιά: PP-plot για  $\Delta_m$  δοθέντος  $\mathfrak{K}_m$ .

$i$	10	9	8	7	6	5	4
IBNR	317326	17805	6670	3950	3215	2453	1883
RBNS	1137206	853576	511176	298823	168321	92933	57285
Απόθεμα	1454532	871381	517846	302773	171536	95386	59168

Πίνακας 4.4. Οι διαδοχικές εκτιμήσεις αποθεμάτων χωρίζονται σε IBNR και RBNS απαιτήσεις.

Είδαμε στην υποπαράγραφο 4.7.1 ότι η IBNR καθυστέρηση είναι (κατά μέσο όρο) μικρότερη από την RBNS καθυστέρηση, και ως εκ τούτου το RBNS απόθεμα αναμένεται να είναι μεγαλύτερο από το IBNR απόθεμα. Οι πραγματικές εκτιμήσεις διαιρούν τα αποθέματα έτσι ώστε το RBNS απόθεμα να καταλαμβάνει το 91,6% του συνολικού αποθεματικού και το IBNR απόθεμα μόνο 8,4%: περίπου 1:10 όπως προτείνεται στην υποπαράγραφο 4.7.1.

Όπως προτείνεται στην υποπαράγραφο 4.6.2, είναι σημαντικό να συγκρίνουμε τον πιο εξελιγμένο εκτιμητή  $\tilde{\mu}_i^{rbns}$ , με τον απλό εκτιμητή  $\tilde{\mu}_{ij}^{rbns}$ ,  $1 \leq i \leq 10$ . Αυτό γίνεται στον Πίνακα 4.5 παρακάτω.



Σχήμα 4.3. Επάνω αριστερά: τα σημεία  $\{(j, F_{N_{ij}}(N_{ij}), 1 \leq j \leq m - i\}, 1 \leq i \leq m$ . Επάνω δεξιά: τα σημεία  $\{(i, F_{N_{ij}}(N_{ij}), 1 \leq i \leq m - j\}, 0 \leq j \leq m - 1$ . Κάτω αριστερά: τα σημεία  $\{(j, F_{X_{ij}}(X_{ij}), 1 \leq j \leq m - i\}, 1 \leq i \leq m$ . Επάνω δεξιά: τα σημεία  $\{(i, F_{X_{ij}}(X_{ij}), 1 \leq i \leq m - j\}, 0 \leq j \leq m - 1$ .

i	10	9	8	7	6	5	4
$\mu_i(\sim)/(\delta/\nu)$	5601	4204	2518	1472	829	458	282
$\mu_i(v)/(\delta/\nu)$	5594	4197	2509	1473	845	466	215

Πίνακας 4.5. Η πρώτη σειρά είναι ο αριθμός των RBNS απαιτήσεων που εκτιμάται χρησιμοποιώντας την (4.6.2). Η κάτω σειρά είναι ο αριθμός των RBNS απαιτήσεων που υπολογίζονται με  $\tilde{\mu}_{i0}^{rbns} + \dots + \tilde{\mu}_{i,9-i}^{rbns}$ ,  $4 \leq i \leq 10$ , όπου  $\tilde{\mu}_{ij}^{rbns}$ ,  $(i, j) \in Am$ , είναι όπως ορίζεται στην υποπαράγραφο 4.6.2.

Λέγεται στην εισαγωγή ότι η χρήση περισσότερων δεδομένων από, για παράδειγμα, την Chain Ladder μέθοδο επιτρέπει μια πιο εξελιγμένη ανάλυση. Ωστόσο, όπως γίνεται συχνά, μπορούμε να χρησιμοποιήσουμε τη Chain Ladder μέθοδο ως σημείο αναφοράς στο οποίο μπορεί κανείς να συσχετίσει τα αποτελέσματα σε αυτήν την ενότητα. Στην επόμενη υποπαράγραφο συγκρίνουμε τα αποτελέσματα από το μοντέλο (4.4.1) και (4.4.2) και τις εκτιμήσεις για το Chain Ladder απόθεμα. Θα μπορούσε κάποιος να χρησιμοποιήσει την bootstrap μέθοδο για να βρει τα χαρακτηριστικά κατανομής για τις εκτιμήσεις της Chain Ladder μεθόδου. Οι προβλεπόμενες κατανομές στο μοντέλο (4.4.1) και (4.4.2) δίδονται ρητά στις υποπαράγραφους 4.6.1 και 4.6.2.

#### 4.7.4 Σύγκριση με τα αποτελέσματα της Chain Ladder Μεθόδου

Μία από τις πιο δημοφιλείς μεθόδους για τον υπολογισμό των αποθεμάτων στην πράξη είναι η μέθοδος Chain Ladder. Χρησιμοποιώντας το τρίγωνο εξέλιξης αποζημιώσεων,  $\Delta_m$ , είναι εύκολο να εκτιμήσουμε τα αποθέματα. Τα αποθέματα της Chain Ladder μεθόδου περιλαμβάνουν τόσο IBNR όσο και μέρος των RBNS απαιτήσεων. Δεν είναι όμως δυνατόν να χωριστούν ρητά. Οι εκτιμήσεις αποθεμάτων δίνονται στον Πίνακα 4.6.

i	10	9	8	7	6	5	4
Απόθεμα	1459860	763919	475992	249349	173802	101158	53791

Πίνακας 4.6. Εκτιμήσεις των Chain Ladder αποθεμάτων.

Οι εκτιμήσεις για το Chain Ladder απόθεμα είναι κοντά στο αποτέλεσμα της υποπαράγραφου 4.7.3. Το συνολικό Chain Ladder απόθεμα για τις σειρές  $i = 4, \dots, 10$ , είναι 3277871 σε σύγκριση με το 3472622 στην υποπαράγραφο 4.7.3. Η απόκλιση θα

μπορούσε να εξηγηθεί από το γεγονός ότι δεν εκτιμώνται όλες οι RBNS απαιτήσεις με τη Chain Ladder μέθοδο. Θα μπορούσε επίσης να οφείλεται στην αβεβαιότητα του μοντέλου.

Σε σύνολα δεδομένων με σχετικά λίγες απαιτήσεις, υπάρχουν εμπειρικές ενδείξεις ότι η εκτίμηση στο (4.4.1) και (4.4.2) δίνει σφάλμα εκτίμησης των εκτιμήσεων των αποθεμάτων που είναι μικρότερο από ότι σε σύγκριση με την Chain Ladder.

## 4.8 Παράρτημα Α

Σε αυτήν την ενότητα υποθέτουμε το μοντέλο (4.4.1) και (4.4.2).

### 4.8.1 Η συνάρτηση πιθανοφάνειας για τα $\aleph_m$

Στα κείμενα των Hachemeister και Standard (1975) και Mack (1991) φαίνεται ότι

$$e^{\hat{c}_{i,m-1}} = \sum_{j=0}^{m-1} e^{\hat{a}_i + \hat{\beta}_j}, \quad 1 \leq i \leq m$$

όπου  $e^{\hat{a}_i}$  και  $e^{\hat{\beta}_j}$  δηλώνουν τους εκτιμητές της μέγιστης πιθανότητας των  $e^{a_i}$  και  $e^{\beta_j}$ . Δεδομένου ότι το μοντέλο είναι αναγνωρίσιμο με την επιβολή του περιορισμού

$$\sum_{j=0}^{m-1} e^{\beta_j} = 1$$

λαμβάνουμε άμεσα

$$e^{\hat{c}_{i,m-1}} = e^{\hat{a}_i}, \quad 1 \leq i \leq m$$

Επιπλέον, ακολουθεί άμεσα αυτό

$$\frac{e^{\hat{c}_{m,j}} - e^{\hat{c}_{m,j-1}}}{e^{\hat{a}_m}} = \frac{1}{e^{\hat{a}_m}} \left( \sum_{k=0}^j e^{\hat{a}_m + \hat{\beta}_k} - \sum_{k=0}^{j-1} e^{\hat{a}_m + \hat{\beta}_k} \right) = e^{\hat{\beta}_j}$$

για  $0 \leq j \leq m - 1$ .

### 4.8.2 Η συνάρτηση πιθανοφάνειας για $\Delta_m$ δοθέντος $\aleph_m$

Εκτός από τους όρους της προηγούμενης υποπαραγράφου υποθέτουμε ότι τα ποσά πληρωμών,  $X_{ij} = x_{ij}$ ,  $(i, j) \in Am$ , είναι επίσης γνωστά. Προσδιορίζουμε  $x_i := x_{i0} + \dots + x_{i,m-i}$ ,  $n_{ij} := n_{ij0} + \dots + n_{ijd}$  και  $n_{ij}^{paid} := n_{ij0} + n_{i,j-1,1} + \dots + n_{i,j-d,d}$  όπου  $n_{ijl} = 0$  για  $l < 0$  κατά σύμβαση. Η συνάρτηση υπό όρους πιθανοφάνειας του  $\Delta_m$  δοθέντος  $\aleph_m$  που ορίζεται στο (4.5.1) δίδεται ως

$$\mathcal{L}_{\Delta_m|\mathfrak{K}_m}(\delta, \nu, p_0, \dots, p_d) = \prod_{i=1}^m \prod_{j=0}^{m-i} \sum_{n_{ij0}=0}^{n_{ij}} \sum_{n_{ij1}=0}^{n_{ij}-n_{ij0}} \dots \sum_{n_{ijd}=0}^{n_{ij}-n_{ij.}} \left[ \binom{n_{ij}}{n_{ij0} \dots n_{ijd} \cdot (n_{ij} - n_{ij.})} p_0^{n_{ij0}} \dots p_d^{n_{ijd}} Q^{n_{ij}-n_{ij.}} \right. \\ \left. \prod_{k=0}^{m-i} \frac{\nu^{n_{ik}^{paid} \delta}}{\Gamma(n_{ik}^{paid} \delta)} x_{ik}^{n_{ik}^{paid} \delta - 1} e^{-\nu x_{ik}} \right].$$

Εδώ

$$\prod_{k=0}^{m-i} \frac{\nu^{n_{ik}^{paid} \delta}}{\Gamma(n_{ik}^{paid} \delta)} x_{ik}^{n_{ik}^{paid} \delta - 1} e^{-\nu x_{ik}} = e^{-\nu x_{i.} - \sum_{j=0}^{m-i} \log(x_{ij})} \prod_{k=0}^{m-i} \frac{\nu^{n_{ik}^{paid} \delta}}{\Gamma(n_{ik}^{paid} \delta)} x_{ik}^{n_{ik}^{paid} \delta}$$

και

$$\prod_{l=0}^d p_l^{n_{ijl}} Q^{n_{ij}-n_{ij.}} = Q^{n_{ij}} \prod_{l=0}^d \left(\frac{p_l}{Q}\right)^{n_{ijl}}$$

τέτοια ώστε το  $\mathcal{L}_{\Delta_m|\mathfrak{K}_m}$  μπορεί να γραφτεί ως

$$\prod_{i=1}^m e^{-\nu x_{i.} - \sum_{j=0}^{m-i} \log(x_{ij})} \sum_{j=0}^{m-i} Q^{n_{ij}} n_{ij}! \left[ \sum_{n_{ij0}=0}^{n_{ij}} \sum_{n_{ij1}=0}^{n_{ij}-n_{ij0}} \dots \sum_{n_{ijd}=0}^{n_{ij}-n_{ij.}} \left( \prod_{l=0}^d \left(\frac{p_l}{Q}\right)^{n_{ijl}} \frac{1}{n_{ijl}!} \right) \right. \\ \left. \frac{1}{(n_{ij} - n_{ij.})!} \prod_{k=0}^{m-i} \frac{(\nu x_{ik})^{n_{ik}^{paid} \delta}}{\Gamma(n_{ik}^{paid} \delta)} \right].$$

Λαμβάνοντας λογάριθμους, αποκτάται τελικά

$$\log(\mathcal{L}_{\Delta_m|\mathfrak{K}_m}(\delta, \nu, p_0, \dots, p_d)) = C' + \sum_{i=1}^m \left[ -\nu x_{i.} + \log \left( \sum_{j=0}^{m-i} Q^{n_{ij}} n_{ij}! \left[ \sum_{n_{ij0}=0}^{n_{ij}} \left(\frac{p_0}{Q}\right)^{n_{ij0}} \frac{1}{n_{ij0}!} \right. \right. \right. \\ \left. \left. \left. \sum_{n_{ij1}=0}^{n_{ij}-n_{ij0}} \left(\frac{p_1}{Q}\right)^{n_{ij1}} \frac{1}{n_{ij1}!} \dots \sum_{n_{ijd}=0}^{n_{ij}-n_{ij.}} \left(\frac{p_d}{Q}\right)^{n_{ijd}} \frac{1}{n_{ijd}!} \frac{1}{(n_{ij} - n_{ij.})} \prod_{k=0}^{m-i} \frac{(\nu x_{ik})^{n_{ik}^{paid} \delta}}{\Gamma(n_{ik}^{paid} \delta)} \right] \right) \right]$$

όπου  $C' = \sum_{i=1}^m \sum_{j=0}^{m-i} \log(x_{ij})$ .

## 4.9 Παράρτημα Β

Σε αυτήν την ενότητα προτείνουμε μια προσέγγιση προσέγγισης Γενικευμένου Γραμμικού Μοντέλου για την εκτίμηση των παραμέτρων  $(\nu, \delta, p_0, \dots, p_d)$ . Η Ενότητα 9 στον Neuhaus W (2004) χρησιμοποιείται για τη δημιουργία μιας συνάρτησης πιθανοφάνειας. Χρειαζόμαστε, ωστόσο, την ακόλουθη σύμβαση: ισχύει η σχέση  $VY_{11}^{(1)}/EY_{11}^{(1)} \gg \max_{l=0, \dots, d}(1 - p_l)$ , όπου  $\gg$  σημαίνει ότι η δεξιά πλευρά είναι ανεπιθύμητη σε σχέση με την αριστερή πλευρά. Αυτό παρατηρείται εμπειρικά. Στην Ενότητα 9.3 στον Neuhaus W (2004) προτείνεται ότι οι τυχαίες μεταβλητές  $X_{ij}$  θα πρέπει να θεωρούνται (υπό όρους δοθέντος  $\kappa_m$ ) ανεξάρτητες καθώς η δομή συνδιακύμανσης των  $\Delta_m$  δοθέντος  $\kappa_m$  καθορίζεται πλήρως από τις παραμέτρους  $(\nu, \delta, p_0, \dots, p_d)$ . Χρησιμοποιώντας τον Neuhaus W (2004), σελ. 326, η εκτίμηση βασίζεται αποκλειστικά στη σχέση μεταξύ του μέσου όρου και της διακύμανσης των  $X_{ij}$ ,  $(i, j) \in Am$  δεδομένου  $\Delta_m$ . Για  $(i, j) \in Am$  βλέπουμε ότι

$$E(X_{ij}|\kappa_m) = \sum_{l=0}^d N_{i,j-l}^{rep} \psi_l$$

$$V(X_{ij}|\kappa_m) = \sum_{l=0}^d N_{i,j-l}^{rep} \psi_l \left[ \frac{VY_{11}^{(1)}}{EY_{11}^{(1)}} + (1 - p_l) \right] \cong \phi \sum_{l=0}^d N_{i,j-l}^{rep} \psi_l, \quad (4.9.1)$$

όπου  $\psi_l = p_l EY_{11}^{(1)}$ ,  $0 \leq l \leq d$ , και  $\phi = VY_{11}^{(1)}/EY_{11}^{(1)}$ . Ο κατά προσέγγιση τύπος (4.9.1) συνδυάζεται με το χωρίο 6.2.3 και τον Πίνακα 9.1 από τον Neuhaus W (2004) προτείνει την ακόλουθη συνάρτηση λογαριθμικής πιθανοφάνειας υπό όρους για  $\Delta_m$  δοθέντος  $\kappa_m$ ,

$$\sum_{(i,j) \in Am} \left[ X_{ij} \log \left( \sum_{l=0}^d N_{i,j-l}^{rep} \psi_l \right) - \sum_{l=0}^d N_{i,j-l}^{rep} \psi_l \right].$$

Δηλώστε τους μεγιστοποιητές της τελευταίας ποσότητας  $\hat{\psi}_l$ ,  $0 \leq l \leq d$ . Ακολουθώντας την Ενότητα 9.2 στον Neuhaus W (2004), ο λόγος,  $\phi$ , μπορεί να εκτιμηθεί από

$$\hat{\phi} = \frac{2}{m(m+1)} \sum_{(i,j) \in Am} \frac{(X_{ij} - \sum_{l=0}^d N_{i,j-l}^{rep} \hat{\psi}_l)^2}{\sum_{l=0}^d N_{i,j-l}^{rep} \hat{\psi}_l}.$$

Οι πραγματικές εκτιμήσεις παραμέτρων λαμβάνονται στη συνέχεια με τη μέθοδο των ροπών:

$$\hat{\nu} = \frac{\sum_{k=0}^d \hat{\psi}_k}{\hat{\phi}(1-Q)}, \quad \hat{\delta} = \hat{\nu} \frac{\sum_{k=0}^d \hat{\psi}_k}{(1-Q)},$$



$$p_l = \frac{\hat{\psi}_l(1-Q)}{\sum_{k=0}^d \hat{\psi}_k}, \quad 1 \leq l \leq d. \quad (4.9.2)$$

## Κεφάλαιο Πέμπτο

### 5. Πρόβλεψη αποζημιώσεων σε ένα μοντέλο συμπλέγματος Poisson

#### 5.1 Περίληψη

Θεωρούμε ένα απλό μοντέλο συμπλέγματος Poisson για τους αριθμούς αποζημιώσεων και τις αντίστοιχες συνολικές αποζημιώσεις για ασφαλιστικές απαιτήσεις σε ένα δεδομένο έτος. Λόγω της δομής Poisson μπορεί κανείς να δώσει λογικά σαφείς εκφράσεις για την πρόβλεψη των αριθμών αποζημίωσης και των συνολικών αποζημιώσεων σε μελλοντικές περιόδους, λαμβάνοντας υπόψη τις προηγούμενες παρατηρήσεις των αριθμών αποζημίωσης. Κάποιος μπορεί επίσης να αντλήσει εύλογες ρητές εκφράσεις για τα αντίστοιχα σφάλματα πρόβλεψης. Στην (a,b)-κατηγορία κατανομών μεγέθους απαιτήσεων του Panjer, αυτές οι εκφράσεις μπορούν να αξιολογηθούν με απλούς αναδρομικούς αλγόριθμους. Μελετάμε τις συνθήκες υπό τις οποίες οι προβλέψεις είναι ασυμπτωτικά γραμμικές καθώς ο αριθμός των προηγούμενων αποζημιώσεων γίνεται μεγάλος. Αποδεικνύουμε επίσης ότι, σε άλλα καθεστάτα, η πρόβλεψη ίσως απέχει πολύ από τη γραμμική. Για παράδειγμα, μπορεί να εμφανιστεί και μοτίβο τύπου αλυσίδας. Απεικονίζουμε πώς λειτουργεί η θεωρία σε πραγματικά δεδομένα, επίσης σε σύγκριση με τη μέθοδο Chain Ladder.

#### 5.2 Εισαγωγή

Έχουμε το  $N_k$  να είναι ο αριθμός των αποζημιώσεων για απαιτήσεις που έφτασαν σε ένα ασφαλιστικό χαρτοφυλάκιο το έτος 0 και εκτελέστηκαν το έτος  $k \in \{0,1,\dots\}$ . Επιπλέον, έχουμε το  $S_k$  να είναι το αντίστοιχο συνολικό ποσό των αποζημιώσεων που πραγματοποιήθηκαν το έτος  $k$ . Εάν κάποιος έχει παρατηρήσει τις μετρήσεις  $N_k$ ,  $k = 0, \dots, j$ , για κάποια  $j \geq 0$ , ένα σημαντικό πρόβλημα για μια ασφαλιστική εταιρεία είναι να καθορίσει ένα απόθεμα για τα έτη  $j+1, j+2, \dots$ . Αυτό ισοδυναμεί με την πρόβλεψη των ζευγών  $(N_{j+\ell+1}, S_{j+\ell+1})$  για  $\ell = 0,1,\dots$ . Σε αυτό το πλαίσιο φυσικοί εκτιμητές δίνονται από τις υπό όρους προσδοκίες δεδομένων των προηγούμενων τιμών  $N_0, \dots, N_j$ , δηλ.,

$$\hat{N}_{j+\ell+1} = E(N_{j+\ell+1} | \mathcal{F}_j) \quad \text{και} \quad \hat{S}_{j+\ell+1} = E(S_{j+\ell+1} | \mathcal{F}_j), \quad (5.2.1)$$

$\ell = 0,1,\dots$  όπου  $\mathcal{F}_j = \sigma(N_0, \dots, N_j)$ ,  $j = 0,1,\dots$ . Υποθέτοντας ότι  $\text{var}(S_{j+\ell+1}) < \infty$ ,  $S = \hat{S}_{j+\ell+1}$  είναι ο ασυμπτωτικός μοναδικός ελαχιστοποιητής του μέσου τετραγωνικού σφάλματος  $E((S_{j+\ell+1} - S)^2)$  στην τάξη των τετραγωνικών ενσωματώσιμων τυχαίων

μεταβλητών  $S$  που είναι μετρήσιμες συναρτήσεις των  $N_0, \dots, N_j$ , και μια παρόμοια παρατήρηση ισχύει για  $\hat{N}_{j+\ell+1}$ . Επιπλέον, το  $s = \hat{S}_{j+\ell+1}$  ελαχιστοποιεί το υπό όρους μέσο τετραγωνικό σφάλμα  $E((S_{j+\ell+1} - S)^2 | \mathcal{F}_j)$ .

Μία από τις δημοφιλείς διαδικασίες σε αυτό το πλαίσιο προτάθηκε από τον Mack, βλέπε Mack (1993,1994,1994) και Mack και Braun (2006). Στην απλούστερη εκδοχή του, η διαδικασία του Mack δηλώνει ότι οι προβλέψεις  $\hat{S}_{j+\ell+1}$  και  $\hat{N}_{j+\ell+1}$  είναι γραμμικές συναρτήσεις του  $S_0 + \dots + S_j$  ή  $N_j = N_0 + \dots + N_j$ , αντίστοιχα. Για παράδειγμα,

$$\hat{N}_{j+1} = (f_j - 1)(N_0 + \dots + N_j), \quad j = 0, 1, \dots, \quad (5.2.2)$$

για σταθερές  $f_j \geq 1$ .

Ας υποθέσουμε ότι παρατηρεί κανείς το τρίγωνο εξέλιξης

$$(N_{i,i+k}, S_{i,i+k}), \quad i = 1, \dots, n, \quad k = 0, \dots, n-1, \quad (5.2.3)$$

όπου  $((N_{i,i+k}, S_{i,i+k}))_{k=0, \dots, j-i}$  είναι τα ποσά των αποζημιώσεων και οι συνολικές αποζημιώσεις για αιτήσεις που έχουν εισέλθει στην ασφαλιστική εταιρεία στο έτος  $i$  και έχουν εκπληρωθεί στο έτος  $i+k$ , υποθέτουμε ότι  $((N_{i,i+k}, S_{i,i+k}))_{k=0, 1, \dots}$  είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες επαναδιατυπώσεις των  $((N_k, S_k))_{k=0, 1, \dots}$ . Εδώ  $n$  είναι το τελευταίο έτος για τις αποζημιώσεις το οποίο παρατηρούμε. Οι υποθέσεις του Mack (5.2.2) δίνουν ώθηση στην κατασκευή φυσικών εκτιμήσεων  $\hat{f}_j$  των  $f_j$  που ονομάζονται σαν Chain Ladder εκτιμητές. Τότε, αντικαθιστώντας τις άγνωστες παραμέτρους  $f_j$  στην δεξιά πλευρά του τύπου (5.2.2) με τις εκτιμήσεις τους  $\hat{f}_j$ , τότε λαμβάνουμε την πρόβλεψη του  $N_{j+1}$ . Στην παράγραφο 6 θα εισάγουμε τους Chain Ladder εκτιμητές και θα συγκρίνουμε την απόδοσή τους με τους αντίστοιχους εκτιμητές που προτείνονται σε αυτό το κεφάλαιο.

Η διαδικασία του Mack (5.2.2) δεν ορίζει την δυναμική από μία στοχαστική διαδικασία. Για παράδειγμα, δεν μπορεί κάποιος να αναπαράγει τη διαδικασία  $(N_j)_{j=0, 1, \dots}$  από την (5.2.2). Επιπρόσθετα, η γραμμικότητα του εκτιμητή (5.2.2) είναι μια απλοποίηση στην οποία είναι δύσκολο να συμβιβαστούμε με φυσικά στοχαστικά μοντέλα για την διαδικασία μέτρησης.

Σε αυτό το κεφάλαιο, θεωρούμε ένα απλό στοχαστικό μοντέλο για τις μετρήσεις  $N_j$ , με  $j = 0, 1, \dots$ , και οι αντίστοιχες αποζημιώσεις  $S_j$ , με  $j = 0, 1, \dots$ . Το μοντέλο δίνεται από τις ακόλουθες υποθέσεις, τις οποίες υποθέτουμε στη πορεία του κεφαλαίου.

### **Το μοντέλο**

Έχουμε  $M$  να είναι ο αριθμός των απαιτήσεων που εισέρχονται από ένα δοσμένο έτος με κατανομή

$$q_m = P(M = m), \quad m = 0, 1, \dots$$

Η  $m$ -οστή απαίτηση προκαλεί μία ροή από  $K_m$  αποζημιώσεις από τον ασφαλιστή στον ασφαλισμένο στα επόμενα έτη. Υποθέτουμε ότι η  $k$ -οστή από αυτές τις αποζημιώσεις και εκπληρώνεται στο έτος  $Y_{mk}$ . Υποθέτουμε ακόμα ότι τα  $(K_m)$  είναι ανεξάρτητα και ομοιόμορφα κατανομημένα σε μία αλληλουχία Poisson ( $\mu$ ) κατανομημένες τυχαίες μεταβλητές και ότι  $(Y_{mk})_{m,k=1,2,\dots}$  αποτελούν μια ανεξάρτητη και ομοιόμορφα κατανομημένη οικογένεια με κοινή κατανομή

$$p_j = P(Y_{11} = j), \quad j = 0, 1, \dots \quad (5.2.4)$$

Τέλος, υποθέτουμε ότι τα  $M$ ,  $(K_m)$  και  $(Y_{mk})$  είναι ανεξάρτητα. Οπότε έχουμε

$$N_j = \sum_{m=1}^M \sum_{k=1}^{K_m} I_{\{Y_{mk}=j\}}, \quad j = 0, 1, \dots,$$

όπου  $N_j$  είναι ο αριθμός των αποζημιώσεων για αποζημιώσεις που εισέρχονται στην ασφαλιστική εταιρεία σε ένα δεδομένο έτος και εκπληρώνονται σε ένα έτος  $j$ . Υποθέτουμε περαιτέρω ότι  $(X_{mk})_{m,k=1,2,\dots}$  είναι μία ανεξάρτητη και ομοιόμορφα κατανομημένη οικογένεια από μη αρνητικές τυχαίες μεταβλητές ανεξάρτητες των  $M$ ,  $(K_m)$  και  $(Y_{mk})$ . Ερμηνεύουμε το  $X_{mk}$  σαν την  $k$ -οστή αποζημίωση για την  $m$ -οστή απαίτηση. Τότε

$$S_j = \sum_{m=1}^M \sum_{k=1}^{K_m} X_{mk} I_{\{Y_{mk}=j\}}, \quad j = 0, 1, \dots,$$

είναι οι συνολικές αποζημιώσεις που έρχονται στο έτος 0 και εκπληρώνονται στο έτος  $j$ .

Και οι δύο διαδικασίες  $(N_j)$  και  $(S_j)$  μπορούν εύκολα να προσομοιωθούν. Σκοπός μας είναι να δείξουμε ότι οι εκτιμητές από την (5.2.1) και τα σφάλματα μπορούν να υπολογιστούν ευθέως και απορρέουν από αριθμητικές μεθόδους για ειδικές περιπτώσεις της κατανομής του  $M$ . Οι εκφράσεις για τις προβλέψεις των  $N_{j+1}$  είναι μη γραμμικές συναρτήσεις των  $N_0 + \dots + N_j$  σε αντίθεση με τη διαδικασία του Mack (5.2.2). Ωστόσο, υπό κάποιες προϋποθέσεις στην κατανομή του  $M$  αυτές οι προβλέψεις είναι ασυμπτωτικές γραμμικές συναρτήσεις των  $N_0 + \dots + N_j$  αν η τελευταία ποσότητα αυξάνεται στο άπειρο. Σε άλλες περιπτώσεις, οι προβλέψεις παραμένουν μη γραμμικές ακόμα και στο όριο.

Το κεφάλαιο αυτό είναι οργανωμένο στα εξής. Αρχίζουμε στην παράγραφο 3 δίνοντας της αντίστοιχες συναρτήσεις για τις προβλέψεις των  $S_{j+\ell+1}$ ,  $\ell = 0, 1, \dots$ . Επειδή

$S_j = N_j$  αν  $X_{mi} = 1$  για όλα τα  $m, i$  η πρόβλεψη των  $N_{j+\ell+1}$ ,  $\ell = 0, 1, \dots$ , είναι ιδιαίτερη περίπτωση. Ορίζουμε επίσης τα σφάλματα της πρόβλεψης. Οι προβλέψεις και τα υπό συνθήκη προβλεπόμενα σφάλματα περιλαμβάνουν ορισμένα παράγωγα από τον μετασχηματισμό Laplace-Stieltjes του  $M$ . Γενικά, αυτά τα παράγωγα είναι δύσκολο να παραχθούν. Εν τούτοις, στην (a,b)-τάξη των κατανομών του Panjer, συμπεριλαμβανομένων των παρακάτω κατανομών : Poisson, Διωνυμική, και Αρνητική Διωνυμική, υπάρχουν απλοί αναδρομικοί αλγόριθμοι, για τον υπολογισμό αυτών των παραγώγων, Παράγραφος 4. Στην Παράγραφο 5 αναλύουμε την ασυμπτωτική συμπεριφορά των προβλέψεων ως τον αριθμό από των προηγούμενων παρατηρούμενων αποζημιώσεων που αυξάνονται. Συγκεκριμένα, δίνουμε περιορισμούς υπό τους οποίους οι προβλέψεις  $\hat{S}_{j+\ell+1}$ , είναι ασυμπτωτικές γραμμικές συναρτήσεις των  $N_0 + \dots + N_j$ . Θεωρούμε επίσης και άλλες περιπτώσεις, όπου διαφορετικά ασυμπτωτικά μοτίβα της συμπεριφοράς των προβλέψεων εγείρονται. Ένα ενδιαφέρον χαρακτηριστικό είναι το μοτίβο τύπου αλυσίδας που αναλύουμε σε αυτό το κεφάλαιο. Στην Παράγραφο 6 εφαρμόζουμε τις προβλέψεις σε δεδομένα ασφαλίσεων ζημίας. Συγκρίνουμε την επίδοση των προβλέψεων με τους αντίστοιχους της Chain Ladder μεθόδου.

### 5.3 Το σφάλμα της πρόβλεψης

Θέλουμε να προβλέψουμε μελλοντικούς αριθμούς αποζημιώσεων  $N_{j+\ell+1}$  και τα αντίστοιχα συνολικά ποσά απαιτήσεων  $S_{j+\ell+1}$ ,  $\ell = 0, 1, \dots$ , δεδομένων των προηγούμενων αριθμών αποζημιώσεων  $N_0 + \dots + N_j$ . Αυτό σημαίνει ότι θα υπολογίσουμε τις προβλέψεις  $\hat{S}_{j+\ell+1}$  και  $\hat{N}_{j+\ell+1}$  από την (1) προσπαθώντας να είναι καλά παραγμένες αυτές οι ποσότητες. Εφόσον υποθέτουμε πάντα τις υποθέσεις του μοντέλου όπως στην προηγούμενη παράγραφο, η ανεξαρτησία των  $(X_{mi})$  και των υπολοίπων τυχαίων παραμέτρων του μοντέλου υποδηλώνουν ότι

$$\hat{S}_{j+\ell+1} = EX_{11}E(E(N_{j+\ell+1}|\mathcal{F}_j)) = EX_{11}\hat{N}_{j+\ell+1} \quad (5.3.1)$$

με  $\ell = 0, 1, \dots$ . Άρα το πρόβλημα για τα  $S_{j+\ell+1}$  μειώνεται μόνο για τα  $N_{j+\ell+1}$ . Αντίστροφα, αν τα  $X_{mi} = 1$  για όλα τα  $m, i$  τότε  $N_j = S_j$ ,  $j = 0, 1, \dots$ . Επομένως, αρκεί να μελετήσουμε την πρόβλεψη των  $S_{j+\ell+1}$  δοθέντος  $N_0 + \dots + N_j$ .

Θα παράγουμε εκφράσεις για τις προβλέψεις της (5.3.1) και θα ορίσουμε τα σφάλματά τους. Αρχίζουμε με  $\ell = 0$ .

#### 5.3.1 Πρόβλεψη ενός βήματος μπροστά

Θα εισάγουμε κάποιους συμβολισμούς που θα χρησιμοποιηθούν σε αυτό το κεφάλαιο. Θα χρειαστούμε τον μετασχηματισμό Laplace – Stieltjes για το  $M$ ,

$$L(\gamma) = Ee^{-\gamma M} = \sum_{m=0}^{\infty} q_m e^{-\gamma M}, \quad \gamma \geq 0,$$

και τα παράγωγά του

$$L^{(\ell)}(\gamma) = (1)^\ell E(M^\ell e^{-\gamma M}), \quad \gamma > 0, \quad \ell = 0, 1, \dots,$$

με τη συνθήκη ότι  $L^{(0)} = L$ . Ακόμα, ορίζουμε

$$R_\ell(\gamma) = -\frac{L^{(\ell+1)}(\gamma)}{L^{(\ell)}(\gamma)} = \frac{E(M^{\ell+1} e^{-\gamma M})}{E(M^\ell e^{-\gamma M})} \quad \gamma > 0, \quad \ell = 0, 1, \dots.$$

Τέλος, έχουμε

$$\theta_j = \mu \sum_{d=0}^j p_d \quad j = 0, 1, \dots,$$

Υπενθυμίζουμε ότι το  $\mu$  είναι ο λόγος της Poisson ως τον αριθμό των αποζημιώσεων για κάθε απαίτηση, και  $(p_k)$  είναι οι πιθανότητες στην (5.2.4).

Έπειτα, δημιουργούμε το κυρίως αποτέλεσμα στην πρόβλεψη των  $S_{j+\ell+1}$ .

Υπενθυμίζουμε ότι το αντίστοιχο αποτέλεσμα για τα  $N_{j+\ell+1}$  έρχεται όταν  $X_{mi} = 1$  για όλα τα  $m, i$ .

### Θεώρημα 5.3.1

Υποθέτουμε ότι  $EM < \infty$  και  $EX_{11} < \infty$ .

i) Η πρόβλεψη  $\hat{S}_{j+1}$  των  $S_{j+1}$  δοθέντος  $N_0 + \dots + N_j$  έχει την μορφή

$$E(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) = \mu p_{j+1} EX_{11} R_{n_0 + \dots + n_j}(\theta_j), \quad (5.3.2)$$

ii) Υποθέτουμε επιπλέον ότι  $var(M) < \infty$  και  $var(X) < \infty$ . Τότε το χωρίς συνθήκες σφάλμα της πρόβλεψης για  $S_{j+1}$ ,  $j = 0, 1, \dots$ , δίνεται από την

$$E((S_{j+1} - \hat{S}_{j+1})^2) = E(X_{11}^2) \mu p_{j+1} EM + (EX_{11} \mu p_{j+1})^2 E(M^2) - E(\hat{S}_{j+1}^2). \quad (5.3.3)$$

iii) Υποθέτουμε σε πρόσθεση ότι  $var(M) < \infty$  και  $var(X) < \infty$ . Τότε το υπό συνθήκη προβλεπόμενο σφάλμα για  $S_{j+1}$  δοθέντος τις προηγούμενες παρατηρήσεις  $N_0, \dots, N_j$ ,  $j = 0, 1, \dots$ , είναι

$$\begin{aligned} var(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) &= E(X_{11}^2) \mu p_{j+1} R_{n_0 + \dots + n_j}(\theta_j) + \\ &+ (EX_{11} \mu p_{j+1})^2 R_{n_0 + \dots + n_j}(\theta_j) [R_{n_0 + \dots + n_{j+1}}(\theta_j) - R_{n_0 + \dots + n_j}(\theta_j)]. \end{aligned} \quad (5.3.4)$$

### Παρατήρηση 5.3.1

Έχουμε  $\ell_j = n_0 + \dots + n_j, j = 0, 1, \dots$ , παρατηρούμε σε συνέχεια της (5.3.2) ότι

$$\begin{aligned} E(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) &= \\ &= E(S_{j+1} | N_0 = \ell_0, N_0 + N_1 = \ell_1, \dots, N_0 + \dots + N_j = \ell_j) = \\ &= E(S_{j+1} | N_0 + \dots + N_j = \ell_j), \end{aligned}$$

ή αντίστοιχα

$$\begin{aligned} E(S_0 + \dots + S_{j+1} | N_0 + \dots + N_j = \ell_j) &= \\ &= EX_{11} \ell_j + (S_{j+1} | N_0 + \dots + N_j = \ell_j). \end{aligned} \quad (5.3.5)$$

Λόγω της (5.3.2), η συμβιβασμένη εκτίμηση (5.3.5) είναι γενικά μη γραμμική συνάρτηση της  $\ell_j$ , σύμφωνα με τη διαδικασία του Mack (5.2.2). Στην Παράγραφο 5 θα δοθούν συνθήκες για την κατανομή του  $M$  σιγουρεύοντας ότι η (5.3.5) είναι ασυμπτωματικά γραμμική όταν  $\ell_j \rightarrow \infty$ .

### Παρατήρηση 5.3.2

Στην Παράγραφο 4 θα δοθεί αναδρομικός αλγόριθμος για να αξιολογηθούν οι ποσότητες  $L^{(\ell)}$  όταν η κατανομή του  $M$  ανήκει στην (a,b)-τάξη που χρησιμοποιήθηκε για την αναδρομή του Panjer.

#### Απόδειξη :

Με την διαχώριση της Poisson διαδικασίας,  $(N_j)$  σταθερές, υπό συνθήκη για το  $M$ , μια αλληλουχία από ανεξάρτητες τυχαίες μεταβλητές Poisson. Τότε

$$E(N_{j+1} | N_0, \dots, N_j, M) = M \mu p_{j+1}, \quad j = 0, 1, \dots,$$

και

$$\hat{S}_{j+1} = EX_{11} \mu p_{j+1} E(M | \mathcal{F}_j), \quad j = 0, 1, \dots, \quad (5.3.6)$$

Ακόμα πιο αναλυτικά, έχουμε  $Z_{jl}$  σαν τον αριθμό των αποζημιώσεων στην  $l$ -ιοστή ροή αποζημίωσης,

$l = 1, 2, \dots$ , τα οποία εξοφλούνται στο έτος  $j = 0, 1, \dots$ . Τότε  $(Z_{jl})$  αποτελεί έναν διπλό πίνακα από ανεξάρτητες τυχαίες μεταβλητές με  $Z_{jl} \sim \text{Poisson}(\mu p_j)$ . Τότε για οποιοδήποτε  $m, j = 0, 1, \dots$  και  $n_j = 0, 1, \dots$ ,

$$P(N_0 = n_0, \dots, N_j = n_j, M = m) =$$

$$\begin{aligned}
q_m P\left(\sum_{l=1}^m Z_{0l} = n_0\right) \cdots P\left(\sum_{l=1}^m Z_{jl} = n_j\right) &= \\
q_m \prod_{d=0}^j e^{-m\mu p_d} \frac{(m\mu p_d)^{n_d}}{n_d!} &= \\
= q_m e^{-m\theta_j} m^{\sum_{d=0}^j n_d} \prod_{d=0}^j \frac{(\mu p_d)^{n_d}}{n_d!}. & \quad (5.3.7)
\end{aligned}$$

Καταλήγουμε στο ότι για  $j \geq 0$ ,

$$\begin{aligned}
P(M = m | N_0 = n_0, \dots, N_j = n_j) &= \frac{q_m e^{-m\theta_j} m^{\sum_{d=0}^j n_d}}{\sum_{r=0}^{\infty} q_r e^{-r\theta_j} r^{\sum_{d=0}^j n_d}} \\
m, n_0, \dots, n_j &= 0, 1, \dots. \quad (5.3.8)
\end{aligned}$$

Πιο συγκεκριμένα.

$$\begin{aligned}
E(M = m | N_0 = n_0, \dots, N_j = n_j) &= \frac{\sum_{m=0}^{\infty} q_m e^{-m\theta_j} m^{\sum_{d=0}^j n_d}}{\sum_{r=0}^{\infty} q_r e^{-r\theta_j} r^{\sum_{d=0}^j n_d}} = \\
&= R_{n_0 + \dots + n_j}(\theta_j). \quad (5.3.9)
\end{aligned}$$

Καταλήγουμε, χρησιμοποιώντας την (5.3.6) ότι η (5.3.2) μετατρέπεται.

(i) & (ii) Ξεκινάμε με τον υπολογισμό του σφάλματος της πρόβλεψης των  $S_{j+1}$  δοθέντος των τιμών  $N_0, \dots, N_j$ . Παρατηρούμε πρώτα ότι

$$vars(S_{j+1} | N_0, \dots, N_j, M) = ME(X_{11}^2) \mu p_{j+1}.$$

Λαμβάνοντας υπόψιν αυτήν την σχέση και την (5.3.1), βλέπουμε ότι το υπό συνθήκες προβλεπόμενο σφάλμα μπορεί να γραφτεί ως εξής

$$vars(S_{j+1} | \mathcal{F}_j) = E(X_{11}^2) E(N_{j+1} | \mathcal{F}_j) + (EX_{11} \mu p_{j+1})^2 var(M | \mathcal{F}_j). \quad (5.3.10)$$

Χρησιμοποιώντας την (5.3.8), μπορούμε να αντικαταστήσουμε τις υπό συνθήκες ροπές του  $M$  με τα αντίστοιχα παράγωγα του  $L$ , που οδηγεί (5.3.4). Παίρνοντας τις αναμενόμενες τιμές στην (5.3.10) λαμβάνουμε το σφάλμα της πρόβλεψης

$$E\left((S_{j+1} - \hat{S}_{j+1})^2\right) = E[vars(S_{j+1} | \mathcal{F}_j)] =$$

$$\begin{aligned}
&= E(X_{11}^2)EN_{j+1} + (EX_{11}\mu p_{j+1})^2 \left[ E(M^2) - E \left[ \left( E(M|\mathcal{F}_j) \right)^2 \right] \right] = \\
&= E(X_{11}^2)\mu p_{j+1}EM + (EX_{11}\mu p_{j+1})^2 E(M^2) - E(\hat{S}_{j+1}^2).
\end{aligned}$$

Και ολοκληρώθηκε η απόδειξη. □

### Παρατήρηση 5.3.3

Ένα απλό άνω όριο του υπό συνθήκης σφάλμα πρόβλεψης της (5.3.3) δίνεται από

$$E \left[ (S_{j+1} - \hat{S}_{j+1}^2)^2 \right] \leq E(X_{11}^2)\mu p_{j+1}EM + (EX_{11}\mu p_{j+1})^2 E(M^2).$$

Η εκτίμηση της  $E(\hat{S}_{j+1}^2)$  στην (5.3.3) είναι πολύπλοκη. Ακολουθώντας τα βήματα της προηγούμενης απόδειξης, μπορούμε να παράξουμε έναν πιο άμεσο τρόπο για αυτόν τον όρο.

$$\begin{aligned}
E(\hat{S}_{j+1}^2) &= (EX_{11}\mu p_{j+1})^2 E \left[ \left( E(M|\mathcal{F}_j) \right)^2 \right] = \\
&= (EX_{11}\mu p_{j+1})^2 E \left[ \left( R_{N_0+\dots+N_j}(\theta_j) \right)^2 \right] = \\
&= (EX_{11}\mu p_{j+1})^2 \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} (N_0 + \dots + N_j = k, M = m) \left( R_k(\theta_j) \right)^2.
\end{aligned}$$

Εφαρμόζοντας την (5.3.7) στο δεξιό άθροισμα μετατρέπεται σε

$$\begin{aligned}
&\sum_{k=0}^{\infty} \left[ \sum_{m=0}^{\infty} q_m e^{-m\theta_j} m^k \right] \sum_{n_0+\dots+n_j=k} \prod_{d=0}^j \frac{(\mu p_d)^{n_d}}{n_d!} \left( R_k(\theta_j) \right)^2 = \\
&= e^{\theta_j} \sum_{k=0}^{\infty} (-1)^k \frac{\left( L^{(k+1)}(\theta_j) \right)^2}{L^{(k)}(\theta_j)} P(\theta_j = k) = \\
&= e^{\theta_j} E \left( (-1)_j^{\theta_j} \frac{\left( L^{(\theta_j+1)}(\theta_j) \right)^2}{L^{(\theta_j)}(\theta_j)} \right),
\end{aligned}$$

Όπου  $\theta_j$  ακολουθεί την κατανομή Poisson( $\theta_j$ ).

### 5.3.2 Πρόβλεψη πολλαπλών βημάτων μπροστά



Σε αυτή την ενότητα θεωρούμε το μοντέλο της πρόβλεψης για  $\ell + 1$  περιόδους μπροστά. Αυτό σημαίνει ότι ενδιαφερόμαστε στις ποσότητες  $\hat{S}_{j+\ell+1}$ ,  $\ell = 0, 1, \dots$ , όπως ορίζονται στην (5.2.1) και στα αντίστοιχα σφάλματα πρόβλεψης.

### Θεώρημα 5.3.2

Υποθέτουμε ότι  $EM < \infty$  και  $EX_{11} < \infty$ .

1) Η πρόβλεψη  $\hat{S}_{j+\ell+1}$  του  $S_{j+\ell+1}$ , δοθέντος ότι  $N_0, \dots, N_j$  έχει τη μορφή

$$E(S_{j+\ell+1} | N_0 = n_0, \dots, N_j = n_j) = EX_{11} \mu p_{j+\ell+1} R_{n_0+\dots+n_j}(\theta_j), \quad j, \ell = 0, 1, \dots$$

2) Υποθέτουμε επιπρόσθετα ότι  $var(M) < \infty$  και  $var(X_{11}) < \infty$ . Τότε το χωρίς όρους σφάλμα της πρόβλεψης για  $S_{j+\ell+1}$ ,  $j, \ell = 0, 1, \dots$ , δίνεται από

$$E[(S_{j+\ell+1} - \hat{S}_{j+\ell+1})^2] = E(X_{11}^2) \mu p_{j+\ell+1} EM + (EX_{11} \mu p_{j+\ell+1})^2 E(M^2) - E(\hat{S}_{j+\ell+1}^2).$$

3) Υποθέτουμε ακόμα ότι  $var(M) < \infty$  και  $var(X_{11}) < \infty$ . Τότε το χωρίς όρους σφάλμα της πρόβλεψης για  $S_{j+\ell+1}$ ,  $j, \ell = 0, 1, \dots$ , δίνεται από

$$\begin{aligned} var(S_{j+\ell+1} | N_0 = n_0, \dots, N_j = n_j) &= \\ &= E(X_{11}^2) \mu p_{j+\ell+1} R_{n_0+\dots+n_j}(\theta_j) \\ &\quad + (EX_{11} \mu p_{j+\ell+1})^2 R_{n_0+\dots+n_j}(\theta_j) [R_{n_0+\dots+n_{j+1}}(\theta_j) - R_{n_0+\dots+n_j}(\theta_j)]. \end{aligned}$$

### Απόδειξη

1) Ξεκινάμε παρατηρώντας για τιμές όπου  $\ell \geq 0$

$$\hat{S}_{j+\ell+1} = E[E(S_{j+\ell+1} | \mathcal{F}_{j+\ell}) | \mathcal{F}_j].$$

τότε με την χρήση της (5.3.6), λαμβάνουμε ότι

$$\begin{aligned} \hat{S}_{j+\ell+1} &= EX_{11} \mu p_{j+\ell+1} E[E(M | \mathcal{F}_{j+\ell}) | \mathcal{F}_j] = \\ &= EX_{11} \mu p_{j+\ell+1} E(M | \mathcal{F}_j). \end{aligned}$$

Τώρα θα χρησιμοποιήσουμε την σχέση (5.3.9).

2) και 3) για το υπό συνθήκη σφάλμα της πρόβλεψης για το  $S_{j+\ell+1}$  παρατηρούμε ότι από την (5.3.10) έχουμε

$$var(S_{j+\ell+1} | \mathcal{F}_j) =$$

$$\begin{aligned}
&= E(\text{var}(S_{j+\ell+1}|\mathcal{F}_{j+\ell})|\mathcal{F}_j) + \text{var}(E(S_{j+\ell+1}|\mathcal{F}_{j+\ell})|\mathcal{F}_j) = \\
&= E(X_{11}^2)\widehat{N}_{j+\ell+1} + (EX_{11}\mu p_{j+\ell+1})^2 E(\text{var}(M|\mathcal{F}_{j+\ell})|\mathcal{F}_j) \\
&\quad + (EX_{11}\mu p_{j+\ell+1})^2 \text{var}(E(M|\mathcal{F}_{j+\ell})|\mathcal{F}_j) = \\
&= E(X_{11}^2)\widehat{N}_{j+\ell+1} + (EX_{11}\mu p_{j+\ell+1})^2 \text{var}(M|\mathcal{F}_j).
\end{aligned}$$

Οι υπό συνθήκες ροπές του  $M$  μπορούν να εκφραστούν ακόμα μία φορά σύμφωνα με την (5.3.8). Λαμβάνοντας τις προβλέψεις παίρνουμε το χωρίς όρους σφάλμα της πρόβλεψης

$$\begin{aligned}
&E[(S_{j+\ell+1} - \widehat{S}_{j+\ell+1})^2] = E[\text{var}(S_{j+\ell+1}|\mathcal{F}_j)] = \\
&= E(X_{11}^2)\mu p_{j+\ell+1}EM + (EX_{11}\mu p_{j+\ell+1})^2 \left[ E(M^2) - E\left(\left(E(M|\mathcal{F}_j)\right)^2\right) \right] = \\
&= E(X_{11}^2)\mu p_{j+\ell+1}EM + (EX_{11}\mu p_{j+\ell+1})^2 E(M^2) - E(\widehat{S}_{j+\ell+1}^2). \quad \square
\end{aligned}$$

#### Παρατήρηση 5.3.4

Παρατηρούμε ότι

$$\widehat{S}_{j+\ell+1} = \frac{p_{j+\ell+1}}{p_{j+1}} \widehat{S}_{j+1}, \quad \ell = 0, 1, \dots,$$

με την προϋπόθεση ότι  $p_{j+1} > 0$ . Επίσης, αν το  $p_{j+\ell+1} = 0$  τότε και το  $\widehat{S}_{j+\ell+1} = 0$ .

### 5.3.3 Υποθετικές ανεξάρτητες αποζημιώσεις

Σε αυτήν την υποπαράγραφο θεωρούμε ένα λίγο πιο γενικό μοντέλο. Όπως πριν, υποθέτουμε ότι οι συνέχειες  $(X_{mk})_{k=1,2,\dots}$ ,  $m = 1, 2, \dots$ , είναι ανεξάρτητες και ομοιόμορφα κατανομημένες και ανεξάρτητες από τις υπόλοιπες τυχαίες μεταβλητές που ορίζουν το μοντέλο. Υποθέτουμε περαιτέρω, ότι κάθε συνέχεια  $(X_{mk})_{k=1,2,\dots}$ , αποτελείται από υποθετικές ανεξάρτητες και ομοιόμορφα κατανομημένες τυχαίες μεταβλητές ή, ισοδύναμα, ότι  $(X_{mk})_{k=1,2,\dots}$  είναι ανταλλάξιμη. Η κατάσταση αυτή είναι παρόμοια με τα μοντέλα της θεωρίας αξιοπιστίας, όπου τα μεγέθη των απαιτήσεων που συμβαίνει σε μία ατομική ασφάλιση υποθέτουμε ότι είναι υπό όρους ανεξάρτητα και ομοιόμορφα κατανομημένα, βλ. Mikosch (2004), Κεφάλαια 5 και 6.

Όταν τυχαίες μεταβλητές  $(X_{mk})_{k=1,2,\dots}$  και  $(Y_{mk})_{k=1,2,\dots}$  είναι ανεξάρτητες η μορφή της πρόβλεψης του ενός βήματος μπροστά δίνεται από (5.3.1) αλλά το σφάλμα της πρόβλεψης αλλάζει.

### Πρόταση 5.3.5

Υποθέτουμε ότι  $var(M) < \infty$  και  $var(X_{11}) < \infty$ . Τότε το χωρίς όρους σφάλμα της πρόβλεψης για  $S_{j+\ell+1}$ ,  $j, \ell = 0, 1, \dots$ , δίνεται από

$$\begin{aligned} & E \left( (S_{j+1} - \hat{S}_{j+1})^2 \right) = \\ & = EM \left[ E(X_{11}^2) \mu p_{j+1} + cov(X_{11}, X_{12}) (\mu p_{j+1})^2 \right] + (EX_{11} \mu p_{j+1})^2 E(M^2) - E(\hat{S}_{j+1}^2), \end{aligned}$$

$j = 0, 1, \dots$ .

Το υπό όρους σφάλμα της πρόβλεψης του  $S_{j+1}$  δοθέντος  $N_0, \dots, N_j, j = 0, 1, \dots$ , έχει τη μορφή

$$\begin{aligned} & var(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) = \\ & = \left[ E(X_{11}^2) \mu p_{j+1} + cov(X_{11}, X_{12}) (\mu p_{j+1})^2 \right] R_{n_0 + \dots + n_j}(\theta_j) \\ & \quad + (EX_{11} \mu p_{j+1})^2 R_{n_0 + \dots + n_j}(\theta_j) \left[ R_{n_0 + \dots + n_{j+1}}(\theta_j) - R_{n_0 + \dots + n_j}(\theta_j) \right]. \end{aligned}$$

Μία σύγκριση με το αποτέλεσμα του Θεωρήματος 5.1 δείχνει ότι το σφάλμα της πρόβλεψης αυξάνεται με την εισαγωγή του πολλαπλασιαστή  $cov(X_{11}, X_{12})$ . Είναι μη αρνητική σαν διαδικασία από την υπό όρους ανισότητα Jensen.

### Απόδειξη

Ξεκινάμε υπολογίζοντας

$$\begin{aligned} var(S_{j+1} | N_0, \dots, N_j, M) &= M var \left( \sum_{k=1}^{K_1} X_{1k} I_{\{Y_{1k}=j+1\}} | \mathcal{F}_j \right) = \\ &= ME \left( \left( \sum_{k=1}^{K_1} X_{1k} I_{\{Y_{1k}=j+1\}} \right)^2 | \mathcal{F}_j \right) - M (EX_{11} \mu p_{j+1})^2. \end{aligned}$$

Παρατηρούμε ότι

$$E \left( \left( \sum_{k=1}^{K_1} X_{1k} I_{\{Y_{1k}=j+1\}} \right)^2 | \mathcal{F}_j \right) = E \left( \sum_{k=1}^{K_1} (X_{1k})^2 I_{\{Y_{1k}=j+1\}} | \mathcal{F}_j \right) +$$

$$\begin{aligned}
& + E \left( \sum_{k=1}^{K_1} \sum_{l=1, l \neq k}^{K_1} X_{1k} X_{1l} I_{\{Y_{1k}=j+1\}} I_{\{Y_{1l}=j+1\}} \mid \mathcal{F}_j \right) = \\
& = [E(X_{11}^2) - E(X_{11}X_{12})] \mu p_{j+1} + E(X_{11}X_{12}) E \left( \left( \sum_{k=1}^{K_1} I_{\{Y_{1k}=j+1\}} \right)^2 \mid \mathcal{F}_j \right) = \\
& = E(X_{11}^2) \mu p_{j+1} + E(X_{11}X_{12}) (\mu p_{j+1})^2.
\end{aligned}$$

Εδώ χρησιμοποιούμε το γεγονός ότι, από ανταλλακτικότητα,  $E(X_{11}X_{12}) = E(X_{1k}X_{1l})$  για  $k \neq l$ . Εν τέλη, λαμβάνουμε

$$var(S_{j+1} | N_0, \dots, N_j, M) = M \left[ E(X_{11}^2) \mu p_{j+1} + cov(X_{11}X_{12}) (\mu p_{j+1})^2 \right].$$

Τότε

$$\begin{aligned}
var(S_{j+1} | \mathcal{F}_j) &= E(M | \mathcal{F}_j) \left( E(X_{11}^2) \mu p_{j+1} + cov(X_{11}X_{12}) (\mu p_{j+1})^2 \right) + \\
& \quad + (EX_{11} \mu p_{j+1})^2 var(M | \mathcal{F}_j).
\end{aligned}$$

Για το χωρίς προϋποθέσεις σφάλμα της πρόβλεψης έχουμε

$$\begin{aligned}
E \left[ (S_{j+1} - \hat{S}_{j+1})^2 \right] &= EM \left[ E(X_{11}^2) \mu p_{j+1} + cov(X_{11}, X_{12}) (\mu p_{j+1})^2 \right] + (EX_{11} \mu p_{j+1})^2 E(M^2) - \\
& \quad - (EX_{11} \mu p_{j+1})^2 E \left[ \left( E(M | \mathcal{F}_j) \right)^2 \right] = \\
& = EM \left[ E(X_{11}^2) \mu p_{j+1} + cov(X_{11}, X_{12}) (\mu p_{j+1})^2 \right] + (EX_{11} \mu p_{j+1})^2 E(M^2) - E(\hat{S}_{j+1}^2)
\end{aligned}$$

□

## 5.4 Πρόβλεψη στην (a,b) – τάξη

Στις προηγούμενες ενότητες δείξαμε ότι, για την πρόβλεψη των τιμών  $N_{j+\ell+1}$  και  $S_{j+\ell+1}$ , για  $\ell \geq 0$ , δοθέντος  $N_0, \dots, N_j$ , είναι αναγκαίο να μπορούμε να εκτιμήσουμε τα παράγωγα  $(-1)^l L^{(l)}(\gamma) = E(M^l e^{-\gamma M})$ . Σε αυτήν την παράγραφο υποθέτουμε ότι η κατανομή του  $M$  ανήκει στην (a,b) – τάξη η οποία χρησιμοποιήθηκε στον αναδρομικό αλγόριθμο του Panjer, βλ. Mikosch (2004), Ενότητα 3.3. Η τάξη αυτή δίνεται από την αναδρομική σχέση

$$q_0 > 0, \quad q_m = \left( a + \frac{b}{m} \right) q_{m-1}, \quad m = 1, 2, \dots, \quad a, b \in \mathbb{R}. \quad (5.4.1)$$

Η τάξη αυτή περιέχει ακριβώς τρεις μη εκφυλισμένες κατανομές.

a) Την κατανομή Poisson( $b$ ) με  $a = 0, b > 0$ .

b) Την Διωνυμική κατανομή Bin( $n, p$ ) με  $a < 0, a = -p/(1-p), b = -a(n+1)$  και  $p \in (0,1), n \geq 1$ .

c) Την Αρνητική Διωνυμική κατανομή με παραμέτρους  $(p, v)$  :

$$q_m = \binom{v+m-1}{m} p^v (1-p)^m, \quad m = 0, 1, \dots, \quad p \in (0,1), \quad v > 0.$$

όπου  $0 < a = 1-p, b = (1-p)(v-1)$ .

Θα παράξουμε μία αναδρομή για το παρακάτω  $(-1)^l L^{(l)}(\gamma) = E(M^l e^{-\gamma M})$ .

Χρησιμοποιώντας την (a,b) – δομή και τον τύπο της Διωνυμικής κατανομής, θα έχουμε για  $l \geq 1$ ,

$$\begin{aligned} (-1)^l L^{(l)}(\gamma) &= \sum_{m=1}^{\infty} e^{-\gamma m} m^l \left(a + \frac{b}{m}\right) q_{m-1} = \\ &= a e^{-\gamma} \sum_{m=0}^{\infty} e^{-\gamma m} (m+1)^l q_m + b e^{-\gamma} \sum_{m=0}^{\infty} e^{-\gamma m} (m+1)^{l-1} q_m = \\ &= a e^{-\gamma} \sum_{m=0}^{\infty} e^{-\gamma m} \sum_{r=0}^l \binom{l}{r} m^r q_m + b e^{-\gamma} \sum_{m=0}^{\infty} e^{-\gamma m} \sum_{r=0}^{l-1} \binom{l-1}{r} m^r q_m = \\ &= a e^{-\gamma} (-1)^l L^{(l)}(\gamma) + e^{-\gamma} \sum_{r=0}^{l-1} \left[ a \binom{l}{r} + b \binom{l-1}{r} \right] (-1)^r L^{(r)}(\gamma). \end{aligned}$$

τότε,

$$(-1)^l L^{(l)}(\gamma) = \frac{e^{-\gamma}}{1 - a e^{-\gamma}} \sum_{r=0}^{l-1} \left[ a \binom{l}{r} + b \binom{l-1}{r} \right] (-1)^r L^{(r)}(\gamma). \quad (5.4.2)$$

Παρατηρούμε ότι αυτοί οι τύποι έχουν νόημα γιατί το  $a$  είναι πάντα μικρότερο από το 1. Θεωρούμε τις τρεις διαφορετικές τάξεις των  $(a, b)$  – κατανομών.

#### Πρόταση 5.4.1

Υποθέτουμε ότι η κατανομή του  $M$  ανήκει στην  $(a, b)$  – τάξη.

a) Στην περίπτωση που έχουμε Poisson( $b$ ),  $a = 0, b > 0$ , έχουμε

$$L(\gamma) = e^{-b(1-e^{-\gamma})}, \quad \gamma \geq 0$$

$$(-1)^l L^{(l)}(\gamma) = e^{-\gamma} b \sum_{r=0}^{l-1} \binom{l-1}{r} (-1)^r L^{(r)}(\gamma), \quad l \geq 1.$$

b) Στην περίπτωση της Διωνυμικής κατανομής  $\text{Bin}(n, p)$ ,  $a = -\frac{p}{1-p}$ ,  $b = p(n+1)(1-p)$ ,  $n = 1, 2, \dots$  και  $p \in (0, 1)$  έχουμε για  $l \geq 1$ ,

$$(-1)^l L^{(l)}(\gamma) = \frac{pe^{-\gamma}}{1-p(1-e^{-\gamma})} \sum_{r=0}^{l-1} \binom{l}{r} \left[ (n+1) \frac{l-r}{l} - 1 \right] (-1)^r L^{(r)}(\gamma).$$

c) Στην περίπτωση της Αρνητικής Διωνυμικής,  $a = 1-p$ ,  $b = (1-p)(v-1)$ ,  $p \in (0, 1)$  και  $v > 0$  έχουμε για  $l \geq 1$ ,

$$L(\gamma) = \left( \frac{p}{1-(1-p)e^{-\gamma}} \right)^v, \quad \gamma \geq 0$$

$$(-1)^l L^{(l)}(\gamma) = \frac{(1-p)e^{-\gamma}}{1-(1-p)e^{-\gamma}} \sum_{r=0}^{l-1} \binom{l}{r} \left[ (v-1) \frac{l-r}{l} - 1 \right] (-1)^r L^{(r)}(\gamma).$$

#### Παρατήρηση 5.4.1

Στην περίπτωση της Poisson, κάποιος μπορεί να πάρει μία διαφορετική αναδρομή για το  $L^{(k)}$ . Εισάγουμε το πολυώνυμο  $H_n$  το οποίο είναι  $n$  βαθμών με την αναδρομική σχέση

$$H_0(x) = 0, H_n(x) = -x[H'_{n-1}(x) + H_{n-1}(x)], n \geq 1, x > 0.$$

Ο υπολογισμός τείνει σε

$$L^{(k)}(\gamma) = H_k(be^{-\gamma})L(\gamma), k \geq 0, \gamma > 0.$$

Πιο συγκεκριμένα,  $R_k(\gamma)$  είναι μία λογική συνάρτηση του  $be^{-\gamma}$  για κάθε  $k \geq 0$ :

$$R_k(\gamma) = -\frac{H_{k+1}(be^{-\gamma})}{H_k(be^{-\gamma})} = be^{-\gamma} \left( 1 + \frac{H'_k(be^{-\gamma})}{H_k(be^{-\gamma})} \right).$$

#### Παρατήρηση 5.4.2

Ελεγκτάσεις της  $(a, b)$  – τάξης θεωρούνται στον Hess (2002). Εισάγονται κατανομές  $(q_m)_{m=0,1,\dots}$  που ικανοποιούν την  $(a, b)$  συνθήκη (5.4.1) με  $q_0, \dots, q_k = 0$  για ορισμένα  $k \geq 0$  και  $q_{k+1} > 0$ . Οι υπολογισμοί οδηγούν στην αναδρομική σχέση (5.4.2) για  $L^{(l)}$  σταθερό στην περίπτωση αυτή ομοίως.

### Παρατήρηση 5.4.3

Οι ποσότητες  $(-1)^l L^{(l)}(\gamma)$  αυξάνονται ραγδαία σαν μία συνάρτηση του  $l$  και τότε με σταθερούς υπολογισμούς αποδίδουν την τιμή  $\infty$  ακόμα και για σχετικά μεγάλες τιμές του  $l$ . Αυτό το αριθμητικό πρόβλημα μπορεί να αποφευχθεί γράφοντας την (5.4.2) σε όρους των λόγων  $R_r(\gamma)$  που είναι σχετικοί για την πρόβλεψη όπως θεωρήθηκαν σε προηγούμενες παραγράφους:

$$R_l(\gamma) = \frac{e^{-\gamma}}{1 - \alpha e^{-\gamma}} \sum_{r=0}^l \left[ a \binom{l+1}{r} + b \binom{l}{r} \right] (R_r(\gamma) \cdots R_{l-1}(\gamma))^{-1}.$$

Η τελευταία αναδρομή του  $R_l$  αποφεύγει τον άμεσο υπολογισμό μεγάλων ποσοτήτων της  $[L^{(l)}(\gamma)]$ .

## 5.5 Η ασυμπτωτική συμπεριφορά της πρόβλεψης

### 5.5.1 Η συμπεριφορά των $R_k(\gamma)$ όταν $k \rightarrow \infty$ .

Σε αυτήν την υποπαράγραφο θα μελετήσουμε την ασυμπτωτική συμπεριφορά των προβλέψεων  $E(S_{j+1} | N_0 = n_0, \dots, N_j = n_j), j = 0, 1, \dots$ , όταν ο αριθμός των αποζημιώσεων  $N_0 + \dots + N_j = n_0 + \dots + n_j = k \rightarrow \infty$ . Η ίδια ανάλυση θα γίνει ισοδύναμα για τις πολυσταδιακές προβλέψεις  $E(S_{j+\ell+1} | N_0 = n_0, \dots, N_j = n_j), j, \ell = 0, 1, \dots$ . Παρατηρώντας τα αποτελέσματα του θεωρήματος 5.3.1 μπορούμε να μελετήσουμε την ασυμπτωτική συμπεριφορά των δεικτών  $R_k(\gamma)$  όταν  $k \rightarrow \infty$ .

Το ενδιαφέρον στην ασυμπτωτική συμπεριφορά των  $R_k(\gamma)$  όταν  $k \rightarrow \infty$  γίνεται έντονο, πιο συγκεκριμένα, σε μία σύγκριση με την διαδικασία του Mack (5.2.2). Η τελευταία δηλώνει την πρόβλεψη του  $N_{j+\ell+1}$  δοθέντος  $N_0, \dots, N_j$  να είναι μία γραμμική συνάρτηση του  $k = N_0 + \dots + N_j$ . Στην επιλογές μας, η πρόβλεψη είναι ένα πολλαπλάσιο του  $R_k(\gamma)$  το οποίο δεν έχει λόγο να είναι γραμμικό. Ωστόσο, η παρατήρηση αυτή δεν εξαιρείται από την περίπτωση που το όριο  $k^{-1} R_k(\gamma)$  υπάρχει, είναι πεπερασμένο και θετικό. Σε τέτοιες περιπτώσεις το  $R_k(\gamma)$  γίνεται προσεγγιστικά γραμμικό για μεγάλα  $k$ , όπως στην διαδικασία του Mack.

Το παρακάτω αποτέλεσμα τείνει σε μία επαρκή συνθήκη για ασυμπτωτική γραμμικότητα του  $R_k(\gamma)$ .

#### Λήμμα 5.5.1

Υποθέτουμε ότι  $q_m > 0$  για  $m \geq m_0$  και το όριο

$$\lim_{m \rightarrow \infty} \frac{q_m}{q_{m-1}} = e^{-r} \in (0, 1] \quad (5.5.1)$$

υπάρχει. Τότε

$$\lim_{k \rightarrow \infty} \frac{R_k(\gamma)}{k} = \frac{1}{\gamma+r}. \quad (5.5.2)$$

### Απόδειξη

Έχουμε  $\epsilon \in (0,1)$ . Αναλύουμε την  $(-1)^k L^{(k)}(\gamma)$  για ειδικό  $\gamma$ :

$$\begin{aligned} (-1)^k L^{(k)}(\gamma) &= \left( \sum_{m < \frac{k(1-\epsilon)}{\gamma+r}} + \sum_{m \in \left[ \frac{k(1-\epsilon)}{\gamma+r}, \frac{k(1+\epsilon)}{\gamma+r} \right]} + \sum_{m > \frac{k(1+\epsilon)}{\gamma+r}} \right) m^k q_m e^{-\gamma m} = \\ &= I_1(\epsilon) + I_2(\epsilon) + I_3(\epsilon). \end{aligned}$$

Αρχίζουμε μελετώντας την  $I_3(\epsilon)$ . Για  $m > k(1+\epsilon)/(\gamma+r)$ ,

$$\begin{aligned} \frac{(m+1)^k e^{-(m+1)\gamma}}{m^k e^{-m\gamma}} &= e^{-\gamma} (1+m^{-1})^k \\ &\leq e^{-\gamma} ((1+m^{-1})^m)^{\frac{\gamma+r}{1+\epsilon}} \leq e^{-\gamma + \frac{\gamma+r}{1+\epsilon}}. \end{aligned}$$

Διαλέγουμε  $\delta \in (0, \frac{\epsilon}{1+\epsilon})$ . Για  $k$  αρκετά μεγάλο, υποδηλώνει ότι  $m > \frac{k(1+\epsilon)}{\gamma+r}$  είναι αρκετά μεγάλο, και τώρα έχουμε στην (5.5.1)

$$\frac{q_m}{q_m - 1} \leq e^{-r(1-\delta)}.$$

Με την ένωση αυτών των δύο ορίων λαμβάνουμε

$$\frac{q_{m+1}(m+1)^k e^{-(m+1)\gamma}}{q_m m^k e^{-m\gamma}} \leq e^{-\gamma + \frac{\gamma+r}{1+\epsilon}} e^{-r(1-\delta)} := f(\epsilon, \delta) < 1,$$

όπου χρησιμοποιήσαμε το γεγονός ότι το  $\delta \in (0, \frac{\epsilon}{1+\epsilon})$ . Τότε για μεγάλα  $k$ ,

$$I_3(\epsilon) \leq \frac{1}{1-f(\epsilon, \delta)} q_{\left\lceil \frac{k(1+\epsilon)}{\gamma+r} \right\rceil} \left\lceil \frac{k(1+\epsilon)}{\gamma+r} \right\rceil^k e^{-\left\lceil \frac{k(1+\epsilon)}{\gamma+r} \right\rceil \gamma}. \quad (5.5.3)$$

Ακόμα, για μεγάλα  $k$  ο δείκτης ορίζεται στην  $I_2(\epsilon)$  περιέχει το σημείο  $\left\lceil \frac{k}{\gamma+r} \right\rceil$ . Τότε λαμβάνουμε ένα κάτω όριο, ικανό για μεγάλα  $k$ :

$$I_2(\epsilon) \geq q_{\left\lfloor \frac{k}{\gamma+r} \right\rfloor} \left\lfloor \frac{k}{\gamma+r} \right\rfloor^k e^{-\gamma \left\lfloor \frac{k}{\gamma+r} \right\rfloor}. \quad (5.5.4)$$

Ένας συνδυασμός των (5.5.3) και (5.5.4) αποδίδει



$$\frac{I_3(\epsilon)}{I_2(\epsilon)} \leq \frac{1}{1-f(\epsilon, \delta)} \frac{q_{\lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor}}{q_{\lfloor \frac{k}{\gamma+r} \rfloor}} \left( \frac{\lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor}{\lfloor \frac{k}{\gamma+r} \rfloor} \right)^k e^{-\left[ \lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor - \lfloor \frac{k}{\gamma+r} \rfloor \right] \gamma}.$$

Δυνάμει της (5.5.1), για μικρό  $a \in (0,1)$  και μεγάλα  $k$ ,

$$\frac{q_{\lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor}}{q_{\lfloor \frac{k}{\gamma+r} \rfloor}} \leq e^{-r(1-a)\left(\lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor - \lfloor \frac{k}{\gamma+r} \rfloor\right)} \leq e^{-\left(\frac{r}{\gamma+r}\right)k(1-a)\epsilon}.$$

Επίσης, για μεγάλα  $k$  και κάποιες θετικές σταθερές  $c_1$ ,

$$\begin{aligned} \left( \frac{\lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor}{\lfloor \frac{k}{\gamma+r} \rfloor} \right)^k &\leq \left( \frac{\frac{k(1+\epsilon)}{\gamma+r} + 1}{\frac{k}{\gamma+r} - 1} \right)^k = (1+\epsilon)^k \left( 1 + \frac{1}{\frac{k(1+\epsilon)}{\gamma+r}} \right)^k \left( 1 - \frac{1}{\frac{k}{\gamma+r}} \right)^{-k} \\ &\leq c_1(1+\epsilon)^k. \end{aligned}$$

Τέλος,

$$e^{-\left[ \lfloor \frac{k(1+\epsilon)}{\gamma+r} \rfloor - \lfloor \frac{k}{\gamma+r} \rfloor \right] \gamma} \leq e^{-\frac{\gamma k \epsilon}{\gamma+r}}.$$

Συγκεντρώνοντας τα άνω όρια και επιλέγοντας ένα  $a$  τέτοιο ώστε

$$e^{\frac{\epsilon(r(1-a)+\gamma)}{\gamma+r}} > (1+a)(1+\epsilon),$$

λαμβάνουμε. για κάποιες θετικές σταθερές  $c_2$ ,

$$\frac{I_3(\epsilon)}{I_2(\epsilon)} \leq c_2(1+\epsilon)^k e^{-\frac{k\epsilon(r(1-a)+\gamma)}{\gamma+r}} \leq c_2 \left( \frac{1}{1+a} \right)^k.$$

Άρα η  $I_3(\epsilon) = o(I_2(\epsilon))$  όταν  $k \rightarrow \infty$ .

Έπειτα εργαζόμαστε πάνω στην εκτίμηση της  $I_1(\epsilon)$ . Παρόμοια, έχουμε  $\delta \in (0,1)$  να είναι μικρό και επιλέγουμε  $m_0$  τόσο μεγάλο ώστε

$$\frac{q_{m+1}}{q_m} \geq e^{-r(1+\delta)} \quad \left( 1 + \frac{1}{m} \right)^m \geq e^{1-\delta}, \quad m \geq m_0 \quad (5.5.5)$$

Αναλύουμε περαιτέρω την  $I_1(\epsilon)$ :

$$I_1(\epsilon) = \left( \sum_{m < m_0} + \sum_{m \in [m_0, k \frac{1-\epsilon}{\gamma+r})} \right) m^k q_m e^{-\gamma m} = I_{11}(\epsilon) + I_{12}(\epsilon).$$

Επιπρόλαια,  $I_{11}(\epsilon) = o(I_2(\epsilon))$  όταν  $k \rightarrow \infty$ . Για  $m \in [k, k \frac{1-\epsilon}{\gamma+r})$  έχουμε από την (5.5.5)

$$(1 + m^{-1})^k \geq e^{\left(\frac{k}{m}\right)(1-\delta)} \geq e^{\frac{(\gamma+r)(1-\delta)}{1-\epsilon}},$$

και τότε,

$$\frac{q_{m+1}(m+1)^k e^{-(m+1)\gamma}}{q_m m^k e^{-\gamma m}} \geq e^{-\gamma} e^{-r(1+\delta)} e^{-\frac{(\gamma+r)(1-\delta)}{1-\epsilon}} = g(\epsilon, \delta) > 1,$$

αν διαλέξουμε  $\delta \in (0,1)$  τόσο μικρό ώστε  $\frac{1-\delta}{1-\epsilon} < 1 + \delta$ . Τότε για μεγάλα  $k$ ,

$$I_{12}(\epsilon) \leq \frac{1}{1 - (g(\epsilon, \delta))^{-1}} q_{\lfloor \frac{k(1-\epsilon)}{\gamma+r} \rfloor} \left[ \frac{k(1-\epsilon)}{\gamma+r} \right]^k e^{-\lfloor \frac{k(1-\epsilon)}{\gamma+r} \rfloor \gamma},$$

και με ίδια επιχειρηματολογία στα παραπάνω δηλώνει ότι  $I_{12}(\epsilon) = o(I_2(\epsilon))$  όταν  $k \rightarrow \infty$  για κάθε  $\epsilon \in (0,1)$ .

Συνοψίζουμε ότι για  $\epsilon \in (0,1)$  όταν  $k \rightarrow \infty$

$$R_k(\gamma) \sim \frac{\sum_{m \in \left[ \frac{k(1-\epsilon)}{\gamma+r}, \frac{k(1+\epsilon)}{\gamma+r} \right]} m^{k+1} q_m e^{-\gamma m}}{\sum_{m \in \left[ \frac{k(1-\epsilon)}{\gamma+r}, \frac{k(1+\epsilon)}{\gamma+r} \right]} m^k q_m e^{-\gamma m}} \in \left[ \frac{k(1-\epsilon)}{\gamma+r}, \frac{k(1+\epsilon)}{\gamma+r} \right],$$

άρα, η σχέση (5.5.2) είναι άμεση. Αυτό ολοκληρώνει την απόδειξη.  $\square$

Βλέποντας το Λήμμα 5.5.1 και το Θεώρημα 5.3.1 καταλήγουμε στο ότι, υπό κατάλληλες προϋποθέσεις στην κατανομή του  $M$ , η πρόβλεψη για το  $S_{j+1}$  δοθέντος  $N_0, \dots, N_j$  είναι ασυμπτωματικά γραμμική.

### Πόρισμα 5.5.1

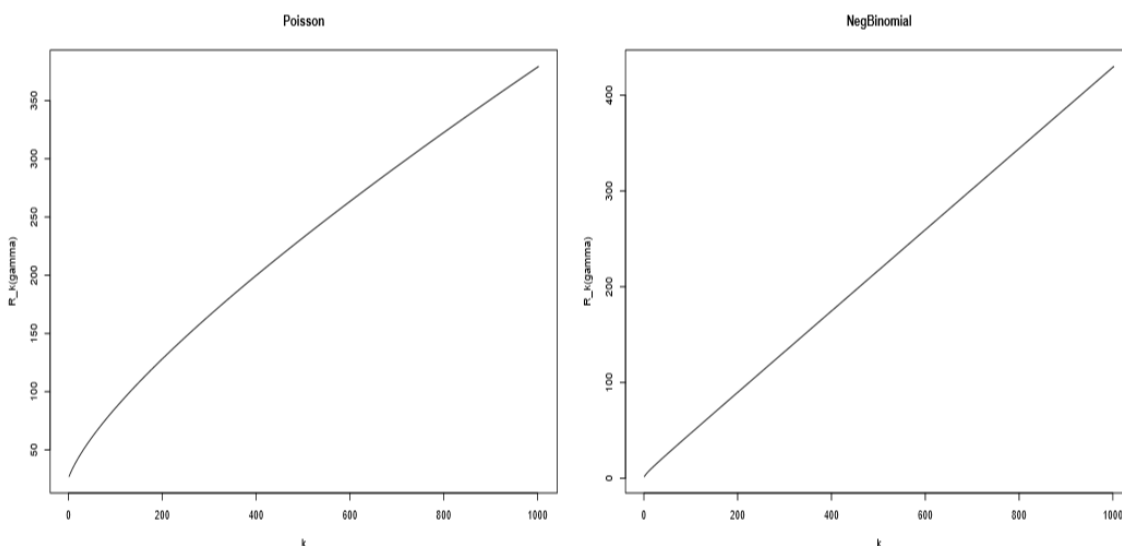
Υποθέτουμε ότι  $EX_{11} < \infty$  και ότι η κατανομή του  $M$  ικανοποιεί την συνθήκη (5.5.1). Τότε όταν  $n_0 + \dots + n_j \rightarrow \infty$ ,

$$E(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) \sim \frac{EX_{11} \mu p_{j+1}}{\tau + \theta_j} (n_0 + \dots + n_j).$$

Στην υπόλοιπη υποενότητα θα μελετήσουμε την συμπεριφορά του  $R_k(\gamma)$  για μεγάλα  $k$  για τις κατανομές της  $(a, b)$  – τάξης που εισάχθηκε στην παράγραφο 4.

### 5.5.1.1 Η Αρνητική Διωνυμική Κατανομή

Η Αρνητική Διωνυμική Κατανομή είναι το μόνο μέλος της  $(a, b)$  – τάξης που ικανοποιεί την συνθήκη (5.5.1) με  $e^{-r} = 1 - p$ . Άρα το Πόρισμα 5.5.1 ισχύει. Η ασυμπτωματική γραμμική συμπεριφορά του  $R_k(\gamma)$  είναι διατυπωμένη στο δεξιό γράφημα του Σχήματος 5.5.1.



Σχήμα 5.5.1. Ο δείκτης  $R_k(2)$  για  $Poisson(200)$  της κατανομής του  $M$  (αριστερά) και για την Αρνητική Διωνυμική Κατανομή του  $M$  με παραμέτρους  $p = 0.1, v = 12.1$  (δεξιά).

### 5.5.1.2 Η Διωνυμική Κατανομή

Σε αυτήν την περίπτωση, είναι φανερό πως, όταν  $k \rightarrow \infty$ ,

$$R_k(\gamma) = \frac{\sum_{m=1}^n \left(\frac{m}{n}\right)^k m e^{-\gamma m} q_m}{\sum_{m=1}^n \left(\frac{m}{n}\right)^k e^{-\gamma m} q_m} \rightarrow n.$$

Το ίδιο αποτέλεσμα ισχύει για κάθε κατανομή  $(q_m)_{m=0,1,\dots,n}$ ,  $n \geq 1$ , με  $q_n > 0$ .

Σε αυτήν την περίπτωση, αν  $EX_{11} < \infty$  τότε  $n_0 + \dots + n_j \rightarrow \infty$ ,

$$E(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) \rightarrow n EX_{11} \mu p_{j+1}.$$

### 5.5.1.3 Η Poisson Κατανομή

Αν  $M$  είναι κατανομημένο σε μία  $\text{Poisson}(b)$ , τότε

$$E(M^k e^{-\gamma M}) = e^{-b(1-e^{-\gamma})} \sum_{m=1}^{\infty} m^k \left(\frac{b}{e^{\gamma}}\right)^m \frac{1}{m!} e^{-be^{-\gamma}}, \quad k \geq 1.$$

Άρα, ο δείκτης  $R_k(\gamma)$  είναι ίσος με τον λόγο των ροπών  $\frac{E(M')^{k+1}}{E(M')^k}$  για κάποια  $\text{Poisson}$  τυχαία μεταβλητή  $M'$  με διαφορετικό μέσο, ας πούμε,  $\lambda$ . Στην συνέχεια μελετούμε, την ασυμπτωματική συμπεριφορά, όταν  $k \rightarrow \infty$ , από κάποιους λόγους. Για ευκολία, θα χρησιμοποιήσουμε την χρήση των συμβόλων  $\frac{E(M)^{k+1}}{E(M)^k}$  αντί για τα κατάλληλα  $\frac{E(M')^{k+1}}{E(M')^k}$ .

Η απόδειξη του παρακάτω Λήμματος είναι παρόμοια με την απόδειξη στο Λήμμα 5.5.1.

Παραθέτουμε το επιχείρημα.

### Λήμμα 5.5.2

Έχουμε  $M$  να είναι κατανομημένο σε μία  $\text{Poisson}(\lambda)$ . Τότε, όταν  $k \rightarrow \infty$ ,

$$\frac{E(M)^{k+1}}{E(M)^k} \sim \frac{k}{\log k}. \quad (5.5.6)$$

### Απόδειξη

Για  $\epsilon \in (0,1)$  αναλύουμε το  $E(M)^k$  ως

$$\begin{aligned} E(M)^k &= \left( \sum_{m < \frac{k(1-\epsilon)}{\log k}} + \sum_{m \in \left[ \frac{k(1-\epsilon)}{\log k}, \frac{k(1+\epsilon)}{\log k} \right]} + \sum_{m > \frac{k(1+\epsilon)}{\log k}} \right) m^k e^{-\lambda} \frac{\lambda^m}{m!} = \\ &= I_1(\epsilon) + I_2(\epsilon) + I_3(\epsilon). \end{aligned}$$

Ξεκινάμε με το  $I_3(\epsilon)$ , είναι επακόλουθο, με την χρήση του τύπου του Stirling ότι

$$\sup_{m > \frac{k(1+\epsilon)}{\log k}} \frac{(m+1)^k \lambda^{m+1}}{(m+1)!} \frac{1}{m^k \frac{\lambda^m}{m!}} \leq \frac{\lambda \log k}{1+\epsilon} k^{-\frac{\epsilon}{1+\epsilon}} \rightarrow 0$$

όταν  $k \rightarrow \infty$ . Άρα, όταν  $k \rightarrow \infty$ ,

$$I_3(\epsilon) = (1 + o(1)) (m_*^{+, \epsilon})^k e^{-\lambda} \frac{\lambda^{m_*^{+, \epsilon}}}{(m_*^{+, \epsilon})!}$$

όπου

$$m_* = \lfloor k / \log(k) \rfloor, m_*^{+, \epsilon} = \lceil (1 + \epsilon)m_* \rceil.$$

Είναι φανερό ότι για μεγάλα  $k$ ,

$$I_3(\epsilon) \geq m_*^k e^{-\lambda} \frac{\lambda^{m_*}}{m_*!}.$$

Τότε, όταν  $k \rightarrow \infty$ ,

$$\frac{I_3(\epsilon)}{I_2(\epsilon)} = e^{k(\log(1+\epsilon) - \epsilon + o(1))}.$$

Επειδή,  $\log(1 + \epsilon) < \epsilon$ , συμπεραίνουμε ότι για κάθε  $\epsilon > 0$ ,

$$\frac{I_3(\epsilon)}{I_2(\epsilon)} \rightarrow 0 \text{ όταν } k \rightarrow \infty.$$

Όμοια, για όταν  $k \rightarrow \infty$ ,

$$I_1(\epsilon) = (1 + o(1))(m_*^{-, \epsilon})^k e^{-\lambda} \frac{\lambda^{m_*^{-, \epsilon}}}{(m_*^{-, \epsilon})!},$$

όπου

$$m_*^{-, \epsilon} = \lfloor (1 - \epsilon)m_* \rfloor,$$

για το οποίο είναι εύκολο να ελέγξουμε ότι, όταν  $k \rightarrow \infty$ ,

$$\frac{I_1(\epsilon)}{I_2(\epsilon)} = e^{k(\log(1-\epsilon) - \epsilon + o(1))}.$$

Αφού,  $-\log(1 - \epsilon) > \epsilon$  για  $0 < \epsilon < 1$ , καταλήγουμε ότι για κάθε  $\epsilon > 0$ ,

$$\frac{I_1(\epsilon)}{I_2(\epsilon)} \rightarrow 0 \text{ όταν } k \rightarrow \infty.$$

Έτσι για κάθε  $\epsilon > 0$ ,

$$EM^k = (1 + o(1)) \sum_{m \in \left[ \frac{k(1-\epsilon)}{\log k}, \frac{k(1+\epsilon)}{\log k} \right]} m^k e^{-\lambda} \frac{\lambda^m}{m!},$$

όταν  $k \rightarrow \infty$ , και άρα,

$$\frac{E(M)^{k+1}}{E(M)^k} = (1 + o(1)) \frac{\sum_{m \in \left[\frac{k(1-\epsilon)}{\log k}, \frac{k(1+\epsilon)}{\log k}\right]} m^{k+1} e^{-\lambda} \frac{\lambda^m}{m!}}{\sum_{m \in \left[\frac{k(1-\epsilon)}{\log k}, \frac{k(1+\epsilon)}{\log k}\right]} m^k e^{-\lambda} \frac{\lambda^m}{m!}}$$

$$\in \left[ (1 + o(1)) \frac{k(1-\epsilon)}{\log k}, (1 + o(1)) \frac{k(1+\epsilon)}{\log k} \right].$$

Άρα και η δήλωση του Λήμματος. □

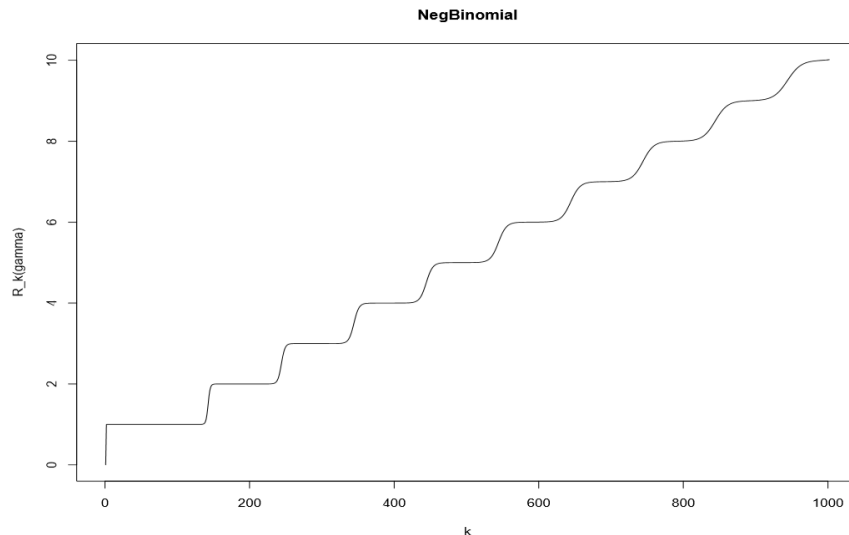
Παίρνοντας υπόψιν το Λήμμα 5.5.2 και τη παρατήρηση πριν από αυτό, καταλήγουμε στο ότι, αν  $M$  ακολουθεί μία κατανομή Poisson( $b$ ), τότε, όταν  $n_0 + \dots + n_j \rightarrow \infty$ ,

$$E(S_{j+1} | N_0 = n_0, \dots, N_j = n_j) \sim EX_{11} \mu p_{j+1} \frac{n_0 + \dots + n_j}{\log(n_0 + \dots + n_j)}.$$

Η ασυμπτωτική συμπεριφορά του  $R_k(\gamma)$  ορισμένη από την (5.5.6) έχει αναπαρασταθεί στο αριστερό γράφημα του Σχήματος 5.5.1.

### 5.5.2 Η συμπεριφορά των $R_k(\gamma)$ για μεγάλα $\gamma$

Όταν αναλύουμε τον δείκτη  $R_k(\gamma)$  αριθμητικά, παρατηρούμε ένα ασυνήθιστο φαινόμενο για μεγάλες τιμές του  $\gamma$ : το  $R_k(\gamma)$  κυμαίνεται σχετικά πολύ για σχετικά μεγάλες τιμές του  $k$ , όπου η επίδραση αυτή σταδιακά εξαφανίζεται όταν το  $k$  γίνεται ακόμα πιο μεγάλο. Για μικρές τιμές του  $\gamma$  η συμπεριφορά αυτή δεν μπορεί να παρατηρηθεί. Ένα υπολογιστικό γράφημα του  $R_k(\gamma)$  το οποίο επιδεικνύει την σταδιακή συμπεριφορά απεικονίζεται στο Σχήμα 5.5.2. Παρακάτω δίνουμε ένα οριακό θεώρημα που εξηγεί αυτό το φαινόμενο.



Σχήμα 5.5.2 Ο δείκτης  $R_k(100)$  για την Αρνητική Διωνυμική κατανομή του  $M$  με παραμέτρους  $p = 0.1$  και  $v = 12.1$ .

Υποθέτουμε ότι

$$\gamma \rightarrow \infty, \quad k = k(\gamma) \text{ και } \frac{k}{\gamma} \rightarrow t \in (0, \infty). \quad (5.5.7)$$

Για  $t > 0$  θεωρούμε την αυστηρή συνάρτηση κυρτότητας

$$h_t(s) = s - t \log s, \quad s > 0,$$

το οποίο φτάνει στο ελάχιστο του όταν  $s = t$ . Τότε η αλληλουχία  $(h_t(m))_{m=1,2,\dots}$  φτάνει στο ελάχιστο είτε όταν  $m = m_-(t) = \lfloor t \rfloor$  αν  $m_-(t) \geq 1$  είτε όταν  $m = m_+(t) = \lceil t \rceil$  και είναι πιθανό ότι οι τιμές της  $h_t$  στα  $m_-(t)$  και  $m_+(t)$  συμπίπτουν. Υποθέτουμε ακόμα ότι υπάρχει μία μοναδική τιμή  $m = m_*(t)$  στην οποία ελαχιστοποιείται η  $h_t$ . Έχουμε για  $k \geq 1$ ,

$$E(M^k e^{-\gamma M}) = q_{m_*(t)} (m_*(t))^k e^{-\gamma m_*(t)} + \sum_{m \neq m_*(t)} m^k e^{-\gamma m} q_m.$$

Δυνάμει της (5.5.7),

$$q_{m_*(t)} (m_*(t))^k e^{-\gamma m_*(t)} = q_{m_*(t)} e^{-\frac{\gamma h_k(m_*(t))}{\gamma}} \rightarrow q_{m_*(t)} e^{-\gamma h_t(m_*(t))}. \quad (5.5.8)$$

Η υπόθεση ότι το ελάχιστο της  $h_t(m)$  είναι μοναδικό για ένα δεδομένο  $t$  δηλώνει ότι το ελάχιστο της  $h_u(m)$  είναι επίσης μοναδικό για  $u$  σε ένα συγγενικό του  $t$  και επιτυγχάνεται στο  $m_*(t)$ . Άρα, για  $u$  αρκετά κοντά στο  $t$ ,

$$\begin{aligned} \min_{m \neq m_*(t)} h_u(m) &= \min(h_u(m_*(t) - 1), h_u(m_*(t) + 1)) \rightarrow \\ &\rightarrow \min(h_t(m_* - 1), h_t(m_* + 1)), u \rightarrow t \end{aligned} \quad (5.5.9)$$

Τότε για  $\gamma$  αρκετά μεγάλο,

$$\begin{aligned} \sum_{m \neq m_*(t)} m^k e^{-\gamma m} q_m &= \sum_{m \neq m_*(t)} e^{-\frac{\gamma h_k(m)}{\gamma}} q_m \leq \\ &\leq e^{-\gamma \min\left(\frac{h_k(m_*(t)-1)}{\gamma}, \frac{h_k(m_*(t)+1)}{\gamma}\right)}. \end{aligned} \quad (5.5.10)$$

Ακολουθεί από (5.5.8)-(5.5.10) ότι

$$E(M^k e^{-\gamma M}) \sim q_{m_*(t)} (m_*(t))^k e^{-\gamma m_*(t)},$$

υπό την προϋπόθεση ότι  $q_{m_*(t)} \neq 0$ , και, τότε, υπό την υπόθεση (5.5.7),

$$R_k(\gamma) \rightarrow m_*(t). \quad (5.5.11)$$

Για  $j = 1, 2, \dots$ , έχουμε

$$a_j = \inf\{t \geq 0: m_*(t) = j\} = \frac{1}{\log j - \log(j-1)}.$$

Τότε  $a_j < a_{j+1}$  και από την σχέση (4.11) λαμβάνουμε το ακόλουθο αποτέλεσμα.

### Πρόταση 5.5.1

Για  $j = 1, 2, \dots$ , έχουμε  $t \in (a_j, a_{j+1})$  να είναι τέτοιο ώστε  $q_{m_*(t)} \neq 0$ . Αν  $k$  και  $\gamma$  αυξάνονται ανάλογα με την (5.5.7), τότε

$$R_k(\gamma) \rightarrow j.$$

Αυτό είναι, αν  $k$  και  $\gamma$  είναι μεγάλα, και  $k/\lambda \in (a_j, a_{j+1})$  για  $j$  όχι πολύ μεγάλο, τότε  $R_k(\gamma)$  θα είναι κοντά στο  $j$ . Πάλι η σταδιακή συμπεριφορά είναι φανερή στο Σχήμα 5.5.2.

## 5.6 Πρόβλεψη σε ένα σύνολο δεδομένων από ασφαλίσεις ζημιών

Σε αυτήν την υποπαράγραφο θεωρούμε ένα σύνολο δεδομένων από ασφαλίσεις ζημιών το οποίο ήταν παροχή του Alois Gisler. Η πολιτική της εταιρείας είναι άγνωστη σε εμάς. Στόχος μας είναι να μελετήσουμε την επίδοση των προβλέψεών μας σε αυτό το σύνολο δεδομένων, δεδομένων επιθυμητών υποθέσεων στις κατανομές των  $(q_m)$  και  $(p_j)$  και στον παράμετρο της Poisson κατανομής  $\mu$ . Επιπρόσθετα, θα συγκρίνουμε τις προβλέψεις μας με αυτές που έχουν δοθεί από την μέθοδο Chain Ladder υπό τις προϋποθέσεις του Mack, παράγραφος 2. Θα εστιάσουμε στις προβλέψεις των αριθμών των αποζημιώσεων.

Τα δεδομένα που χρησιμοποιούμε περιέχουν απαιτήσεις που καταφθάνουν σε ένα έτος (1985) και οι ατομικές διαδικασίες αποζημιώσεων για κάθε απαίτηση, συμπεριλαμβανομένων ημερομηνία άφιξης στο 1985 και όλες τις ημερομηνίες και τα ποσά των εξοφλημένων αποζημιώσεων. Συνολικά, 7,302 απαιτήσεις έχουν συμβεί οι οποίες δημιούργησαν 24,606 αποζημιώσεις σε περισσότερα από 10 έτη. Αφού θέλουμε να συγκρίνουμε την μέθοδό μας με την διαδικασία της Chain Ladder πρόβλεψης του Mack, ένα από τα προβλήματα που αντιμετωπίζουμε είναι το ακόλουθο. Η μέθοδος Chain Ladder χρειάζεται ένα τρίγωνο εξέλιξης με δεδομένα από διαφορετικά έτη, βλ. (5.2.3). Τα δεδομένα αυτά χρειάζονται για την κατασκευή των εκτιμητών της Chain Ladder για τους παράγοντες  $f_j$  στην (5.2.2) :



$$\hat{f}_j = \frac{\sum_{i=1}^{n-j-1} \sum_{r=0}^{j+1} N_{i,i+r}}{\sum_{i=1}^{n-j-1} \sum_{r=0}^j N_{i,i+r}}, \quad (5.6.1)$$

όπου  $n$  να είναι ο αριθμός των ετών για τα οποία το τρίγωνο εξέλιξης είναι διαθέσιμο. Αλλά έχουμε μόνο ένα έτος απαιτήσεων που έχουν εισέλθει στην διάθεση μας.

Λύνουμε το πρόβλημα αυτό με την αλλαγή των ετών σε μήνες. Τότε έχουμε  $n = 12$  μήνες απαιτήσεων που έχουν εισέλθει και οι αντίστοιχες ατομικές αποζημιώσεις υποθέτουμε πως είναι υπολογισμένες σε μήνες. Ο Πίνακας 5.1 περιέχει την μηνιαία πληροφορία για αριθμούς απαιτήσεων που έχουν εισέλθει  $M_j, j = 1, \dots, 12$ , δείχνοντας μια καθαρή εποχικότητα στα δεδομένα.

$j$	1	2	3	4	5	6	7	8	9	10	11	12
$M_j$	414	391	493	552	641	765	748	762	741	635	628	532
$100 \times p_j$	2.41	11.36	10.37	7.95	6.49	5.26	4.30	4.04	3.00	2.94	2.60	2.09

Πίνακας 5.1 Οι μηνιαίοι αριθμοί απαιτήσεων  $M_j$  και οι εκτιμώμενες πιθανότητες  $p_j, j = 1, \dots, 12$ .

Ο Πίνακας 5.2 περιέχει τους αριθμούς αποζημιώσεων σε μία μορφή ενός τριγώνου εξέλιξης. Η  $i$ -οστή σειρά περιέχει τους αριθμούς αποζημιώσεων  $N_{ii}, \dots, N_{i,12}$  για απαιτήσεις που έρχονται στον  $i$ -οστό μήνα και του οποίου οι αποζημιώσεις εξοφλούνται στον μήνα  $i+k \in \{1, \dots, 12\}$ . Τα δεδομένα αυτά υποτίθεται ότι είναι γνωστά (παρατηρούμενες τιμές). Για το σύνολο των δεδομένων μας, γνωρίζουμε επίσης τους μελλοντικούς αριθμούς αποζημιώσεων  $N_{i,12+k}, k \geq 1$ , των οποίων θέλουμε να προβλέψουμε. Αυτοί παρουσιάζονται στον Πίνακα 5.3.

Στο μοντέλο μας, υποθέτουμε πως οι μηνιαίοι αριθμοί απαιτήσεων  $M_j, j = 1, \dots, 12$ , ακολουθούν την Poisson κατανομή. Αφού παρατηρείται μια εποχικότητα στα δεδομένα δεν υποθέτουμε ότι τα  $M_j$  είναι ομοιόμορφα κατανομημένα και απλώς παίρνουμε τα  $M_j$  αυτούσια ως υποκατάστατα των εκτιμητών τους  $EM_j$ . Η υπόθεση της κατανομής Poisson στα  $M_j$  είναι επί τούτου, αλλά έχουμε μόνο ένα σύνολο δεδομένων, το οποίο το κάνει πολύ δύσκολο να εκτιμήσουμε τις κατανομές του  $M_j$ . Γενικά, τέτοια εκτίμηση έχει να κάνει με ιστορικά δεδομένα. Εκτιμάμε μαζί τον μέσο Poisson αριθμό των αποζημιώσεων ανά απαίτηση,  $\mu = 3.37$ , και η κατανομή ( $p_j$ ) από την εμπειρική κατανομή των δεδομένων. Οι παρατηρούμενες τιμές  $M_j$  και οι εκτιμώμενες πιθανότητες  $p_j, j = 1, \dots, 12$ , δίνονται στον Πίνακα 5.1.

Μήνες	1	2	3	4	5	6	7	8	9	10	11	12
1	40	147	141	120	99	84	67	63	42	44	47	35
2		40	144	132	100	85	71	53	63	52	50	25
3			39	179	159	125	118	80	85	85	85	48
4				36	209	169	158	117	91	103	98	54
5					60	217	224	159	136	125	110	86
6						75	280	229	186	191	145	121
7							70	234	192	184	157	131
8								63	242	250	197	142
9									62	246	227	195
10										64	200	175
11											59	197
12												28

Πίνακας 5.2. Το τρίγωνο εξέλιξης με τον αριθμό των αποζημιώσεων από 12 μήνες των εισερχομένων απαιτήσεων. Η  $i$ -οστή σειρά περιέχει τους παρατηρούμενους αριθμούς αποζημιώσεων  $N_{ii}, \dots, N_{i,12}$ .

Μήνες	13	14	15	16	17	18	19	20	21	22	23
1											
2	33										
3	52	42									
4	51	45	32								
5	68	53	56	42							
6	107	79	68	75	55						
7	120	83	66	98	68	54					
8	150	105	112	107	88	75	59				
9	148	140	115	131	107	85	80	56			
10	174	118	122	103	82	77	76	57	45		
11	235	167	143	134	116	89	65	66	50	61	
12	177	161	138	139	120	101	113	63	65	86	56

Πίνακας 5.3. Οι 'μελλοντικοί' αριθμοί αποζημιώσεων αντίστοιχοι στις παρατηρήσεις του Πίνακα 5.2. Η  $i$ -οστή σειρά περιέχει τους μηνιαίους αριθμούς αποζημιώσεων  $N_{i,13}, \dots, N_{i,12+i-1}$ . Οι αριθμοί αυτοί πρέπει να προβλεφθούν.

Στον Πίνακα 5.4 παρουσιάζουμε τα αποτελέσματα από την διαδικασία της πρόβλεψης. Αφού γνωρίζουμε τις τιμές που θέλουμε να προβλέψουμε (Πίνακας 5.3) υπολογίζουμε το σχετικό σφάλμα της πρόβλεψης. Η  $i$ -οστή σειρά στον Πίνακα 5.4 δείχνει το σχετικό σφάλμα στην διαδικασία της πρόβλεψης μας για τις τιμές  $N_{i,12+k}$ ,  $k = 1, \dots, i - 1$ , βασισμένοι στις τιμές  $N_{ii}, \dots, N_{i,12}$ . Η πρόβλεψη του τριγώνου έχει επιλεγθεί

επειδή είναι έτσι η διαδικασία στον υπολογισμό της μεθόδου Chain Ladder. Τα σχετικά σφάλματα κυμαίνονται ανάμεσα σε αρνητικές και θετικές κατευθύνσεις. Οι λόγοι για αυτές τις αποκλίσεις είναι τα προβλήματα στην επιλογή των κατάλληλων κατανομών για τα  $(q_m)$ ,  $(p_j)$  αλλά επίσης την στατιστική αβεβαιότητα όταν βαθμονομούμε το μοντέλο από ένα μοναδιαίο σύνολο δεδομένων. Ακόμα, δεν μπορούμε να αναμένουμε θαύματα από τις μέσες τετραγωνικές προβλέψεις με τις παρούσες κατανομές οι οποίες είναι αρκετά διαφορετικές από την Γκαουσιανή.

Μήνες	13	14	15	16	17	18	19	20	21	22	23
1											
2	-11.6										
3	-0.8	0.7									
4	-29.8	-33.0	-48.9								
5	-32.9	-56.1	-33.6	-41.9							
6	-2.4	-7.9	-14.8	5.9	-2.2						
7	-1.8	-36.5	-33.5	17.7	-7.3	-7.6					
8	6.0	-15.6	-0.5	18.2	8.9	3.3	2.2				
9	0.4	7.9	3.5	21.5	25.2	13.8	17.2	5.8			
10	14.1	-6.1	10.3	8.5	-6.5	11.7	18.1	1.3	0.5		
11	17.1	3.2	5.4	11.7	12.2	-6.1	-13.0	-1.9	-21.6	20.6	
12	-4.0	-6.1	-2.7	14.7	13.6	11.6	26.7	-2.2	9.3	38.0	24.2

Πίνακας 5.4. Σχετικό σφάλμα πρόβλεψης σε ποσοστά για την μέθοδό μας. Οι τιμές έχουν υπολογιστεί βάσει τις παρατηρούμενες τιμές στους πίνακες 5.2 και 5.3.

Μήνες	13	14	15	16	17	18	19	20	21	22	23
1											
2	3.4										
3	17.9	2.5									
4	-13.3	-6.5	-28.4								
5	-13.8	-22.0	0.7	-18.7							
6	-11.2	-17.0	-17.1	11.8	-11.7						
7	8.3	-21.7	-21.1	35.9	15.3	-1.4					
8	14.3	-10.5	-0.2	20.8	15.2	20.1	1.7				
9	-6.9	7.3	-1.4	17.4	21.5	11.9	28.8	-2.9			
10	14.3	-8.3	15.4	9.0	-9.3	7.9	23.6	13.3	-3.7		
11	28.0	9.5	11.0	26.6	22.6	-1.6	-9.0	7.2	-0.7	30.4	
12	67.5	68.0	73.3	106.5	117.1	104.4	139.1	68.9	10.21	227.0	129.2

Πίνακας 5.5. Σχετικό σφάλμα πρόβλεψης σε ποσοστά για την μέθοδο Chain Ladder. Οι τιμές έχουν υπολογιστεί βάσει τις παρατηρούμενες τιμές στους πίνακες 5.2 και 5.3.

Στην αριστερή στήλη του Σχήματος 5.3 δείχνουμε την επίδοση της διαδικασίας της πρόβλεψης για δύο διαφορετικές περιπτώσεις. Στο πάνω γράφημα είμαστε σε θέση της 11<sup>ης</sup> γραμμής των Πινάκων 5.2 και 5.3. Προβλέπουμε τους αριθμούς των αποζημιώσεων

για απαιτήσεις που εισέρχονται τον 11<sup>ο</sup> μήνα. Δύο αριθμοί αποζημιώσεων έχουν ήδη παρατηρηθεί από τους οποίους οι αριθμοί αποζημιώσεων για τους μήνες 13-22 έχουν προβλεφθεί (μονόχρωμες γραμμές). Οι κουκκίδες δηλώνουν τους παρατηρούμενους αριθμούς αποζημιώσεων και οι ζώνες γύρω από τις προβλέψεις αντιπροσωπεύουν το  $\pm 1.96$  φορές το τετράγωνο του υπό προϋποθέσεις σφάλματος πρόβλεψης που είχε δοθεί στο Θεώρημα 2.2. Στο κάτω γράφημα είμαστε σε θέση της 7<sup>ης</sup> γραμμής των Πινάκων 5.2 και 5.3. Προβλέπουμε τους αριθμούς των αποζημιώσεων απαιτήσεων για τον 7<sup>ο</sup> μήνα. Έξι αποζημιώσεις έχουν ήδη παρατηρηθεί εκ των οποίων οι αριθμοί αποζημιώσεων για τους μήνες 13-18 έχουν προβλεφθεί.

Οι προβλέψεις μας συγκρίνονται με εκείνους από την μέθοδο Chain Ladder του Mack (5.2.2). Χρησιμοποιώντας τους προηγούμενους συμβολισμούς και τους Chain Ladder εκτιμητές που έχουν δοθεί στην (5.6.1), οι Chain Ladder προβλέψεις  $\tilde{N}_{i,12+k}$  του  $N_{i,12+k}$  δίνονται από

$$\tilde{N}_{i,12+k} = (\hat{f}_{12-i+k-1} - 1)\hat{f}_{12-i+k-2} \cdots \hat{f}_{12-i}(N_{ii}, \dots, N_{i,12}), \quad k = 1, 2, \dots.$$

Για να δημιουργήσουμε αυτές τις προβλέψεις προτείνεται να γίνει πολλαπλή χρήση της υπόθεσης του Mack (5.2.2) :

$$\begin{aligned} E(N_{i,12+k} | N_{ii}, \dots, N_{i,12}) &= \\ &= E(E(N_{i,12+k} | N_{ii}, \dots, N_{i,12+k-1}) | N_{ii}, \dots, N_{i,12}) = \\ &= (f_{12-i+k-1} - 1)E(N_{ii} + \dots + N_{i,12+k-1} | N_{ii}, \dots, N_{i,12}) = \\ &= (f_{12-i+k-1} - 1)f_{12-i+k-2}E(N_{ii} + \dots + N_{i,12+k-2} | N_{ii}, \dots, N_{i,12}) = \\ &= (f_{12-i+k-1} - 1)f_{12-i+k-2} \cdots f_{12-i}(N_{ii}, \dots, N_{i,12}). \end{aligned}$$

Η ποσότητα  $\tilde{N}_{i,12+k}$  λαμβάνεται όταν αντικαταστήσουμε τους παράγοντες  $f$  με τους  $\hat{f}$  εκτιμητές τους.

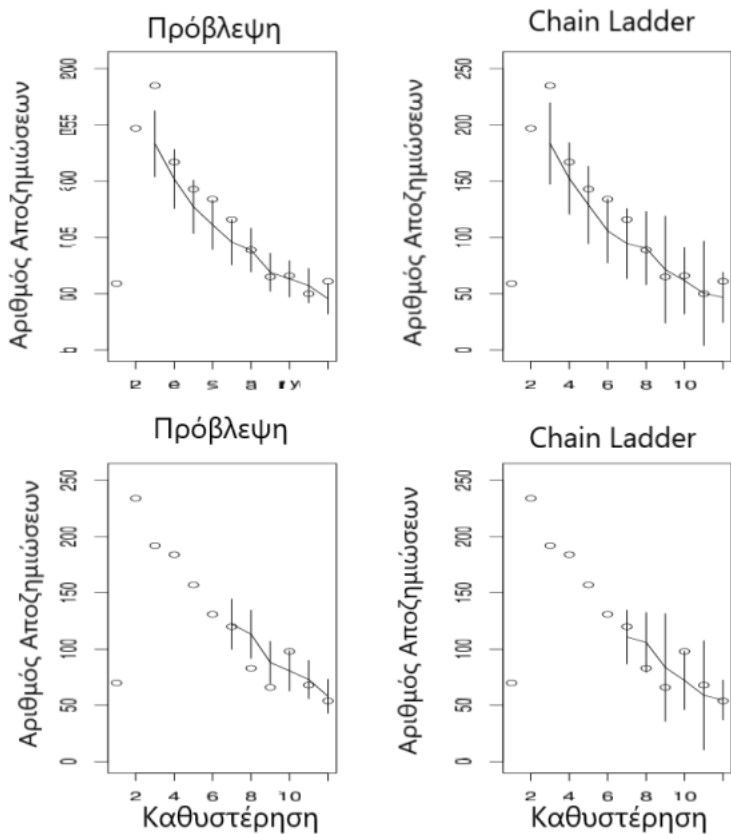
Στον Πίνακα 5.5 τα σχετικά σφάλματα πρόβλεψης των Chain Ladder προβλέψεων  $\tilde{N}_{i,12+k}$ ,  $k = 1, \dots, 12 - i + 1$ , υπολογίζονται από τους Πίνακες 5.2 και 5.3. Πάλι, τα σχετικά σφάλματα της πρόβλεψης κυμαίνονται ανάμεσα σε αρνητικές και θετικές τιμές. Μία σύγκριση του Πίνακα 5.4 δείχνει ότι καμία από τις δυο μεθόδους φαίνεται να ξεπερνάει η μία την άλλη, με την εξαίρεση της τελευταίας γραμμής στον Πίνακα 5.5 που απέχει πολύ από τις πραγματικές τιμές αποζημιώσεων. Είναι ακόμα άτοπο να πούμε σε ποια περιοχή των Πινάκων 5.4 και 5.5 κάποια από τις μεθόδους έχει κάποια πλεονεκτήματα.

Στην δεξιά πλευρά του Σχήματος 5.3 βλέπουμε την πρόβλεψη από την Chain Ladder για τις ίδιες περιπτώσεις με την αριστερή στήλη. Πάλι, μία άμεση σύγκριση φαίνεται να είναι δύσκολη αν και τα σφάλματα στην Chain Ladder φαίνεται να είναι μεγαλύτερα

από την δικιά μας μέθοδο. Στην δεξιά στήλη, τα όρια γύρω από τις προβλέψεις δηλώνουν  $\pm 1.96$  φορές το τετράγωνο του υπό προϋποθέσεις σφάλματος της πρόβλεψης που δίνεται από τον Mack (σελ. 363).

Αν υποθέσουμε  $M_j$  ακολουθούν την κατανομή Poisson, το Λήμμα 5.5.2 δείχνει την κύρια υπόθεση την Chain Ladder προσέγγισης, γραμμικότητα της υπό όρους εκτίμησης στην (5.2.2), δεν ικανοποιείται ακόμα και σε ασυμπτωτικό νόημα. Δοθέντος το γεγονός ότι μόνο 12 μήνες εισέρχονται και οι αντίστοιχες αποζημιώσεις δεν μπορούν να επαληθευθούν. Παρόλες τις ατέλειες, η μελέτη μας σε πραγματικά δεδομένα δείχνει άμεσα πως η μέθοδος πρόβλεψης δεν είναι χειρότερη από την πρόβλεψη της Chain Ladder σε σχέση με το πόσο κοντά είναι στους πραγματικούς αριθμούς των αποζημιώσεων και σε μέγεθος από τα υπό όρους σφάλματα.

Για τις τρεις κατανομές της  $(a, b)$  – τάξης η μέθοδος πρόβλεψης μας είναι εύκολα εφαρμοσμένη με την χρήση ενός απλού λογισμικού. Όταν παίρνουμε την Παρατήρηση 5.4.4, παίρνουμε γρήγορα αριθμητικά αποτελέσματα για το πρόβλημα της πρόβλεψης. Παρατηρούμε τις κατανομές της  $(a, b)$  – τάξης να είναι συχνά χρησιμοποιημένες και σε άλλες εφαρμογές. Το μοντέλο μας απαιτεί τη γνώση των κατανομών  $(q_m)$ ,  $(p_j)$  και την παράμετρο Poisson  $\mu$ . Οι ποσότητες αυτές είναι εύκολο να εκτιμηθούν αν υπάρχουν δεδομένα διαθέσιμα από διαφορετικές περιόδους.



*Σχήμα 5.3. Προβλέψεις (σταθερές γραμμές) των μηνιαίων αριθμών αποζημιώσεων με την χρήση του θεωρήματος της υποπαραγράφου 3.2 (αριστερά γραφήματα) και η προσέγγιση της Chain Ladder (δεξιά γραφήματα). Οι κουκίδες αντιπροσωπεύουν τους αριθμούς των αποζημιώσεων για πραγματικά δεδομένα. Οι κύκλοι γύρω από τις προβλέψεις αντιπροσωπεύουν  $\pm 1.96$  φορές το τετράγωνο των υπό όρων σφαλμάτων της πρόβλεψης. Τα πάνω γραφήματα αντιστοιχούν στην 11<sup>η</sup> γραμμή στους Πίνακες 5.2 και 5.3, και τα κάτω στην γραμμή 7.*

Τα θεωρητικά και τα εμπειρικά αποτελέσματα αυτού του κεφαλαίου δείχνουν ότι είναι αξιόλογο να θεωρήσουμε την στοχαστική διαδικασία για την παραγωγή της δυναμικής σε ένα χαρτοφυλάκιο ασφαλίσεων ζημιών. Ακόμα και αν δεν έχουμε αρκετή πληροφορία σε όλα τα στοιχεία του μοντέλου, το στοχαστικό μοντέλο επιτρέπει την ακριβή πρόβλεψη των μελλοντικών αποζημιώσεων. Επιπρόσθετα, το μοντέλο αυτό μπορεί εύκολα να προσομοιωθεί.

## Βιβλιογραφία

- 1) Χατζηκωνσταντινίδης Ε. (Πανεπιστημιακές Σημειώσεις στις Ασφαλίσεις Ζημιών)
- 2) Benktander, G. : «An approach to credibility in calculating IBNR for casualty excess reinsurance.», Actuar. Rev. 3, no. 2, 7, (1976)
- 3) Boor, J.: «Estimating tail development factors-What to do when the triangle runs out.», CAS Forum Winter 2006, 245-390. (2006)
- 4) Bornhuetter, R.L. and Ferguson, R.E. : «The Actuary and IBNR. Proceedings of the Casualty Actuarial Society», vol. LIX, 181-195. (1972)
- 5) Buchwalder, M., Bühlmann, H., Merz, M. and Wüthrich, M. : «The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited)». Astin Bulletin, vol 36, 521-543. (2006)
- 6) Cairns, J. G. C. : «A discussion of parameter and model uncertainty in insurance.» Insurance: Mathematics and Economics, vol. 27, 313-330. (2000)
- 7) Casella, G. and Berger, R. L. : «Statistical Inference», 2nd Edt. Duxbury Press. (2002)
- 8) Efron, B. : «Bootstrap methods: Another look at the jackknife. » Ann. Statist., vol. 7, 1-26. (1979)
- 9) England, P. D. and Verrall, R. J. : «Analytic and bootstrap estimates of prediction errors in claims reserving». Insurance: Mathematics and Economics, vol. 25, 281-293. (1999)
- 10) Hachemeister, C.A. and Stanard, J.N. : «IBNR claims counts estimation with static lag function.» 12th ASTIN Colloquium, IAA, Portimao, Portugal. (1975)
- 11) Hedegaard Jessen, A. : «Chain Ladder and its extensions.» Department of Mathematical Sciences, University of Copenhagen. (2009)
- 12) Hedegaard Jessen, A. and Rietdorf N. : «Diagonal Effects in Claims Reserving.» Department of Mathematical Sciences, University of Copenhagen. (2009)
- 13) Hedegaard Jessen, A., Nielsen J. P. and Verrall, R. : «Prediction of RBNS and IBNR claims using claim amounts and claim counts.» Department of Mathematical Sciences, University of Copenhagen. (2009)
- 14) Hedegaard Jessen, A., Mikosch, T and Samorodnitsky G. : «Prediction of outstanding payments in a Poisson Cluster Model.» Department of Mathematical Sciences, University of Copenhagen. (2009)
- 15) K. T. Hess, A. Liewald, and K. D. Schmidt, «An extension of Panjer's recursion,» ASTIN Bull. J. IAA, vol. 32, no. 2, pp. 283-297. (2002)
- 16) Kremer, E. : «IBNR-claims and the two-way model of AVNOVA». Scandinavian Actuarial Journal, 47-55. (1982)
- 17) Kremer, E. : «Einführung in die Versicherungsmathematik». Vandenhoe und Ruprecht, Gottingen. (1985)

- 18) Kuang, D., Nielsen, B. and Nielsen, J.P. : «Forecasting with the age-period-cohort model and the extended chain ladder model.» *Biometrika*, vol. 95, 979-986. (2008a)
- 19) Kuang, D., Nielsen, B. and Nielsen, J. P. : «Identification of the age-period-cohort model and the extended chain ladder model». *Biometrika*, vol. 95, 987-991. (2008b)
- 20) Liu, H. and Verrall, R. : «Predictive Distributions for Reserves which Separate True IBNR and IBNER Claims.» To appear in *Astin Bulletin*. (2009)
- 21) Mack, T. : «A simple parametric model for rating automobile insurance or estimating IBNR claims reserves». *Astin Bulletin*, vol. 21, no. 1, 93-108. (1991)
- 22) Mack, T. : «Distribution-free calculation of the standard error of chain ladder reserve estimates». *Astin Bulletin*, vol. 23, no. 2, 213-225. (1993)
- 23) Mack, T. : «Which stochastic model is underlying the chain ladder method.» *Insur. Math. Econ.*, vol. 15, no. 2-3, pp. 133-138 (1994)
- 24) Mack, T. : «Measuring the variability of chain ladder reserve estimates». *Casualty Actuarial Society Forum Spring 1994*, vol. 1, 101-182. (1994)
- 25) Mack, T. : «The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor.» *ASTIN Bull. J. IAA*, vol. 29, no. 2, pp. 361-366. (1999)
- 26) Mack, T. : «Credible claims reserves the Benktander method». *Astin Bulletin*, vol. 30, no. 2, 333-347. (2000)
- 27) McCullagh, P. and Nelder, J. A. : «Generalized Linear Models, second edition.» Chapman & Hall, London. (1989)
- 28) Mikosch, T. : «Non-Life Insurance Mathematics.» Springer-Verlag Berlin Heidelberg. (2004)
- 29) Neuhaus, W. : «On the estimation of outstanding claims.» *Australian Actuarial Journal*, vol. 10, 485-518. (2004)
- 30) Nelder, J.A. and Pregibon, D. : «An extended quasi-likelihood function.» *Biometrika*, vol. 74, 221-232. (1987)
- 31) Norberg, R. : «Prediction of outstanding liabilities in non-life insurance.» *Astin Bulletin*, vol. 23, no. 1, 95-115. (1993)
- 32) Norberg, R. : «Prediction of outstanding claims: Model variations and extensions.» *Astin Bulletin*, vol. 29, no. 1, 5-25. (1999)
- 33) Ntzoufras, I. and Dellaportas, P: «Bayesian modeling of outstanding liabilities incorporating claim count uncertainty.» *North American Actuarial Journal*, vol. 6, no. 1, 113-137. (2002)
- 34) Quarg, G. and Mack, T. : «Munich chain ladder.» *Blätter DGVM*, vol. 26, 597-630. (2004)
- 35) Renshaw, A. E. and Verrall, R. : «The stochastic model underlying the chain ladder technique.» *Actuarial research paper no. 63*. (1994)



- 36) Rietdorf, N. : «The chain ladder in a GLM setup with an extension that includes calendar effects.» Master Thesis, University of Copenhagen, Department of Mathematics Sciences. (2008)
- 37) Schmidt, K. D. : «A bibliography on loss reserving.» Available on: <http://www.math.tu-dresden.de/sto/schmidt/dsvm/reserve.pdf> (2007)
- 38) Schmidt, K.D. and Wünsche, A. : «Chain ladder, marginal sum and maximum likelihood estimation». Blätter, vol. 23, no. 3, 267-277. (1998)
- 39) Schmidt, K. D. and Zocher, M.: «The Bornhuetter – Ferguson principal.» Variance no. 2, 85-110. (2008)
- 40) Schnieper, R. : «Separating true IBNR and IBNER claims.» Astin Bulletin, vol. 21, 111-127. (1991)
- 41) Taylor, G. : «Separation of inflation and other effects from the distribution of non-life insurance claim delays.» Astin Bulletin, vol. 9, 219-230. (1977)
- 42) Taylor, G. : «Loss reserving, An Actuarial Perspective.» KAP Boston. (2000)
- 43) Taylor, G. : «Chain ladder bias». Astin Bulletin, vol. 33, no. 2, 313-330. (2003)
- 44) Taylor, G. : «Loss reserving, An Actuarial Perspective». KAP Boston.
- 45) Taylor, G. C. and McGuire, G. : «Loss reserving with GLM - a case study.» CAS discussion paper program, 327-391. (2004)
- 46) Thomas, M. : «Non-Life Insurance Mathematics.» Springer-Verlag, Berlin. (2004)
- 47) Venter, G. G. : «Discussion of the mean square error of prediction the chain ladder reserving method». Astin bulletin, vol. 36, 566-571. (2006)
- 48) Venter, G. G. : «Refining reserve runoff Ranges». CAS E-forum: <http://casact.org/pubs/forum/o7sforum/o7s-venter2.pdf> (2007)
- 49) Verrall, R. J. : «Bayes and empirical Bayes estimates for the chains ladder». Astin Bulletin, vol. 20, 217-243. (1990)
- 50) Verbeek, H.G. : «An approach to the analysis of claims experience in motor liability excess of loss reinsurance.» Astin Bulletin, vol. 6, 195-202. (1972)
- 51) Verrall, R. J. : «Chain ladder and maximum likelihood». Journal of the institute of actuaries, vol. 118, 489-499. (1991)
- 52) Verrall, R. J. : «A Bayesian generalized linear model for the Bornhuetter Ferguson method of claims reserving.» North American Actuarial Journal, vol. 8, 67-89. (2004)
- 53) Verrall, R. J. and England, P. D. : «Incorporating expert opinion into a stochastic model for the chain-ladder technique». Insurance, mathematic and economics, vol. 37, 355-370 . (2005)
- 54) Wright, T. S. : «A stochastic method for claims reserving in general insurance.» Journal of the Institute of Actuaries, vol. 117, 677-731. (1990)
- 55) Wüthrich, M.V. : «Using a Bayesian approach for claims reserving.» Variance, vol. 1, 292-301. (2007)

- 56) Zehnwirth, B. : «ICRFS-Plus 9.2 Manual. Stochastic loss reserving.» Insureware Pty Ltd. St. Kilda, Australia. (2003)
- 57) [https://en.wikipedia.org/wiki/Chain-ladder\\_method](https://en.wikipedia.org/wiki/Chain-ladder_method)
- 58) [www.casact.org/library/studynotes/Friedland\\_estimating.pdf](http://www.casact.org/library/studynotes/Friedland_estimating.pdf)
- 59) [www.variancejournal.org/issues/01-02/214.pdf](http://www.variancejournal.org/issues/01-02/214.pdf)