

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ
ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

Μέθοδοι εκτίμησης της πιθανότητας αθέτησης για οφειλέτες υψηλής φερεγγυότητας

Δημήτρης Σ. Τσούκας

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

Πειραιάς

Οκτώβριος 2020

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣ του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμόν 12^η/24.06.2019

συνεδρίαση του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου.

Τα μέλη της Επιτροπής ήταν:

- Καθηγητής Μ. Κούτρας (Επιβλέπων)
- Αναπλ. Καθηγητής Κ. Πολίτης
- Αναπλ. Καθηγητής Ε. Χατζηκωνσταντινίδης

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

UNIVERSITY OF PIREUS



DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

**POSTGRADUATE PROGRAM IN ACTUARIAL SCIENCE AND RISK
MANAGEMENT**

**Default probability estimation techniques for
high credibility obligors**

Dimitris S. Tsoukas

MSc Dissertation

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

Piraeus Greece

October 2020

Στην οικογένειά μου

ΕΥΧΑΡΙΣΤΙΕΣ

Με αφορμή την ολοκλήρωση της παρούσας διπλωματικής εργασίας θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κύριο Κούτρα Μάρκο για την καθοδήγηση του για την ολοκλήρωση της διπλωματικής και για το ενδιαφέρον του σε όλη την πορεία μου στο μεταπτυχιακό πρόγραμμα.

Τέλος, οφείλω το μεγαλύτερο ευχαριστώ στην οικογένειά μου για την στήριξη τους σε όλη τη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Ο όρος χαρτοφυλάκια υψηλής φερεγγυότητας χρησιμοποιείται για να περιγράψει χαρτοφυλάκια δανείων στα οποία έχει συμβεί μικρός ή ακόμα και μηδενικός αριθμός αθετήσεων. Λόγω αυτής της ιδιότητας, η εκτίμηση της πιθανότητας αθέτησης στα χαρτοφυλάκια υψηλής φερεγγυότητας καθίσταται δύσκολη. Στην παρούσα διπλωματική εργασία, αναφέρουμε αρχικά, κάποιες βασικές έννοιες και εξηγούμε γιατί σε αυτά τα χαρτοφυλάκια δεν μπορούν να χρησιμοποιηθούν τα κλασσικά στατιστικά μοντέλα. Παρουσιάζουμε την μέθοδο εκτίμησης της πιθανότητας αθέτησης μέσω του πιο συντηρητικού εκτιμητή και της προσέγγισης Bayes. Επιπρόσθετα, εξετάζουμε την καμπύλη ισχύος, η οποία είθισται να χρησιμοποιείται για την αξιολόγηση του μοντέλου βαθμονόμησης, και πως μέσω αυτής μπορούμε να εκτιμήσουμε την πιθανότητα αθέτησης. Τέλος, με χρήση προσομοίωσης βαίνουμε σε αξιολόγηση της αποτελεσματικότητας της κάθε μεθόδου.

ABSTRACT

The term low default portfolios is used to describe loan portfolios in which a small or even zero number of defaults has occurred. Due to this property, it is difficult to estimate the default probability on low default portfolios. In the present master thesis, we first introduce some basic concepts and explain why the most popular statistical models cannot be used in these portfolios. We present in detail how to estimate the default probability through the most prudent estimation and the Bayesian approach. In addition, we present the power curve, which is commonly used to calibrate the model, and illustrate how we can use it to estimate the probability of default. Finally, we perform a simulation study in order to evaluate the effectiveness of the methods presented in this thesis.

Περιεχόμενα

| | |
|---|----|
| Κατάλογος σχημάτων | 10 |
| Κατάλογος πινάκων | 11 |
| 1. Τα μοντέλα πιστοληπτικής βαθμολόγησης και τα χαρτοφυλάκια υψηλής φερεγγυότητας | 13 |
| 1.1 Εισαγωγή..... | 13 |
| 1.2 Τα στατιστικά μοντέλα πιστοληπτικής βαθμολόγησης..... | 13 |
| 1.2.1. Το μοντέλο του Altman | 13 |
| 1.2.2 Το μοντέλο λογιστικής παλινδρόμησης..... | 14 |
| 1.2.3 Τα δέντρα απόφασης | 15 |
| 1.3 Η αξία της πιθανότητας αθέτησης..... | 15 |
| 1.4 Χαρτοφυλάκια υψηλής φερεγγυότητας..... | 16 |
| 2. Η πιο συντηρητική αρχή εκτίμησης..... | 19 |
| 2.1 Καμία αθέτηση και ανεξαρτησία των παρατηρήσεων του δείγματος | 19 |
| 2.2 Μικρός αριθμός αθετήσεων και ανεξαρτησία των παρατηρήσεων του δείγματος | 24 |
| 2.3 Καμία αθέτηση και εξάρτηση των παρατηρήσεων του δείγματος | 28 |
| 2.4 Μικρός αριθμός αθετήσεων και εξάρτηση των παρατηρήσεων του δείγματος | 31 |
| 2.5 Περίοδος παρατήρησης άνω του ενός έτους..... | 33 |
| 2.6 Η χρήση της γλώσσας R στην εκτίμηση των πιθανοτήτων αθέτησης..... | 36 |
| 2.7 Σχολιασμός της μεθόδου..... | 37 |
| 3. Η προσέγγιση Bayes | 39 |
| 3.1 Το θεώρημα Bayes | 39 |
| 3.2 Η κατανομή του δείγματος..... | 40 |
| 3.3 Επιλογή της εκ των προτέρων κατανομής | 41 |
| 3.4 Η εκτίμηση της πιθανότητας αθέτησης..... | 42 |
| 3.4.1 Η κατανομή Beta με δύο παραμέτρους..... | 42 |
| 3.4.2 Η περίπτωση της συντηρητικής εκ των προτέρων κατανομής | 46 |
| 3.4.3 Η ομοιόμορφη εκ των προτέρων κατανομή..... | 48 |
| 3.5 Εξάρτηση του δείγματος και περίοδος παρατήρησης άνω του ενός έτους | 49 |
| 3.6 Σχολιασμός της μεθόδου..... | 50 |
| 4. Η καμπύλη ισχύος..... | 51 |
| 4.1 Η κατασκευή της καμπύλης CAP | 51 |
| 4.2 Μοντελοποίηση της καμπύλης..... | 52 |
| 4.3 Δείκτες αξιολόγησης και η κοιλότητα της καμπύλης | 54 |
| 4.4 Υπολογισμός της κοιλότητας και εκτίμηση της πιθανότητας αθέτησης | 56 |
| 4.5 Η χρήση της γλώσσας R στην εύρεση της παραμέτρου κοιλότητας | 61 |

| | |
|---|----|
| 4.6 Σχολιασμός της μεθόδου | 62 |
| 5. Προσομοίωση και αξιολόγηση των μεθόδων | 63 |
| 5.1 Προσομοίωση και αξιολόγηση της συντηρητικής αρχής εκτίμησης..... | 63 |
| 5.1.1 Αξιολόγηση της μεθόδου υπό την υπόθεση της ανεξαρτησίας..... | 63 |
| 5.1.2 Αξιολόγηση της μεθόδου υπό την υπόθεση της εξάρτησης του δείγματος για μονοετή περίοδο παρατήρησης..... | 65 |
| 5.1.3 Αξιολόγηση της μεθόδου υπό την υπόθεση της εξάρτησης του δείγματος για πολυετή περίοδο παρατήρησης..... | 66 |
| 5.2 Προσομοίωση και αξιολόγηση της προσέγγισης Bayes | 68 |
| 5.2.1 Αξιολόγηση της μεθόδου λαμβάνοντας ως εκ των προτέρων κατανομή την Beta | 68 |
| 5.2.2 Αξιολόγηση της μεθόδου με χρήση της συντηρητικής εκ των προτέρων κατανομής | 69 |
| 5.2.3 Αξιολόγηση της μεθόδου με χρήση της ομοιόμορφης εκ των προτέρων κατανομής | 70 |
| 5.3 Προσομοίωση και αξιολόγηση της καμπύλης CAP..... | 71 |
| 6. Σύνοψη | 75 |
| Παράρτημα..... | 77 |
| Αναφορές..... | 90 |

Κατάλογος σχημάτων

| | |
|--|----|
| Σχήμα 2.1 Γραφική απεικόνιση της pA για κάθε χαρτοφυλάκιο..... | 22 |
| Σχήμα 2.2 Γραφική απεικόνιση της pB για κάθε χαρτοφυλάκιο | 22 |
| Σχήμα 2.3 Γραφική απεικόνιση της $p\Gamma$ για κάθε χαρτοφυλάκιο..... | 22 |
| Σχήμα 2.4 Γραφική απεικόνιση της pA για κάθε τιμή γ | 22 |
| Σχήμα 2.5 Γραφική απεικόνιση της pB για κάθε τιμή γ | 22 |
| Σχήμα 2.6 Γραφική απεικόνιση της $p\Gamma$ για κάθε τιμή γ | 22 |
| Σχήμα 2.7 Γραφική απεικόνιση της pA χωρίς εξάρτηση δείγματος με αθετήσεις | 26 |
| Σχήμα 2.8 Γραφική απεικόνιση της pB χωρίς εξάρτηση δείγματος με αθετήσεις | 26 |
| Σχήμα 2.9 Γραφική απεικόνιση της $p\Gamma$ χωρίς εξάρτηση δείγματος με αθετήσεις | 26 |
| Σχήμα 2.10 Γραφική απεικόνιση της pA για κάθε τιμή του γ με αθετήσεις | 27 |
| Σχήμα 2.11 Γραφική απεικόνιση της pB για κάθε τιμή του γ με αθετήσεις | 26 |
| Σχήμα 2.12 Γραφική απεικόνιση της $p\Gamma$ για κάθε τιμή του γ με αθετήσεις | 27 |
| Σχήμα 2.13 Γραφική απεικόνιση της pA εξάρτηση του δείγματος, χωρίς αθετήσεις..... | 31 |
| Σχήμα 2.14 Γραφική απεικόνιση της pB εξάρτηση του δείγματος, χωρίς αθετήσεις | 31 |
| Σχήμα 2.15 Γραφική απεικόνιση της $p\Gamma$ εξάρτηση του δείγματος, χωρίς αθετήσεις..... | 31 |
| Σχήμα 2.16 Γραφική απεικόνιση της pA με εξάρτηση του δείγματος και αθετήσεις..... | 33 |
| Σχήμα 2.17 Γραφική απεικόνιση της pB με εξάρτηση του δείγματος και αθετήσεις | 33 |
| Σχήμα 2.18 Γραφική απεικόνιση της $p\Gamma$ με εξάρτηση του δείγματος και αθετήσεις | 33 |
| Σχήμα 3.1 Συνάρτηση πυκνότητας πιθανότητας κατανομών Beta..... | 43 |
| Σχήμα 3.2 Πιθανότητα αθέτησης με χρήση Beta εκ των προτέρων κατανομής | 45 |
| Σχήμα 3.3 Γραφική παράσταση της συντηρητικής εκ των προτέρων κατανομής..... | 47 |
| Σχήμα 3.4 Πιθανότητα αθέτησης με conservative prior, Beta prior, prudent estimation και MLE..... | 48 |
| Σχήμα 4.1 Καμπύλες CAP για 3 διαφορετικά μοντέλα αξιολόγησης..... | 52 |
| Σχήμα 4.2 Καμπύλη CAP και $k=5$, $k=20$ και $k=-2$ | 53 |
| Σχήμα 4.3 Πολυπαραμετρική καμπύλη CAP | 54 |
| Σχήμα 4.4 Προσέγγιση της κοιλότητας για διάφορες τιμές του AUC..... | 55 |
| Σχήμα 4.5 Προσέγγιση της κοιλότητας για διάφορες τιμές του AR..... | 56 |

| | |
|--|----|
| Σχήμα 4.6 Προσαρμογή της καμπύλης $y(x)$ για $k = 6.96$ στα δεδομένα του χαρτοφυλακίου 7 | 58 |
| Σχήμα 4.7 Καμπύλες CAP για $k = 6.96$ και $k = 6.67$ στο χαρτοφυλάκιο 7..... | 59 |
| Σχήμα 4.8 Εκτίμηση των $pA, pB, p\Gamma, p\Delta, pE$ για 2 διαφορετικές τιμές του concavity | 59 |
| Σχήμα 5.1 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για την βαθμίδα A | 64 |
| Σχήμα 5.2 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για την βαθμίδα B | 64 |
| Σχήμα 5.3 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για την βαθμίδα Γ | 64 |
| Σχήμα 5.4 Σύγκριση της πραγματικής και της εκτιμώμενης pA για το χαρτοφυλάκιο 8 | 65 |
| Σχήμα 5.5 Σύγκριση της πραγματικής και της εκτιμώμενης pB για το χαρτοφυλάκιο 8 | 65 |
| Σχήμα 5.6 Σύγκριση της πραγματικής και της εκτιμώμενης $p\Gamma$ για το χαρτοφυλάκιο 8 | 66 |
| Σχήμα 5.7 Σύγκριση της πραγματικής και της εκτιμώμενης pA για το χαρτοφυλάκιο 9 | 67 |
| Σχήμα 5.8 Σύγκριση της πραγματικής και της εκτιμώμενης pB για το χαρτοφυλάκιο 9 | 67 |
| Σχήμα 5.9 Σύγκριση της πραγματικής και της εκτιμώμενης $p\Gamma$ για το χαρτοφυλάκιο 9 | 67 |
| Σχήμα 5.10 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 10 με χρήση Beta prior | 69 |
| Σχήμα 5.11 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 11 με χρήση conservative prior και MLE | 70 |
| Σχήμα 5.12 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 12 με χρήση Ομοιόμορφης εκ των προτέρων κατανομής | 71 |
| Σχήμα 5.13 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 13 | 73 |

Κατάλογος πινάκων

| | |
|---|----|
| Πίνακας 2.1 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 1 χωρίς αθετήσεις..... | 21 |
| Πίνακας 2.2 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 2 χωρίς αθετήσεις..... | 21 |
| Πίνακας 2.3 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 3 χωρίς αθετήσεις..... | 21 |
| Πίνακας 2.4 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 4 χωρίς αθετήσεις | 21 |
| Πίνακας 2.5 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 5 χωρίς αθετήσεις..... | 21 |
| Πίνακας 2.6 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 6 χωρίς αθετήσεις..... | 22 |
| Πίνακας 2.7 Ο λόγος pA/pB για όλα τα χαρτοφυλάκια | 23 |
| Πίνακας 2.8 Ο λόγος $pB/p\Gamma$ για όλα τα χαρτοφυλάκια | 23 |
| Πίνακας 2.9 Οι λόγοι $nB + n\Gamma nA + nB + n\Gamma$ και $n\Gamma nB + n\Gamma$ για όλα τα χαρτοφυλάκια..... | 23 |
| Πίνακας 2.10 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 1 όταν έχουν συμβεί αθετήσεις | 25 |
| Πίνακας 2.11 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 2 όταν έχουν συμβεί αθετήσεις | 25 |
| Πίνακας 2.12 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 3 όταν έχουν συμβεί αθετήσεις | 25 |
| Πίνακας 2.13 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 4 όταν έχουν συμβεί αθετήσεις | 26 |
| Πίνακας 2.14 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 5 όταν έχουν συμβεί αθετήσεις | 26 |
| Πίνακας 2.15 Εκτίμηση των $pA, pB, p\Gamma$ για το χαρτοφυλάκιο 6 όταν έχουν συμβεί αθετήσεις | 26 |
| Πίνακας 2.16 Εκτίμηση των $pA, pB, p\Gamma$ για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 1..... | 30 |
| Πίνακας 2.17 Εκτίμηση των $pA, pB, p\Gamma$ για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 2..... | 30 |
| Πίνακας 2.18 Εκτίμηση των $pA, pB, p\Gamma$ για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 3..... | 30 |
| Πίνακας 2.19 Εκτίμηση των $pA, pB, p\Gamma$ για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 4..... | 30 |
| Πίνακας 2.20 Εκτίμηση των $pA, pB, p\Gamma$ για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 5..... | 30 |
| Πίνακας 2.21 Εκτίμηση των $pA, pB, p\Gamma$ για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 6..... | 30 |

| | |
|---|----|
| Πίνακας 2.22 Εκτίμηση των $pA, pB, pΓ$ για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 1 | 32 |
| Πίνακας 2.23 Εκτίμηση των $pA, pB, pΓ$ για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 2 | 32 |
| Πίνακας 2.24 Εκτίμηση των $pA, pB, pΓ$ για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 3 | 32 |
| Πίνακας 2.25 Εκτίμηση των $pA, pB, pΓ$ για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 4 | 32 |
| Πίνακας 2.26 Εκτίμηση των $pA, pB, pΓ$ για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 5 | 32 |
| Πίνακας 2.27 Εκτίμηση των $pA, pB, pΓ$ για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 6 | 32 |
| Πίνακας 2.28 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 1 στην πολυετή περίπτωση | 35 |
| Πίνακας 2.29 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 5 στην πολυετή περίπτωση | 35 |
| Πίνακας 2.30 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 4 στην πολυετή περίπτωση | 36 |
| Πίνακας 3.1 Εκτίμηση της πιθανότητας αθέτησης με χρήση Beta εκ των προτέρων κατανομής | 44 |
| Πίνακας 3.2 Εκτίμηση της πιθανότητας αθέτησης μέσω της συντηρητικής εκ των προτέρων κατανομής, της κατανομής <i>Beta</i> , της συντηρητικής αρχής εκτίμησης και του εκτιμητή μέγιστης πιθανοφάνειας | 47 |
| Πίνακας 3.3 Εκτίμηση της πιθανότητας αθέτησης με την ομοιόμορφη, συντηρητική, $Beta(2.1,300)$, $\theta 0.5, MLE$ | 49 |
| Πίνακας 4.1 Αθροιστικά ποσοστά οφειλετών και αθετήσεων για το χαρτοφυλάκιο 7 | 58 |
| Πίνακας 4.2 Εκτίμηση των πιθανοτήτων αθέτησης για $k = 6.96$ και $k = 6.67$ | 59 |
| Πίνακας 4.3 Εκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 7 με $k = 3.77$ | 60 |
| Πίνακας 5.1 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 8 | 64 |
| Πίνακας 5.2 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 8 για $\rho = 3\%$ | 65 |
| Πίνακας 5.3 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 9 με $T = 5, \rho = 15\%, \tau = 30\%$ | 66 |
| Πίνακας 5.4 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 9 με $T=5, \rho=12\%, \tau=30\%$ | 67 |
| Πίνακας 5.5 Εκτίμηση των $pA, pB, pΓ$ για το χαρτοφυλάκιο 9 με $T = 5, \rho = 12\%, \tau = 20\%$ | 68 |
| Πίνακας 5.6 Εκτίμηση της πιθανότητας αθέτησης με χρήση Beta εκ των προτέρων κατανομής για το χαρτοφυλάκιο 10 | 68 |
| Πίνακας 5.7 Εκτίμηση της πιθανότητας αθέτησης με την ομοιόμορφη κατανομή, την συντηρητική κατανομή και τον συντηρητικό εκτιμητή για $\gamma=50\%$ και 75% | 71 |
| Πίνακας 5.8 Εκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 13 για $k = 10.98$ | 72 |
| Πίνακας 5.9 Εκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 13 για $k = 6.14$ | 72 |
| Πίνακας 5.10 Εκτίμηση της πιθανότητας αθέτησης για τις βαθμίδες A και B με χρήση της συντηρητικής εκ των προτέρων, της ομοιόμορφης στο $(0,1)$ και του συντηρητικού εκτιμητή | 73 |

1. Τα μοντέλα πιστοληπτικής βαθμολόγησης και τα χαρτοφυλάκια υψηλής φερεγγυότητας

1.1 Εισαγωγή

Μία από τις κύριες δραστηριότητες ενός πιστωτικού ιδρύματος είναι η έγκριση ή η απόρριψη του αιτήματος δανειοδότησης ενός φυσικού ή νομικού προσώπου. Προκειμένου ένα πιστωτικό ίδρυμα, συνήθως μια τράπεζα, να αποφασίσει αν θα χορηγήσει δάνειο σε έναν πελάτη, θα πρέπει να αξιολογήσει, βάσει κάποιου κριτηρίου, αν αυτός ο πελάτης δύναται να αποπληρώσει στο μέλλον τις υποχρεώσεις του (“καλός πελάτης”) ή όχι (“κακός πελάτης”). Ο διαχωρισμός αυτός των πελατών είναι ζωτικής σημασίας, καθώς η έγκριση δανείων σε πρόσωπα τα οποία μπορεί να μην ανταποκριθούν στις υποχρεώσεις τους, ενδέχεται να επιφέρει καταστροφικούς κινδύνους για το ίδρυμα. Η ανάγκη αυτή λοιπόν, οδήγησε στη δημιουργία διαφόρων μοντέλων πιστοληπτικής βαθμολόγησης (**credit scoring models**). Για την κατασκευή τέτοιων μοντέλων γίνεται χρήση ιστορικών δεδομένων, που έχουν συλλεχθεί από προηγούμενες περιόδους. Με χρήση των δεδομένων αυτών, συντίθεται το μοντέλο μέσω του οποίου αξιολογείται ο υποψήφιος δανειολήπτης βάσει κάποιων χαρακτηριστικών όπως η ηλικία, ο μισθός και άλλα. Μέσω της αξιολόγησης αυτής αποφασίζεται η χορήγηση ή μη του δανείου. Στα χαρτοφυλάκια υψηλής φερεγγυότητας συνήθως έχουμε ελάχιστες έως μηδενικές περιπτώσεις αθετήσεων με αποτέλεσμα τα πιο δημοφιλή στατιστικά μοντέλα να μην μπορούν να παρέχουν αξιόπιστα αποτελέσματα. Ορισμένα από τα εν λόγω μοντέλα περιγράφονται στην επόμενη ενότητα.

1.2 Τα στατιστικά μοντέλα πιστοληπτικής βαθμολόγησης

Οι στατιστικές τεχνικές που χρησιμοποιούνται για τη μέτρηση της πιστοληπτικής ικανότητας είναι κυρίως:

- a) Η διαχωριστική ανάλυση
- b) Η λογιστική παλινδρόμηση
- c) Τα δέντρα απόφασης

Θα δώσουμε στην συνέχεια κάποιες λεπτομέρειες για αυτές τις τεχνικές και θα αναφερθούμε στα προβλήματα που ανακύπτουν στα χαρτοφυλάκια υψηλής φερεγγυότητας.

1.2.1. Το μοντέλο του Altman

Το μοντέλο του Altman (1968), είναι ευρύτερα γνωστό ως Z-score model και βασίζεται στην τεχνική της διαχωριστικής ανάλυσης. Χρησιμοποιώντας δείγμα 66 επιχειρήσεων, εκ των οποίων οι μισές είχαν χρεοκοπήσει, βρήκε την παρακάτω συνάρτηση που εκφράζει την φερεγγυότητα του πιστούχου μέσω του δείκτη Z:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5$$

όπου,

- $X_1 = \text{κεφάλαιο κίνησης/συνολικό κεφάλαιο}$
 $X_2 = \text{αδιανέμητα κέρδη/συνολικό κεφάλαιο}$
 $X_3 = \text{κέρδη προ τόκων και φόρων/συνολικό κεφάλαιο}$
 $X_4 = \text{αγοραία αξία κεφαλαίων/λογιστική αξία συνολικού παθητικού}$
 $X_5 = \text{πωλήσεις/συνολικό κεφάλαιο}$

Βάσει της συγκεκριμένης συνάρτησης, ο Altman κατασκεύασε έναν κανόνα απόφασης, σύμφωνα με τον οποίο για $Z > 2.99$ ο πελάτης θεωρείται φερέγγυος, ενώ για $Z \leq 1.80$ αφερέγγυος. Για τιμές του $Z \in (1.8, 2.99]$ το μοντέλο δεν μπορεί να κατατάξει τον πελάτη σε καμία από τις 2 κατηγορίες.

1.2.2 Το μοντέλο λογιστικής παλινδρόμησης

Είναι σαφές ότι με το μοντέλο του Altman γίνεται κατάταξη του πελάτη, ωστόσο το μοντέλο δεν βασίζεται στην εκτίμηση της πιθανότητας αθέτησής του, η οποία κάτω από συγκεκριμένες προϋποθέσεις μπορεί να εκτιμηθεί εκ των υστέρων. Αντιθέτως, με το μοντέλο λογιστικής παλινδρόμησης μπορούμε να εκτιμήσουμε την πιθανότητα αθέτησης ενός υποψήφιου δανειολήπτη βάσει κάποιων μεταβλητών (ποσοτικών και ποιοτικών), οι οποίες θεωρούνται σημαντικές για το μοντέλο. Συμβολίζοντας με p την πιθανότητα αθέτησης την οποία θέλουμε να υπολογίσουμε και θέτοντας

$$Y = \begin{cases} 1, & \text{με πιθανότητα } p \\ 0, & \text{με πιθανότητα } 1 - p \end{cases} \quad (1.1)$$

Το κλασικό μοντέλο της λογιστικής παλινδρόμησης έχει την εξής μορφή:

$$\ln \frac{E(Y)}{1-E(Y)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{p-1} X_{p-1} \quad (1.2)$$

Όπως αναφέραμε παραπάνω, μέσω του συγκεκριμένου μοντέλου μπορούμε να χειριστούμε και ποιοτικές μεταβλητές. Αυτό γίνεται εφικτό με τη χρήση δείκτριων συναρτήσεων.

Με γνώμονα τα ιστορικά δεδομένα που διαθέτει το πιστωτικό ίδρυμα, μπορούν να εκτιμηθούν οι τιμές των παραμέτρων $\beta_0, \beta_1, \dots, \beta_{p-1}$, συνήθως με τη μέθοδο της μέγιστης πιθανοφάνειας, και να κατασκευαστούν διαστήματα εμπιστοσύνης. Μάλιστα, γνωρίζοντας τις τιμές των μεταβλητών X_1, X_2, \dots, X_{p-1} μπορεί να υπολογιστεί το δεξί μέλος της (1.2). Από την (1.1) είναι προφανές ότι $E(Y) = p$, άρα η (1.2) μπορεί να γραφεί ως :

$$p = \frac{e^z}{1+e^z} \quad (1.3)$$

όπου $z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{p-1} X_{p-1}$. Μέσω της (1.3), υπολογίζεται η ζητούμενη πιθανότητα αθέτησης και αποφασίζεται αν θα εγκριθεί ή όχι το δάνειο.

1.2.3 Τα δέντρα απόφασης

Τα δένδρα απόφασης ή δένδρα ταξινόμησης αποτελούν ακόμα ένα είδος μοντέλου πιστοληπτικής βαθμολόγησης. Η συγκεκριμένη μέθοδος ξεκινάει με όλο το δείγμα σε μια ομάδα και ανάλογα με τις τιμές που παρατηρούνται στα διάφορα χαρακτηριστικά, τα οποία έχουν κριθεί ως σημαντικά, κατανέμει το δείγμα σε υποομάδες. Η διαδικασία συνεχίζεται, έως ότου ικανοποιηθεί ένα κριτήριο τερματισμού.

Εν γένει, τα δένδρα απόφασης συγκροτούν έναν σχετικά αξιόπιστο και εύκολα κατανοητό τρόπο για να γίνει η κατηγοριοποίηση ενός υποψήφιου δανειολήπτη σε “καλό” ή “κακό” πελάτη.

Επιπρόσθετα, στα δένδρα απόφασης μπορούμε, όπως και στα μοντέλα λογιστικής παλινδρόμησης, να χειριστούμε και ποιοτικές μεταβλητές. Επίσης με τη συγκεκριμένη μέθοδο, η απόφαση για τη δανειοδότηση θα γίνει βάσει της πιθανότητας αθέτησης που θα έχει εκτιμηθεί, μέσω της υποομάδας στην οποία κατατάχθηκε ο υποψήφιος δανειολήπτης.

Η δυσκολία της συγκεκριμένης μεθόδου έγκειται στον διαχωρισμό των παρατηρήσεων στις διάφορες υποομάδες, στην επιλογή των μεταβλητών που επηρεάζουν την εκτιμώμενη πιθανότητα αθέτησης και στα σημεία αποκοπής.

Τέλος, στα δένδρα απόφασης ύψιστη σημασία για την ταξινόμηση έχει ο αλγόριθμος κατασκευής του δένδρου. Τα 4 είδη δένδρων ταξινόμησης που υπάρχουν, ανάλογα με τη μέθοδο κατασκευής τους είναι τα C&R, CHAID, QUEST και C5.0.

Είναι υπαρκτό το ενδεχόμενο με ίδιο δείγμα, αλλά με διαφορετική μέθοδο κατασκευής δένδρου ταξινόμησης να καταλήξουμε σε τελείως διαφορετικά αποτελέσματα.

1.3 Η αξία της πιθανότητας αθέτησης

Όπως αναφέρθηκε, με κάθε νέα αίτηση δανειοδότησης το πιστωτικό ίδρυμα προσπαθεί να εκτιμήσει, αν ο υποψήφιος δανειολήπτης θα αποδειχθεί “καλός” ή “κακός” πελάτης. Προτιμότερο όμως είναι, αντί για την απλή ταξινόμηση του πελάτη σε μία από τις δύο κατηγορίες να γίνει η εκτίμηση της πιθανότητας αθέτησης αυτού. Αν ορίσουμε με p την πιθανότητα αθέτησης, τότε ο πιο απλός διαχωρισμός των πελατών μπορεί να γίνει με βάση τον κάτωθι κανόνα:

$$Ο\ πελάτης\ είναι = \begin{cases} \text{Φερέγγυος, αν } p < 0.5 \\ \text{Αφερέγγυος, αν } p \geq 0.5 \end{cases} \quad (1.4)$$

Παρότι ο παραπάνω είναι ένας κοινά αποδεκτός κανόνας, για τα πιστωτικά ιδρύματα αυτό δεν ισχύει πάντα, καθώς σε περιπτώσεις οικονομικής ύφεσης ή συντηρητικής πολιτικής του ιδρύματος η διαχωριστική τιμή μπορεί να είναι διαφορετική.

Βάσει της πιθανότητας αθέτησης όμως, το πιστωτικό ίδρυμα εκτός από την κατάταξη του πελάτη δύναται να υπολογίσει και άλλες συνιστώσες της σύμβασης, εξίσου σημαντικές. Χαρακτηριστικό παράδειγμα αποτελεί η απόδοση του δανείου η οποία μπορεί να υπολογιστεί με τον παρακάτω τρόπο.

Έστω,

| | |
|---------|--|
| r | η ζητούμενη απόδοση του δανείου |
| p | η πιθανότητα μη αποπληρωμής του δανείου |
| $1 - p$ | η πιθανότητα κανονικής αποπληρωμής του δανείου |
| z | το ποσοστό ανάκτησης σε περίπτωση αθέτησης |
| r_f | η απόδοση μιας <<ακίνδυνης>> επένδυσης |

Τότε, πρέπει να ισχύει η εξής σχέση:

$$1 + r_f = (1 + r)(1 - p) + zp \quad (1.5)$$

Από την (1.5) το πιστωτικό ίδρυμα γνωρίζει το r_f , το οποίο είθισται να λαμβάνεται ως η απόδοση του 10ετούς γερμανικού ομολόγου, έχει εκτιμήσει επίσης το p και έχει αποφασιστεί μεταξύ των συμβαλλομένων το ποσοστό ανάκτησης z .

Με αυτά τα δεδομένα έχουμε ότι:

$$r = \frac{1+r_f-zp}{1-p} - 1 \quad (1.6)$$

Με βάση τα παραπάνω, γίνεται σαφές λοιπόν ότι μοντέλα που μπορούν να υπολογίσουν την πιθανότητα αθέτησης, είναι σαφώς πιο χρήσιμα από μοντέλα διαχωριστικής ανάλυσης.

1.4 Χαρτοφυλάκια υψηλής φερεγγυότητας

Τα χαρτοφυλάκια υψηλής φερεγγυότητας (low default portfolios) μπορούν να οριστούν ως εκείνα τα χαρτοφυλάκια στα οποία παρατηρείται ένας εξαιρετικά χαμηλός, έως και μηδενικός, αριθμός αντισυμβαλλομένων οι οποίοι έχουν αθετήσει τις υποχρεώσεις τους. Αποτελεί γεγονός ότι δεν υπάρχει ρητός ορισμός των χαρτοφυλακίων αυτών. Συνήθως, χαρτοφυλάκια τέτοιου είδους περιέχουν δάνεια σε τράπεζες, ασφαλιστικές εταιρείες, μεγάλες εταιρείες ή πρόκειται για χαρτοφυλάκια με πολύ μικρό αριθμό οφειλετών, εκ των οποίων κανείς δεν έχει σταματήσει να εκπληρώνει τις υποχρεώσεις που έχει απέναντι στο πιστωτικό ίδρυμα. Υπάρχουν αρκετά προβλήματα όσον αφορά τη μοντελοποίηση και τη στάθμιση της πιθανότητας αθέτησης σε χαρτοφυλάκια υψηλής φερεγγυότητας, αλλά ο κυριότερος λόγος ύπαρξης αυτών των προβλημάτων είναι ότι συνήθως υπάρχουν ελάχιστες περιπτώσεις αθετήσεων. Επίσης, είναι ενυπόστατο το σενάριο της μη ύπαρξης αθετήσεων σε ένα χαρτοφυλάκιο, αλλά είναι σαφές ότι η πραγματική πιθανότητα αθέτησης δεν μπορεί να θεωρηθεί μηδενική. Ακόμα και σε μεγάλα χαρτοφυλάκια, όταν παρατηρούνται λίγες αθετήσεις, μπορεί να υπάρξει υποτίμηση της πιθανότητας αθέτησης.

Ένα ακόμα πρόβλημα που παρατηρείται σε τέτοιου είδους χαρτοφυλάκια είναι η μεταβλητότητα των παρατηρούμενων αθετήσεων. Είναι πιθανό, να περάσουν αρκετά χρόνια δίχως να παρατηρηθεί κάποια αθέτηση, αλλά αυτό δεν σημαίνει ότι δεν θα παρατηρηθεί κάποια στιγμή κρίση στο χαρτοφυλάκιο.

Τα θέματα αυτά είναι αρκετά σημαντικά για τα πιστωτικά ιδρύματα, καθώς ένα σημαντικό ποσό της έκθεσής τους στον κίνδυνο μπορεί να προέρχεται από χαρτοφυλάκια υψηλής φερεγγυότητας. Τα ιδρύματα, είναι υποχρεωμένα βάσει νομοθεσίας να κρατήσουν κεφάλαια, προκειμένου να μπορούν να καλύψουν ένα ορισμένο μέρος του κινδύνου. Αν χρησιμοποιηθούν στατιστικές τεχνικές μέσω των οποίων η πιθανότητα αθέτησης υπερεκτιμηθεί, τότε ίδρυμα θα πρέπει να κρατήσει παραπάνω κεφάλαια χάνοντας έτσι μέρος των κερδών που μπορεί να έχει. Πρόβλημα όμως μπορεί να προκύψει και στην περίπτωση της υποεκτίμησης της πιθανότητας αθέτησης. Τότε, το ίδρυμα θα κρατήσει λίγα κεφάλαια και μια ενδεχόμενη κρίση στην εγχώρια ή την παγκόσμια οικονομία θα μπορούσε να έχει καταστρεπτικές συνέπειες για τη λειτουργία του ιδρύματος.

Όπως προαναφέρθηκε, το μεγαλύτερο πρόβλημα που υπάρχει στα χαρτοφυλάκια υψηλής φερεγγυότητας είναι ότι στα ιστορικά δεδομένα υπάρχει υπερβολικά μικρός αριθμός “κακών” πελατών, με αποτέλεσμα την έλλειψη στατιστικών στοιχείων ικανών να μας οδηγήσουν στην κατασκευή ενός κλασικού μοντέλου για την εκτίμηση της πιθανότητας αθέτησης ενός οφειλέτη, που υπάγεται σε ένα τέτοιο χαρτοφυλάκιο.

Η μέθοδος της λογιστικής παλινδρόμησης είναι ένας ικανός τρόπος για την προσέγγιση της πιθανότητας αθέτησης, όμως δε δύναται να είναι αποτελεσματική σε μια τέτοια περίπτωση, καθώς ο ελάχιστος αριθμός αθετήσεων καθιστά ανέφικτη την ικανοποιητική εκτίμηση των παραμέτρων β_i και των διαστημάτων εμπιστοσύνης για τις παραμέτρους αυτές, και ως εκ τούτου την εκτίμηση της πιθανότητας αθέτησης.

Το δένδρο ταξινόμησης δεν αποτελεί μία αξιόπιστη μέθοδο εκτίμησης της πιθανότητας αθέτησης. Η φύση των δεδομένων είναι τέτοια που κάνει σχεδόν ανέφικτη την επιλογή των μεταβλητών και των σημείων αποκοπής αυτών, ώστε να γίνει ο απαραίτητος διαχωρισμός. Επομένως, είναι σαφές ότι ενδέχεται να καταλήξουμε σε ένα δένδρο, το οποίο δε θα μπορεί να εκτιμήσει την πραγματική πιθανότητα αθέτησης ενός δανειολήπτη.

Επίσης, μέθοδοι όπως η προσομοίωση Monte Carlo, που χρησιμοποιείται σε αρκετές περιπτώσεις για να προσομοιώσουμε δεδομένα, αλλά και ο ιστορικός μέσος δεν μπορούν να μας παρέχουν με βεβαιότητα, αξιόπιστα συμπεράσματα.

Τέλος, στα μοντέλα διαχωριστικής ανάλυσης, όπως το μοντέλο του Altman, υπάρχουν προβλήματα που έχουν να κάνουν πάλι με τα δεδομένα. Στόχος των μοντέλων διαχωριστικής ανάλυσης είναι η διαμόρφωση ενός κανόνα απόφασης, προκειμένου ο εκάστοτε υποψήφιος δανειολήπτης να κατατάσσεται είτε ως “καλός” είτε ως “κακός”. Για να επιτευχθεί αυτό, το ΠII βασίζεται στα ιστορικά δεδομένα για

να καθορίσει την κατανομή των “καλών” και “κακών” πελατών, το κόστος της λανθασμένης ταξινόμησης και τα ποσοστά των “καλών” και “κακών” πελατών στο σύνολο του χαρτοφυλακίου.

Ωστόσο, στα χαρτοφυλάκια υψηλής φερεγγυότητας είναι αναμενόμενο να υπάρχουν ελάχιστες αθετήσεις, με αποτέλεσμα να μην μπορεί να καθοριστεί με επαρκή βεβαιότητα η κατανομή των “κακών” πελατών. Στην περίπτωση του μηδενικού αριθμού αθετήσεων στα δεδομένα, η κατανομή των “κακών” δεν μπορεί να εκτιμηθεί καθόλου, ενώ πρόβλημα αποτελεί και το γεγονός ότι όλοι οι πελάτες θεωρούνται “καλοί”.

Άμεση συνέπεια αυτών είναι να μην μπορεί να καθοριστεί ένας κανόνας απόφασης στη συντριπτική πλειοψηφία των περιπτώσεων. Ακόμα όμως και αν το ΠΙ μπορούσε να κατατάξει τους πελάτες στους δύο πληθυσμούς με ικανοποιητικό τρόπο, λόγω της φύσης της μεθόδου δεν θα μπορούσε να εκτιμήσει εκ των προτέρων την πιθανότητα αθέτησης.

Δίχως εκτίμηση για την πιθανότητα αθέτησης δεν θα υπήρχε κανένα μέτρο για να καθορίσει την έκθεση του χαρτοφυλακίου σε κίνδυνο, άρα και κανένα μέτρο για τα απαιτούμενα κεφάλαια που θα πρέπει να κρατήσει το ΠΙ. Όπως αναφέρθηκε, τόσο η διατήρηση υπερβολικά πολλών κεφαλαίων όσο και η διατήρηση λίγων αποτελεί πρόβλημα για το ΠΙ.

2. Η πιο συντηρητική αρχή εκτίμησης

Η πιο ευρέως χρησιμοποιούμενη προσέγγιση για την αντιμετώπιση του κινδύνου υποεκτίμησης της πιθανότητας αθέτησης σε χαρτοφυλάκια υψηλής φερεγγυότητας προτάθηκε από τους Pluto and Tasche (2005) και είναι γνωστή ως η πιο συντηρητική αρχή εκτίμησης (most prudent estimation).

Σύμφωνα με το μοντέλο των Pluto and Tasche (2005), στην αρχή της περιόδου παρατήρησης υπάρχουν στο χαρτοφυλάκιο n οφειλέτες οι οποίοι έχουν την ίδια πιθανότητα αθέτησης $0 < p < 1$. Στο τέλος της περιόδου παρατήρησης έχουν υπάρξει $0 \leq d \leq n$ αθετήσεις.

Η ιδέα που διέπει το μοντέλο είναι να γίνει ένας στατιστικός έλεγχος υποθέσεων με χρήση των ιστορικών δεδομένων, πιο συγκεκριμένα ο εξής:

$$H_0: p \geq p_0 \text{ έναντι της } H_1: p < p_0 .$$

Η τιμή του p_0 είναι ουσιαστικά η μέγιστη τιμή για την οποία η μηδενική υπόθεση H_0 δεν μπορεί να απορριφθεί. Για να υπολογιστεί η τιμή του p_0 σε προκαθορισμένο σφάλμα τύπου 1 ($0 < a < 1$) θέτουμε $P(D \leq d | H_0) \geq a$ και λύνοντας την ανισότητα λαμβάνουμε μία εκτίμηση, έστω p_0^* , για την πιθανότητα αθέτησης του χαρτοφυλακίου. Με αυτό τον τρόπο διασφαλίζεται ότι για οποιαδήποτε τιμή μικρότερη του p_0^* , η μηδενική υπόθεση δεν μπορεί να απορριφθεί. Για την πιθανότητα αθέτησης λαμβάνουμε με τον τρόπο αυτό ένα διάστημα της μορφής $[0, p_0^*]$.

Βασιζόμενοι στον παραπάνω έλεγχο υπόθεσης και λαμβάνοντας ως περίοδο παρατήρησης το τελευταίο έτος, εξετάζουμε ξεχωριστά τις εξής περιπτώσεις για τα χαρτοφυλάκια υψηλής φερεγγυότητας.

2.1 Καμία αθέτηση και ανεξαρτησία των παρατηρήσεων του δείγματος

Στην περίπτωση αυτή, οι υποθέσεις που κάνουμε είναι οι κάτωθι:

- α) Οι οφειλέτες του χαρτοφυλακίου κατατάσσονται σε τρεις βαθμούς διαβάθμισης A, B, Γ και ο αριθμός των οφειλετών σε κάθε βαθμίδα είναι αντίστοιχα n_A, n_B, n_Γ .
- β) Υπάρχει ανεξαρτησία μεταξύ των αθετήσεων.
- γ) Κατά την τελευταία περίοδο παρατήρησης δε σημειώθηκε καμία αθέτηση σε καμία από τις τρεις βαθμίδες.
- δ) Η βαθμίδα A είναι αυτή με την υψηλότερη πιστοληπτική ικανότητα και η Γ αυτή με τη μικρότερη.

Αν p_A, p_B και p_Γ είναι οι πιθανότητες αθέτησης που πρέπει να εκτιμηθούν για την κάθε βαθμίδα, τότε λόγω της υπόθεσης (δ) θα έχουμε:

$$p_A \leq p_B \leq p_\Gamma . \quad (2.1)$$

Προφανώς, για να εκτιμήσουμε την πιο συντηρητική τιμή της p_A , αρκεί να θεωρήσουμε ότι ισχύει $p_A = p_\Gamma$. Σε αυτή την περίπτωση από την (2.1) παίρνουμε:

$$p_A = p_B = p_\Gamma . \quad (2.2)$$

Λόγω της υπόθεσης (β), έχουμε πλέον ένα ομοιογενές χαρτοφυλάκιο με πιθανότητες αθέτησης $p = p_A = p_B = p_\Gamma$ μεγέθους $n_A + n_B + n_\Gamma = n$.

Για τον κάθε οφειλέτη υπάρχουν δύο σενάρια, αυτό της κανονικής αποπληρωμής και αυτό της αθέτησης. Βάσει των παραπάνω και της υπόθεσης (β) προκύπτει άμεσα ότι η τυχαία μεταβλητή που παριστά τον αριθμό των αθετήσεων, ακολουθεί διωνυμική κατανομή με συνάρτηση πιθανότητας

$$p_X(d) = P(X = d) = \binom{n}{d} p_A^d (1 - p_A)^{n-d} . \quad (2.3)$$

Από τη σχέση (2.3) προκύπτει ότι η πιθανότητα να μην έχουμε καμία αθέτηση κατά την περίοδο παρατήρησης είναι $(1 - p_A)^n$. Εφόσον λοιπόν, θέλουμε να εκτιμήσουμε ένα διάστημα εμπιστοσύνης με συντελεστή εμπιστοσύνης γ για την p_A , αρκεί να λύσουμε την παρακάτω ανισότητα

$$\begin{aligned} 1 - \gamma &\leq (1 - p_A)^n \\ \Leftrightarrow p_A &\leq 1 - (1 - \gamma)^{1/n} . \end{aligned} \quad (2.4)$$

Για την εκτίμηση του διαστήματος εμπιστοσύνης για το p_B θα ακολουθήσουμε παρόμοια διαδικασία, αλλά επειδή το ζητούμενο είναι η πιο συντηρητική εκτίμηση και το p_A αποτελεί κάτω όριο της p_B , πλέον αντί της (2.2) έχουμε την σχέση :

$$p_B = p_\Gamma . \quad (2.5)$$

Άμεσο επακόλουθο της σχέσης (2.5) είναι ότι ως μέγεθος του δείγματος θα χρησιμοποιήσουμε το $n_B + n_\Gamma$ και γράφοντας μια σχέση παρόμοια με την (2.4) λαμβάνουμε

$$1 - \gamma \leq (1 - p_B)^{n_B + n_\Gamma} \Leftrightarrow p_B \leq 1 - (1 - \gamma)^{1/n_B + n_\Gamma} . \quad (2.6)$$

Ομοίως για το p_Γ θα έχουμε :

$$(1 - \gamma) \leq (1 - p_\Gamma)^{n_\Gamma} \Leftrightarrow p_\Gamma \leq 1 - (1 - \gamma)^{1/n_\Gamma} . \quad (2.7)$$

Επομένως για τα p_A, p_B, p_Γ τα άνω όρια των διαστημάτων εμπιστοσύνης, με επίπεδο σημαντικότητας $1 - \gamma$ θα δίνονται από τους τύπους

$$1 - (1 - \gamma)^{1/n_A + n_B + n_\Gamma} \leq 1 - (1 - \gamma)^{1/n_B + n_\Gamma} \leq 1 - (1 - \gamma)^{1/n_\Gamma} \quad (2.8)$$

αντίστοιχα.

Για τους πίνακες και τα γραφήματα που παρατίθενται στη συνέχεια χρησιμοποιήθηκε ο τύπος (2.8) και τα εξής χαρτοφυλάκια με μηδενικό αριθμό αθετήσεων:

| Χαρτοφυλάκιο | Οφειλέτες Α βαθμίδας | Οφειλέτες Β βαθμίδας | Οφειλέτες Γ βαθμίδας |
|--------------|-------------------------|-------------------------|-------------------------|
| 1 | 100 | 250 | 150 |
| 2 | 200 | 170 | 230 |
| 3 | 300 | 100 | 250 |
| 4 | 150 | 350 | 200 |
| 5 | 50 | 400 | 300 |
| 6 | 250 | 300 | 450 |

Οι τιμές του γ που χρησιμοποιήθηκαν είναι

$$\gamma = 50\% , 75\% , 90\% , 95\% , 99\% , 99.9\%$$

Στους ακόλουθους πίνακες δίνονται οι εκτιμήσεις των p_A, p_B, p_Γ για κάθε χαρτοφυλάκιο και για όλες τις επιλογές του γ . Γραφική απεικόνιση της p_A για κάθε χαρτοφυλάκιο.

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.14% | 0.28% | 0.46% | 0.60% | 0.92% | 1.37% |
| p_B | 0.17% | 0.35% | 0.57% | 0.75% | 1.14% | 1.71% |
| p_Γ | 0.46% | 0.92% | 1.52% | 1.98% | 3.02% | 4.50% |

Πίνακας 2.1 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 1 χωρίς αθετήσεις

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.12% | 0.23% | 0.38% | 0.50% | 0.76% | 1.14% |
| p_B | 0.17% | 0.35% | 0.57% | 0.75% | 1.14% | 1.71% |
| p_Γ | 0.30% | 0.60% | 1% | 1.29% | 1.98% | 2.96% |

Πίνακας 2.2 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 2 χωρίς αθετήσεις

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.11% | 0.21% | 0.35% | 0.46% | 0.71% | 1.06% |
| p_B | 0.20% | 0.40% | 0.66% | 0.85% | 1.31% | 1.95% |
| p_Γ | 0.28% | 0.55% | 0.92% | 1.19% | 1.83% | 2.73% |

Πίνακας 2.3 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 3 χωρίς αθετήσεις

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.10% | 0.20% | 0.33% | 0.43% | 0.66% | 0.98% |
| p_B | 0.13% | 0.25% | 0.42% | 0.54% | 0.83% | 1.25% |
| p_Γ | 0.35% | 0.69% | 1.14% | 1.49% | 2.28% | 3.39% |

Πίνακας 2.4 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 4 χωρίς αθετήσεις

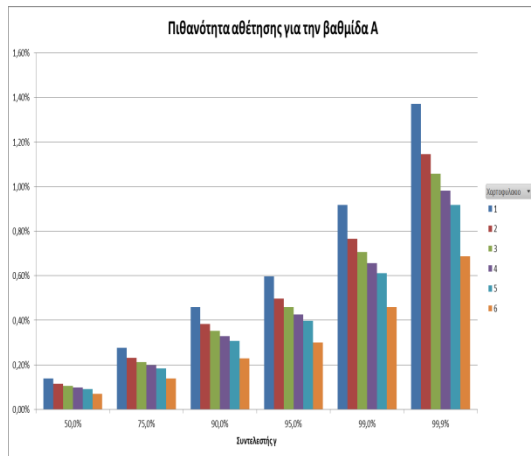
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.09% | 0.19% | 0.31% | 0.40% | 0.61% | 0.92% |
| p_B | 0.10% | 0.20% | 0.33% | 0.43% | 0.66% | 0.98% |
| p_Γ | 0.23% | 0.46% | 0.76% | 0.99% | 1.52% | 2.28% |

Πίνακας 2.5 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 5 χωρίς αθετήσεις

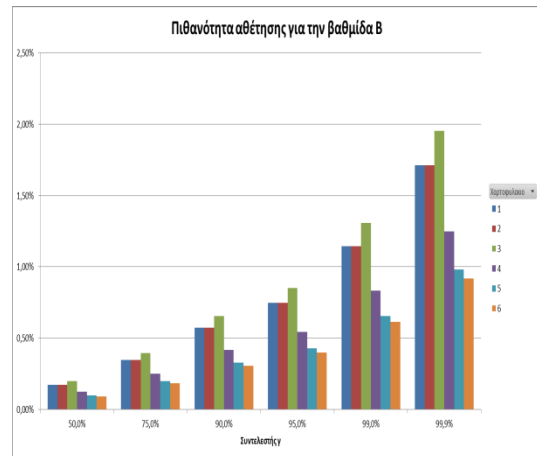
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.07% | 0.14% | 0.23% | 0.30% | 0.46% | 0.69% |
| p_B | 0.09% | 0.19% | 0.31% | 0.40% | 0.61% | 0.92% |
| p_Γ | 0.15% | 0.31% | 0.51% | 0.66% | 1.02% | 1.52% |

Πίνακας 2.6 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 6 χωρίς αθετήσεις

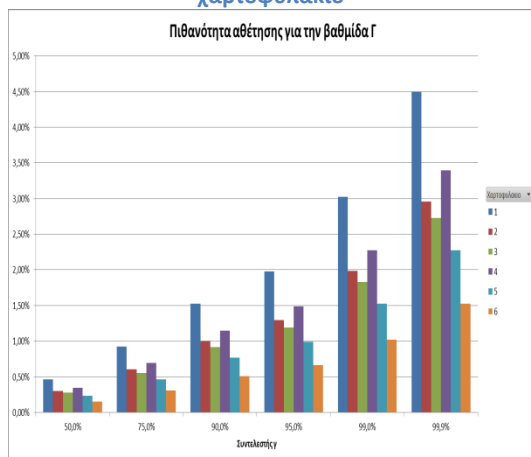
Στα επόμενα σχήματα φαίνονται ξεχωριστά για κάθε βαθμίδα οι εκτιμήσεις για την πιθανότητα αθέτησης για όλα τα χαρτοφυλάκια.



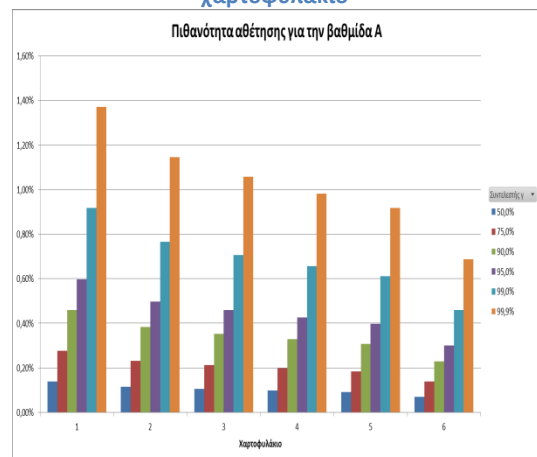
Σχήμα 2.1 Γραφική απεικόνιση της p_A για κάθε χαρτοφυλάκιο



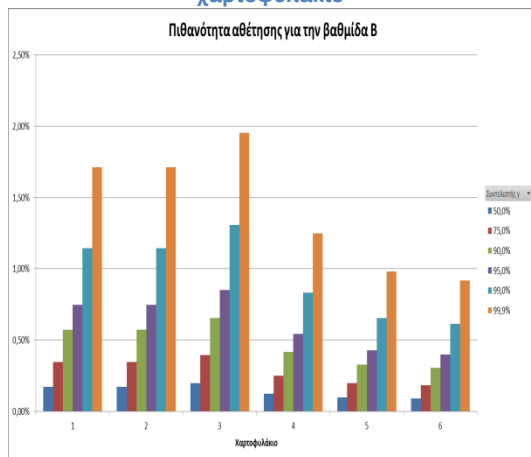
Σχήμα 2.2 Γραφική απεικόνιση της p_B για κάθε χαρτοφυλάκιο



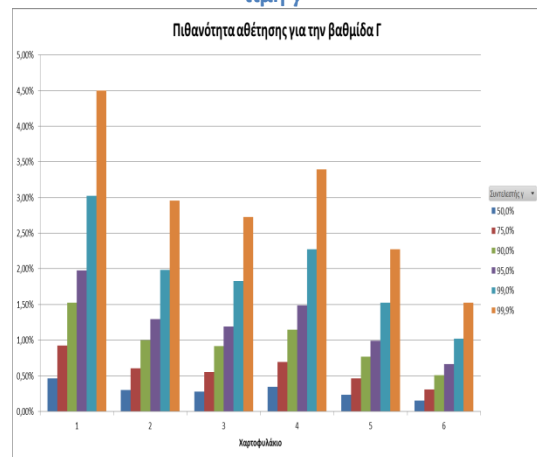
Σχήμα 2.3 Γραφική απεικόνιση της p_Γ για κάθε χαρτοφυλάκιο



Σχήμα 2.4 Γραφική απεικόνιση της p_A για κάθε τιμή γ



Σχήμα 2.5 Γραφική απεικόνιση της p_B για κάθε τιμή γ



Σχήμα 2.6 Γραφική απεικόνιση της p_Γ για κάθε τιμή γ

Όπως ήταν αναμενόμενο τόσο μέσω των πινάκων όσο και από τα γραφήματα 2.1, 2.2, 2.3 φαίνεται πως υπάρχει αρνητική σχέση μεταξύ του αριθμού των οφειλετών που υπάρχουν σε κάθε βαθμίδα και των αντίστοιχων πιθανοτήτων αθέτησης. Από τα γραφήματα 2.4, 2.5, 2.6 μπορούμε εύκολα επίσης, να επισημάνουμε πως για μικρές τιμές του συντελεστή γ οι διαφορές στις εκτιμώμενες πιθανότητες στην εκάστοτε βαθμίδα είναι μικρές, για μεγάλες ωστόσο τιμές του γ υπάρχει σαφής διαφορά.

Όπως γίνεται εύκολα αντιληπτό από τον τύπο (2.8) για κάθε σταθερή τιμή του συντελεστή γ , η εκτίμηση των πιθανοτήτων εξαρτάται από τις συχνότητες n_A, n_B, n_G . Στους παρακάτω πίνακες θα παραθέσουμε τους λόγους $\frac{p_A}{p_B}, \frac{p_B}{p_G}, \frac{n_B+n_G}{n_A+n_B+n_G}$ και $\frac{n_G}{n_B+n_G}$ για όλα τα χαρτοφυλάκια.

| Χαρτοφυλάκιο | $\gamma=50\%$ | $\gamma=75\%$ | $\gamma=90\%$ | $\gamma=95\%$ | $\gamma=99\%$ | $\gamma=99.9\%$ |
|--------------|---------------|---------------|---------------|---------------|---------------|-----------------|
| 1 | 80.01% | 80.03% | 80.05% | 80.06% | 80.09% | 80.14% |
| 2 | 66.69% | 66.71% | 66.73% | 66.75% | 66.79% | 66.86% |
| 3 | 53.87% | 53.90% | 53.93% | 53.95% | 54.01% | 54.09% |
| 4 | 78.58% | 78.59% | 78.61% | 78.62% | 78.64% | 78.68% |
| 5 | 93.34% | 93.34% | 93.34% | 93.35% | 93.35% | 93.36% |
| 6 | 75.01% | 75.02% | 75.03% | 75.04% | 75.06% | 75.09% |

Πίνακας 2.7 Ο λόγος p_A/p_B για όλα τα χαρτοφυλάκια

| Χαρτοφυλάκιο | $\gamma=50\%$ | $\gamma=75\%$ | $\gamma=90\%$ | $\gamma=95\%$ | $\gamma=99\%$ | $\gamma=99.9\%$ |
|--------------|---------------|---------------|---------------|---------------|---------------|-----------------|
| 1 | 37.55% | 37.61% | 37.68% | 37.73% | 37.86% | 38.04% |
| 2 | 57.54% | 57.57% | 57.62% | 57.66% | 57.74% | 57.87% |
| 3 | 71.46% | 71.49% | 71.52% | 71.55% | 71.62% | 71.71% |
| 4 | 36.40% | 36.44% | 36.50% | 36.54% | 36.63% | 36.76% |
| 5 | 42.89% | 42.91% | 42.95% | 42.98% | 43.05% | 43.14% |
| 6 | 60.02% | 60.04% | 60.06% | 60.08% | 60.12% | 60.18% |

Πίνακας 2.8 Ο λόγος p_B/p_G για όλα τα χαρτοφυλάκια

| Χαρτοφυλάκιο | $\frac{n_B + n_G}{n_A + n_B + n_G}$ | $\frac{n_G}{n_B + n_G}$ |
|--------------|-------------------------------------|-------------------------|
| | 1 | 80% |
| 2 | 66.67% | 57.50% |
| 3 | 53.85% | 71.43% |
| 4 | 78.57% | 36.36% |
| 5 | 93.33% | 42.86% |
| 6 | 75% | 60% |

Πίνακας 2.9 Οι λόγοι $\frac{n_B+n_G}{n_A+n_B+n_G}$ και $\frac{n_G}{n_B+n_G}$ για όλα τα χαρτοφυλάκια

Από τους παραπάνω πίνακες μπορούμε να δούμε ότι ο λόγος των πιθανοτήτων είναι σχεδόν σταθερός, ενώ παρατηρούμε ότι ο λόγος $\frac{p_A}{p_B}$ είναι σχεδόν ίσος με τον λόγο $\frac{n_B+n_G}{n_A+n_B+n_G}$ και ο λόγος $\frac{n_G}{n_B+n_G}$ προσεγγίζει αρκετά καλά τον λόγο $\frac{p_B}{p_G}$. Παρακάτω θα παραθέσουμε τον λόγο για τον οποίο ισχύει αυτό.

Γνωρίζουμε ότι,

$$\frac{p_A}{p_B} = \frac{1 - (1 - \gamma)^{1/n_A+n_B+n_\Gamma}}{1 - (1 - \gamma)^{1/n_B+n_\Gamma}}$$

και

$$(1 - \gamma)^\alpha = \sum_{\kappa=0}^{\infty} \binom{\alpha}{\kappa} (-\gamma)^\kappa = 1 + \alpha(-\gamma) + \binom{\alpha}{2} \gamma^2 + \dots$$

Από την παραπάνω σχέση για μικρές τιμές των α, γ ισχύει ότι,

$$1 - (1 - \gamma)^\alpha \cong \alpha \gamma .$$

Τέλος για τον λόγο των πιθανοτήτων έχουμε

$$\frac{p_A}{p_B} = \frac{1 - (1 - \gamma)^{1/n_A+n_B+n_\Gamma}}{1 - (1 - \gamma)^{1/n_B+n_\Gamma}} \approx \frac{\frac{\gamma}{n_A + n_B + n_\Gamma}}{\frac{\gamma}{n_B + n_\Gamma}} = \frac{n_B + n_\Gamma}{n_A + n_B + n_\Gamma} .$$

Αντίστοιχες σχέσεις ισχύουν για τους λόγους $\frac{p_B}{p_\Gamma}$ και $\frac{p_A}{p_\Gamma}$. Πρακτικά λοιπόν, για ένα χαρτοφυλάκιο με μηδενικό αριθμό αθετήσεων, αν με κάποιο τρόπο γνωρίζουμε την πιθανότητα αθέτησης μιας βαθμίδας τότε μια εκτίμηση για την πιθανότητα αθέτησης των άλλων βαθμίδων μπορεί εύκολα να προκύψει από τους λόγους των συχνοτήτων.

2.2 Μικρός αριθμός αθετήσεων και ανεξαρτησία των παρατηρήσεων του δείγματος

Ένα συνηθέστερο σενάριο ακόμα και για ένα χαρτοφυλάκιο υψηλής φερεγγυότητας είναι αυτό που περιγράφεται με τις παρακάτω υποθέσεις :

- α) Οι οφειλέτες του χαρτοφυλακίου κατατάσσονται σε τρεις βαθμούς διαβάθμισης A, B, Γ και n_A, n_B, n_Γ αντίστοιχα είναι ο αριθμός των οφειλετών σε κάθε βαθμίδα.
- β) Υπάρχει ανεξαρτησία μεταξύ των αθετήσεων.
- γ) Κατά την τελευταία περίοδο παρατήρησης σημειώθηκαν a, b, c αθετήσεις στις βαθμίδες A, B, Γ αντίστοιχα, με $a + b + c = d$
- δ) Η βαθμίδα A είναι αυτή με την υψηλότερη πιστοληπτική ικανότητα και η Γ αυτή με την μικρότερη.

Στόχος μας είναι πάλι να καθορίσουμε διαστήματα εμπιστοσύνης για τα p_A, p_B και p_Γ . Προκειμένου να εκτιμήσουμε την πιο συντηρητική τιμή για το p_A υποθέτουμε ότι $p_A = p_B = p_\Gamma$.

Τότε, μπορούμε και πάλι να αντιμετωπίσουμε το χαρτοφυλάκιο ως ομοιογενές δείγμα μεγέθους $n_A + n_B + n_\Gamma$. Επίσης, ισχύει η σχέση (2.3) και η πιθανότητα να παρατηρήσουμε το πολύ d αθετήσεις είναι:

$$\sum_{i=0}^d \binom{n_A + n_B + n_\Gamma}{i} p_A^i (1 - p_A)^{n_A + n_B + n_\Gamma - i} . \quad (2.8)$$

Για να καθορίσουμε λοιπόν διάστημα εμπιστοσύνης για την p_A , αρκεί να λύσουμε την εξής ανισότητα:

$$1 - \gamma \leq \sum_{i=0}^d \binom{n_A + n_B + n_\Gamma}{i} p_A^i (1 - p_A)^{n_A + n_B + n_\Gamma - i} . \quad (2.9)$$

Το δεξί μέλος της ανισότητας είναι η πιθανότητα να παρατηρηθούν το πολύ d αθετήσεις.

Ομοίως για τις p_B, p_Γ πρέπει να λύσουμε αντίστοιχα τις παρακάτω:

$$1 - \gamma \leq \sum_{i=0}^{b+c} \binom{n_B + n_\Gamma}{i} p_B^i (1 - p_B)^{n_B + n_\Gamma - i} \quad (2.10)$$

$$1 - \gamma \leq \sum_{i=0}^c \binom{n_\Gamma}{i} p_\Gamma^i (1 - p_\Gamma)^{n_\Gamma - i} . \quad (2.11)$$

Στα παραδείγματα που ακολουθούν χρησιμοποιήθηκαν τα χαρτοφυλάκια της Ενότητας 2.1. Ο αριθμός των αθετήσεων που θεωρήσαμε για κάθε βαθμίδα στο εκάστοτε χαρτοφυλάκιο δίνεται στον παρακάτω πίνακα.

| Χαρτοφυλάκιο | Αθετήσεις στην βαθμίδα A | Αθετήσεις στην βαθμίδα B | Αθετήσεις στην βαθμίδα Γ | Συνολικές αθετήσεις |
|--------------|----------------------------------|----------------------------------|---------------------------------------|------------------------|
| 1 | 0 | 1 | 1 | 2 |
| 2 | 1 | 0 | 0 | 1 |
| 3 | 1 | 1 | 0 | 2 |
| 4 | 0 | 1 | 2 | 3 |
| 5 | 0 | 2 | 0 | 2 |
| 6 | 0 | 0 | 2 | 2 |

Οι εκτιμήσεις για τις p_A, p_B, p_Γ για κάθε χαρτοφυλάκιο δίνονται στους παρακάτω πίνακες.

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.53% | 0.78% | 1.06% | 1.25% | 1.67% | 2.22% |
| p_B | 0.66% | 0.98% | 1.32% | 1.56% | 2.08% | 2.77% |
| p_Γ | 1.1% | 1.78% | 2.57% | 3.12% | 4.34% | 5.98% |

Πίνακας 2.10 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 1 όταν έχουν συμβεί αθετήσεις

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.28% | 0.45% | 0.65% | 0.79% | 1.10% | 1.53% |
| p_B | 0.17% | 0.35% | 0.57% | 0.75% | 1.14% | 1.71% |
| p_Γ | 0.30% | 0.60% | 1% | 1.29% | 1.98% | 2.96% |

Πίνακας 2.11 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 2 όταν έχουν συμβεί αθετήσεις

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.41% | 0.60% | 0.82% | 0.97% | 1.29% | 1.72% |
| p_B | 0.48% | 0.77% | 1.11% | 1.35% | 1.88% | 2.61% |
| p_Γ | 0.28% | 0.55% | 0.92% | 1.19% | 1.83% | 2.73% |

Πίνακας 2.12 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 3 όταν έχουν συμβεί αθετήσεις

| | | | | | | |
|------------|-------|-------|-------|-------|-------|-------|
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
| p_A | 0.52% | 0.73% | 0.95% | 1.10% | 1.43% | 1.85% |
| p_B | 0.67% | 0.93% | 1.21% | 1.40% | 1.81% | 2.35% |
| p_Γ | 1.33% | 1.95% | 2.64% | 3.11% | 4.13% | 5.48% |

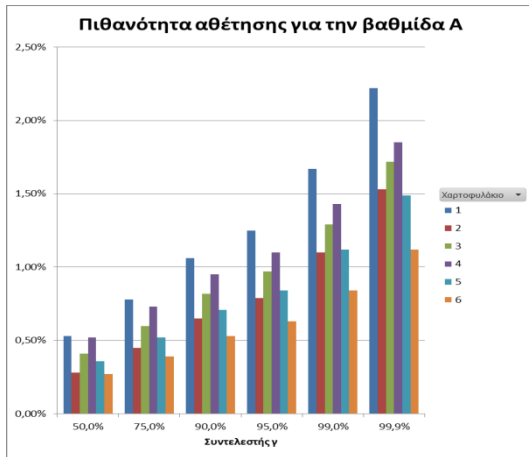
Πίνακας 2.13 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 4 όταν έχουν συμβεί αθετήσεις

| | | | | | | |
|------------|-------|-------|-------|-------|-------|-------|
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
| p_A | 0.36% | 0.52% | 0.71% | 0.84% | 1.12% | 1.49% |
| p_B | 0.38% | 0.56% | 0.76% | 0.90% | 1.20% | 1.59% |
| p_Γ | 0.23% | 0.46% | 0.76% | 0.99% | 1.52% | 2.28% |

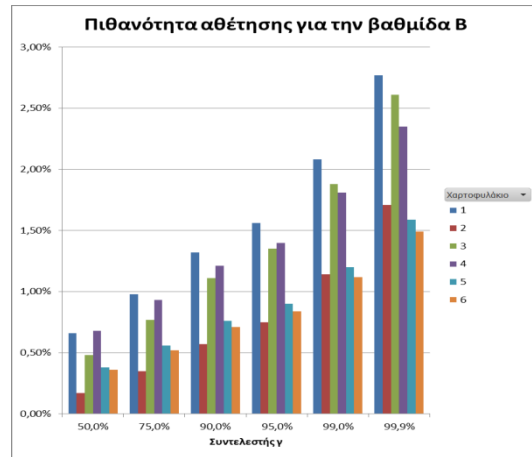
Πίνακας 2.14 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 5 όταν έχουν συμβεί αθετήσεις

| | | | | | | |
|------------|-------|-------|-------|-------|-------|-------|
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
| p_A | 0.27% | 0.39% | 0.53% | 0.63% | 0.84% | 1.12% |
| p_B | 0.36% | 0.52% | 0.71% | 0.84% | 1.12% | 1.49% |
| p_Γ | 0.59% | 0.87% | 1.18% | 1.39% | 1.86% | 2.47% |

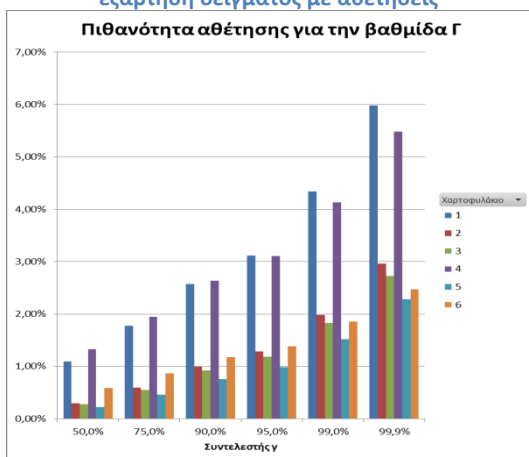
Πίνακας 2.15 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 6 όταν έχουν συμβεί αθετήσεις



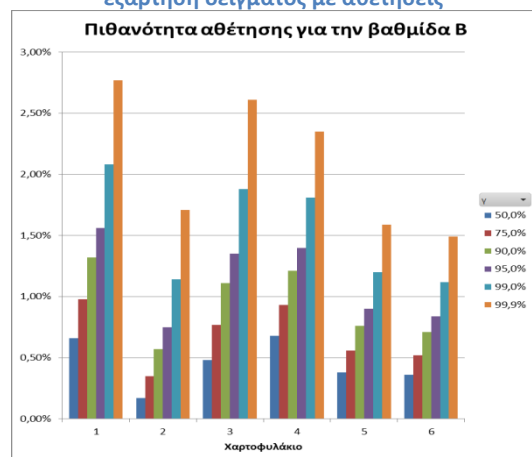
Σχήμα 2.7 Γραφική απεικόνιση της p_A χωρίς εξάρτηση δείγματος με αθετήσεις



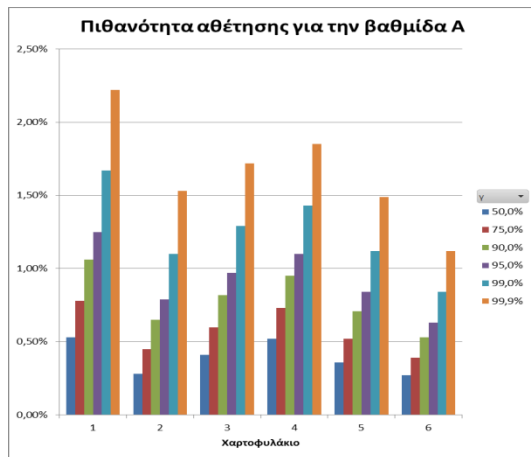
Σχήμα 2.8 Γραφική απεικόνιση της p_B χωρίς εξάρτηση δείγματος με αθετήσεις



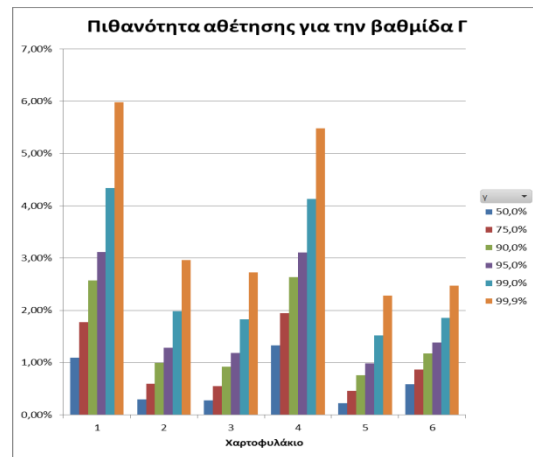
Σχήμα 2.9 Γραφική απεικόνιση της p_Γ χωρίς εξάρτηση δείγματος με αθετήσεις



Σχήμα 2.10 Γραφική απεικόνιση της p_B για κάθε τιμή του γ με αθετήσεις



Σχήμα 2.11 Γραφική απεικόνιση της p_A για κάθε τιμή του γ με αθετήσεις



Σχήμα 2.12 Γραφική απεικόνιση της p_G για κάθε τιμή του γ με αθετήσεις

Βάσει των παραπάνω πινάκων και των γραφημάτων 2.10, 2.11 και 2.12 παρατηρούμε ότι η επιλογή του συντελεστή γ είναι μείζονος σημασίας για την εκτίμηση της πιθανότητας αθέτησης, όπως είχε παρατηρηθεί και στην Ενότητα 2.1. Από τα γραφήματα 2.7, 2.8 και 2.9 γίνεται σαφές ότι ο αριθμός των αθετήσεων είναι ο κύριος παράγοντας που καθορίζει την εκτίμηση των πιθανοτήτων. Ωστόσο με δεδομένο αριθμό οφειλετών στο χαρτοφυλάκιο διασφαλίζεται και η μονοτονία. Χαρακτηριστικά, συγκρίνοντας τα αποτελέσματα των χαρτοφυλακίων 5 και 6 βλέπουμε τα εξής:

Βάσει των αποτελεσμάτων στη βαθμίδα Α, παρατηρούμε ότι οι εκτιμώμενες πιθανότητες για το χαρτοφυλάκιο 6 είναι μικρότερες από του χαρτοφυλακίου 2 για όλες τις τιμές του γ . Το συγκεκριμένο αποτέλεσμα ήταν αναμενόμενο, καθώς για τα δύο χαρτοφυλάκια έχουμε τον ίδιο αριθμό αθετήσεων και το χαρτοφυλάκιο 6 αποτελείται από περισσότερους οφειλότες. Συγκρίνοντας τα αποτελέσματα για τη βαθμίδα Β όμως, παρατηρούμε μια ιδιαιτερότητα. Παρότι στο χαρτοφυλάκιο 5 έχουν σημειωθεί δύο αθετήσεις για την εν λόγω βαθμίδα και στο χαρτοφυλάκιο 6 δεν έχουμε αθετήσεις, οι εκτιμώμενες πιθανότητες για όλες τις τιμές του συντελεστή γ είναι σχεδόν ίδιες. Η ιδιαιτερότητα αυτή οφείλεται στο γεγονός ότι για τον υπολογισμό της πιθανότητας αθέτησης σε κάθε βαθμίδα δε χρησιμοποιούμε τον αριθμό των αθετήσεων που έχουν συμβεί μόνο στην εκάστοτε βαθμίδα, αλλά τον αριθμό των αθετήσεων που έχουν συμβεί έως αυτήν την βαθμίδα.

Αξίζει επίσης να αναφερθούμε στα αποτελέσματα του χαρτοφυλακίου 3. Στο εν λόγω χαρτοφυλάκιο έχει σημειωθεί από μία αθέτηση στις βαθμίδες Α και Β και καμία στη βαθμίδα Γ. Αυτή η ιδιαιτερότητα έχει ως αποτέλεσμα για μικρές τιμές του συντελεστή γ , η πιθανότητα αθέτησης για τη βαθμίδα Γ να είναι μικρότερη των άλλων 2. Παρότι αυτή η σχέση μεταξύ των βαθμίδων Α και Γ αντιστρέφεται, όταν το γ παίρνει τιμές μεγαλύτερες του 90% δεν ισχύει το ίδιο και για τη δεύτερη βαθμίδα. Μία λογική εξήγηση σε αυτήν την περίπτωση είναι πως το σύστημα αξιολόγησης του πιστωτικού ιδρύματος δεν αποδίδει μια σωστή κατάταξη των οφειλετών.

2.3 Καμία αθέτηση και εξάρτηση των παρατηρήσεων του δείγματος

Η υπόθεση της ανεξαρτησίας μεταξύ των αθετήσεων ενός χαρτοφυλακίου δεν είναι ρεαλιστική σε πραγματικά δεδομένα. Για να γίνει αυτό πιο σαφές, αρκεί να σκεφτούμε το σενάριο μιας οικονομικής ύφεσης, όπως αυτή που έζησε τα τελευταία χρόνια η χώρα μας. Σε μία τέτοια περίπτωση και καθώς η ανεργία αυξάνεται, οι εργαζόμενοι χάνουν τις δουλειές τους και προφανώς κάποιοι από αυτούς δεν θα είναι σε θέση να αποπληρώσουν τα δάνειά τους. Εκτός όμως από τους απλούς πολίτες, η οικονομική κατάσταση θα επηρεάσει και τις επιχειρήσεις με παρόμοια αποτελέσματα.

Για τον λόγο αυτό, είναι απαραίτητο να εισαχθεί στο μοντέλο ένας επιπλέον παράγοντας X ο οποίος θα μας δώσει τη δυνατότητα να δημιουργήσουμε μια μορφή εξάρτησης. Συγκεκριμένα, οι Pluto and Tasche (2005) μελέτησαν το εξής χαρτοφυλάκιο:

Στην αρχή της περιόδου παρατήρησης υπάρχουν n οφειλέτες στο χαρτοφυλάκιο. Όλοι οι οφειλέτες του χαρτοφυλακίου έχουν την ίδια πιθανότητα αθέτησης $0 < p < 1$. Στο τέλος της περιόδου παρατήρησης έχουν υπάρξει $0 \leq d \leq n$ αθετήσεις.

Θεωρούμε ότι ένας δανειολήπτης αθετεί τις υποχρεώσεις του, αν η αξία των περιουσιακών του στοιχείων (assets) γίνει μικρότερη ενός συγκεκριμένου ορίου. Η αξία των περιουσιακών στοιχείων του εκάστοτε οφειλέτη εξαρτάται από ένα συστηματικό παράγοντα X ο οποίος μπορεί να θεωρηθεί κοινός για όλο το χαρτοφυλάκιο, καθώς αφορά το σύνολο της οικονομίας και έναν ατομικό παράγοντα Y_i , ο οποίος είναι διαφορετικός για κάθε δανειολήπτη $i (i = 1, 2, \dots, n)$.

Θα παραστήσουμε με D_i την αξία των περιουσιακών στοιχείων του i δανειολήπτη. Τα D_i μπορούν να εκφραστούν ως:

$$D_i = \sqrt{\rho}X + Y_i\sqrt{1 - \rho} . \quad (2.12)$$

Το ρ αποτελεί την συσχέτιση των στοιχείων του ενεργητικού. Οι μεταβλητές X και Y_i είναι ανεξάρτητες και υποθέτουμε ότι ακολουθούν την τυπική κανονική κατανομή. Οπότε τα $D_i, i = 1, 2, \dots, n$ είναι n εξαρτημένες τυχαίες μεταβλητές που ακολουθούν την κατανομή $N(0,1)$.

Βάσει αυτού και θεωρώντας ότι ο i οφειλέτης αθετεί της υποχρεώσεις του, όταν η αξία των περιουσιακών του στοιχείων πέσει κάτω από ένα όριο c , η πιθανότητα αθέτησης δίνεται από την σχέση:

$$P[D_i \leq c] = p . \quad (2.13)$$

Από την τελευταία σχέση προκύπτει,

$$c = \Phi^{-1}(p) \quad (2.14)$$

και συνδυάζοντας τις (2.12) και (2.14) η πιθανότητα αθέτησης δοθείσας μιας τιμής x για τον συστηματικό παράγοντα X μπορεί να γραφεί ως:

$$P[D_i \leq \Phi^{-1}(p)] = P\left[Y_i \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right] = \Phi\left(\frac{\Phi^{-1}(p) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right). \quad (2.15)$$

Όπου Φ είναι η αθροιστική συνάρτηση κατανομής της τυπικής κατανομής και Φ^{-1} είναι η αντίστροφη της Φ .

Όπως και στην περίπτωση που αναλύθηκε στην Ενότητα 2.1, θεωρούμε αρχικά ότι υπάρχουν 3 βαθμίδες μέσα στο χαρτοφυλάκιο και καμία αθέτηση κατά την περίοδο παρατήρησης. Υπό τις υποθέσεις που περιγράφηκαν παραπάνω, η πιθανότητα αθέτησης για τη βαθμίδα A του χαρτοφυλακίου εκτιμάται λύνοντας την παρακάτω ανίσωση:

$$(1 - p_A)^{n_A + n_B + n_\Gamma} \geq 1 - \gamma.$$

Άρα έχοντας μια τιμή για τον συστηματικό παράγοντα X , μπορούμε να εκτιμήσουμε την δεσμευμένη πιθανότητα αθέτησης λύνοντας την ανίσωση:

$$\left(1 - \Phi\left(\frac{\Phi^{-1}(p_A) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right)\right)^{n_A + n_B + n_\Gamma} \geq 1 - \gamma. \quad (2.16)$$

Η εκτίμηση της αδέσμευτης πιθανότητας αθέτησης δίνεται από την παρακάτω σχέση

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(x) (1 - g(\rho, p_A, x))^{n_A + n_B + n_\Gamma} dx \quad (2.17)$$

όπου, $g(\rho, p_A, x) = \Phi\left(\frac{\Phi^{-1}(p_A) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right)$.

Το δεξί μέλος της σχέσης (2.17) είναι η πιθανότητα να μην υπάρξει καμία αθέτηση εντός της τελευταίας περιόδου παρατήρησης.

Αντιστοίχως για τις p_B και p_Γ , αρκεί να λύσουμε τις :

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(x) (1 - g(\rho, p_B, x))^{n_B + n_\Gamma} dx \quad (2.18)$$

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(x) (1 - g(\rho, p_\Gamma, x))^{n_\Gamma} dx \quad (2.19)$$

Από τις παραπάνω εξισώσεις παρατηρούμε ότι με την εισαγωγή του ρ στο μοντέλο μας, η πολυπλοκότητα των αριθμητικών πράξεων έχει αυξηθεί δραματικά. Είναι σημαντικό ωστόσο πως από τις ανισώσεις (2.17), (2.18) και (2.19) μπορούμε να υπολογίσουμε με αριθμητικούς υπολογισμούς τα p_A, p_B, p_Γ .

Ο παρακάτω πίνακας δίνει τις τιμές των p_A, p_B, p_Γ για τα χαρτοφυλάκια της Ενότητας 2.1 για τις τιμές

$$\rho = 18\%$$

$$\gamma = 50\%, 75\%, 90\%, 95\%, 99\%, 99.9\%$$

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|--------|--------|
| p_A | 0.32% | 0.87% | 1.93% | 2.95% | 6.08 % | 12.03% |
| p_B | 0.37% | 1.03% | 2.27% | 3.43% | 6.97% | 13.51% |
| p_{Γ} | 0.87% | 2.21% | 4.56% | 6.65% | 12.50% | 22.20% |

Πίνακας 2.16 Εκτίμηση των p_A, p_B, p_{Γ} για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 1

| Γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|-------|--------|
| p_A | 0.27% | 0.75% | 1.68% | 2.61% | 5.46% | 10.71% |
| p_B | 0.37% | 1.03% | 2.25% | 3.44% | 6.98% | 13.27% |
| p_{Γ} | 0.60% | 1.58% | 3.35% | 4.99% | 9.73% | 17.65% |

Πίνακας 2.17 Εκτίμηση των p_A, p_B, p_{Γ} για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 2

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|-------|--------|
| p_A | 0.25% | 0.71% | 1.59% | 2.46% | 5.16% | 10.28% |
| p_B | 0.42% | 1.15% | 2.48% | 3.75% | 7.51% | 14.23% |
| p_{Γ} | 0.56% | 1.49% | 3.16% | 4.71% | 9.19% | 16.92% |

Πίνακας 2.18 Εκτίμηση των p_A, p_B, p_{Γ} για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 3

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|--------|--------|
| p_A | 0.23% | 0.67% | 1.51% | 2.34% | 4.94% | 10.00% |
| p_B | 0.29% | 0.81% | 1.79% | 2.76% | 5.73% | 11.35% |
| p_{Γ} | 0.67% | 1.77% | 3.70% | 5.48% | 10.53% | 19.12% |

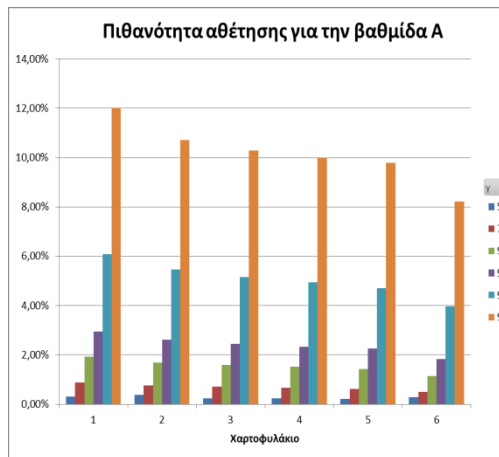
Πίνακας 2.19 Εκτίμηση των p_A, p_B, p_{Γ} για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 4

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|-------|--------|
| p_A | 0.22% | 0.63% | 1.43% | 2.25% | 4.70% | 9.78% |
| p_B | 0.23% | 0.67% | 1.51% | 2.36% | 4.90% | 10.14% |
| p_{Γ} | 0.48% | 1.29% | 2.78% | 4.20% | 8.20% | 15.75% |

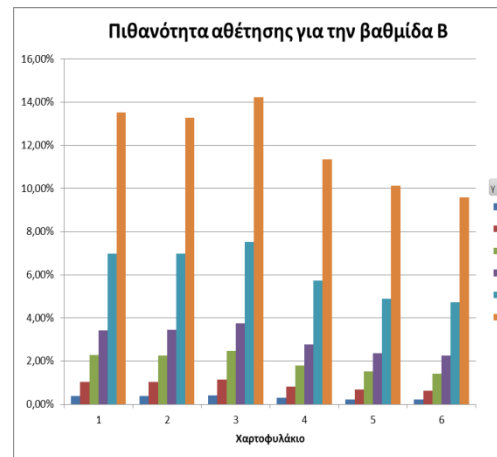
Πίνακας 2.20 Εκτίμηση των p_A, p_B, p_{Γ} για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 5

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|-------|--------|
| p_A | 0.18% | 0.50% | 1.15% | 1.84% | 3.96% | 8.22% |
| p_B | 0.22% | 0.63% | 1.42% | 2.25% | 4.73% | 9.59% |
| p_{Γ} | 0.34% | 0.93% | 2.05% | 3.18% | 6.45% | 12.55% |

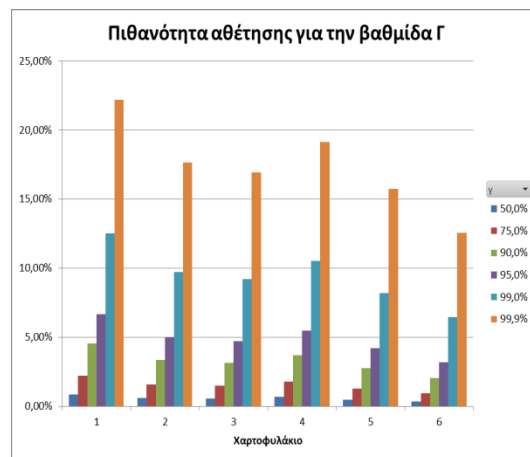
Πίνακας 2.21 Εκτίμηση των p_A, p_B, p_{Γ} για μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 6



Σχήμα 2.13 Γραφική απεικόνιση της p_A εξάρτηση του δείγματος, χωρίς αθετήσεις



Σχήμα 2.14 Γραφική απεικόνιση της p_B εξάρτηση του δείγματος, χωρίς αθετήσεις



Σχήμα 2.15 Γραφική απεικόνιση της p_Γ εξάρτηση του δείγματος, χωρίς αθετήσεις

Συγκρίνοντας τα αποτελέσματα αυτών των πινάκων με τους αντίστοιχους πίνακες της Ενότητας 2.1, παρατηρούμε ότι, παρότι και σε αυτή την περίπτωση υποθέσαμε μηδενικό αριθμό αθετήσεων, οι εκτιμώμενες πιθανότητες αθέτησης και για τις 3 βαθμίδες είναι σαφώς μεγαλύτερες, ιδιαίτερα για τις μεγάλες τιμές του συντελεστή γ .

Αναφορικά με την μέθοδο υπολογισμού, να τονίσουμε ότι βασίζεται σε προσομοιώσεις για τον υπολογισμό του ολοκληρώματος.

2.4 Μικρός αριθμός αθετήσεων και εξάρτηση των παρατηρήσεων του δείγματος

Όπως και στην αντίστοιχη περίπτωση που έχει περιγραφεί στην Ενότητα 2.2, θεωρούμε ότι σημειώθηκαν a , b , c αθετήσεις στις βαθμίδες Α, Β, Γ αντίστοιχα. Πλέον, οι εκτιμήσεις για τις πιθανότητες αθέτησης στην κάθε βαθμίδα καθορίζονται από τις παρακάτω ανισώσεις:

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(x) \sum_{i=0}^{a+b+c} \binom{n}{i} (g(\rho, p_A, x))^i (1 - g(\rho, p_A, x))^{n-i} dx \quad (2.20)$$

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(x) \sum_{i=0}^{b+c} \binom{n_B+n_r}{i} (g(\rho, p_B, x))^i (1 - g(\rho, p_B, x))^{n_B+n_r-i} dx \quad (2.21)$$

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(y) \sum_{i=0}^c \binom{n_r}{i} (g(\rho, p_r, x))^i (1 - g(\rho, p_r, x))^{n_r-i} dx \quad (2.22)$$

Για χαρτοφυλάκια της Ενότητας 2.2 και θεωρώντας την τιμή $\rho = 18\%$ κατασκευάζουμε τον παρακάτω πίνακα:

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|----------|-------|-------|-------|--------|--------|--------|
| p_A | 1% | 2.23% | 4.18% | 5.94% | 10.60% | 18.26% |
| p_B | 1.22% | 2.64% | 4.88% | 6.87% | 12.04% | 20.35% |
| p_r | 1.86% | 4.01% | 7.31% | 10.14% | 17.22% | 27.83% |

Πίνακας 2.22 Εκτίμηση των p_A, p_B, p_r για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 1

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|----------|-------|-------|-------|-------|-------|--------|
| p_A | 0.58% | 1.39% | 2.80% | 4.07% | 7.84% | 14.29% |
| p_B | 0.38% | 1.03% | 2.27% | 3.44% | 7.02% | 13.43% |
| p_r | 0.60% | 1.58% | 3.37% | 4.99% | 9.77% | 17.84% |

Πίνακας 2.23 Εκτίμηση των p_A, p_B, p_r για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 2

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|----------|-------|-------|-------|-------|--------|--------|
| p_A | 0.81% | 1.82% | 3.48% | 4.98% | 9.18% | 15.94% |
| p_B | 0.92% | 2.10% | 4.05% | 5.81% | 10.68% | 18.43% |
| p_r | 0.56% | 1.48% | 3.15% | 4.72% | 9.26% | 16.91% |

Πίνακας 2.24 Εκτίμηση των p_A, p_B, p_r για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 3

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|----------|-------|-------|-------|--------|--------|--------|
| p_A | 1.00% | 2.16% | 4.08% | 5.69% | 10.28% | 17.56% |
| p_B | 1.23% | 2.60% | 4.82% | 6.66% | 11.79% | 19.74% |
| p_r | 2.19% | 4.46% | 7.94% | 10.71% | 17.97% | 28.38% |

Πίνακας 2.25 Εκτίμηση των p_A, p_B, p_r για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 4

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|----------|-------|-------|-------|-------|-------|--------|
| p_A | 0.72% | 1.64% | 3.18% | 4.61% | 8.33% | 15.26% |
| p_B | 0.77% | 1.73% | 3.39% | 4.83% | 8.68% | 15.79% |
| p_r | 0.48% | 1.29% | 2.78% | 4.23% | 8.21% | 15.70% |

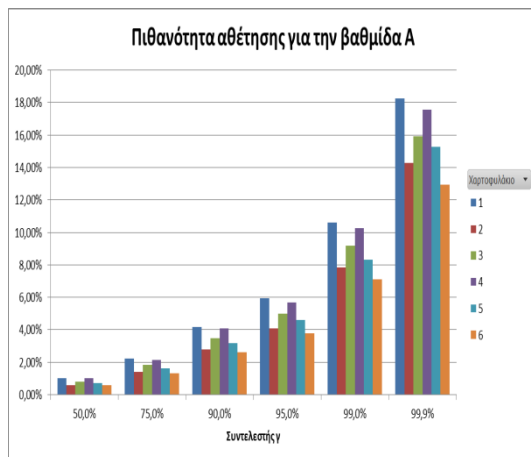
Πίνακας 2.26 Εκτίμηση των p_A, p_B, p_r για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 5

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|----------|-------|-------|-------|-------|--------|--------|
| p_A | 0.57% | 1.32% | 2.61% | 3.78% | 7.13% | 12.96% |
| p_B | 0.72% | 1.64% | 3.18% | 4.57% | 8.44% | 14.98% |
| p_r | 1.11% | 2.43% | 4.54% | 6.39% | 11.33% | 19.28% |

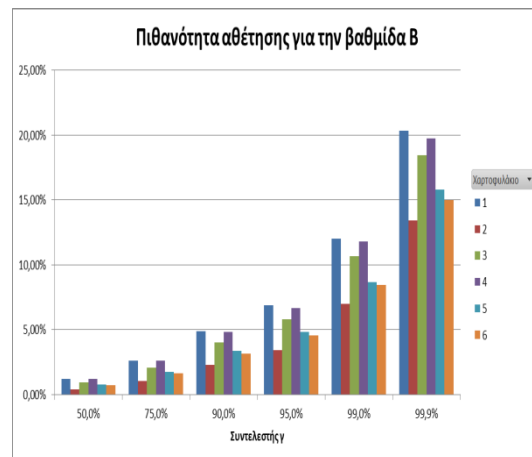
Πίνακας 2.27 Εκτίμηση των p_A, p_B, p_r για μη μηδενικό αριθμό αθετήσεων και εξάρτηση του δείγματος για το χαρτοφυλάκιο 6

Από τους Πίνακες παρατηρούμε πως και σε αυτή την περίπτωση, οι πιθανότητες είναι ιδιαίτερος υψηλές, σε βαθμό που να μη θεωρούνται λογικές για χαρτοφυλάκια υψηλής φερεγγυότητας. Οι εκτιμήσεις αυτές, όπως και στην Ενότητα 2.3, οφείλονται

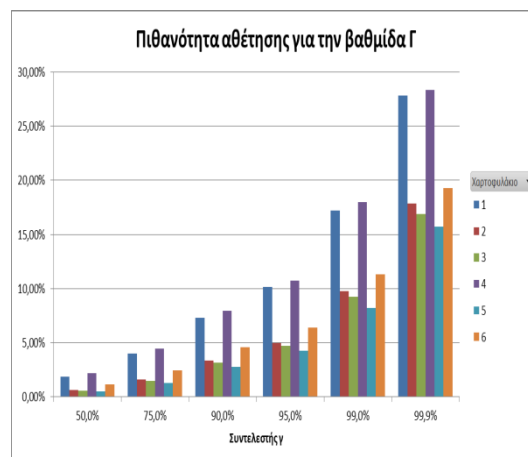
στην εισαγωγή του συντελεστή συσχέτισης ρ . Από τα ακόλουθα διαγράμματα μπορούμε να συμπεράνουμε πως σχεδόν όλες οι ιδιαιτερότητες που είχαμε επισημάνει στην Ενότητα 2.2 εξακολουθούν να υφίστανται.



Σχήμα 2.16 Γραφική απεικόνιση της p_A με εξάρτηση του δείγματος και αθετήσεις



Σχήμα 2.17 Γραφική απεικόνιση της p_B με εξάρτηση του δείγματος και αθετήσεις



Σχήμα 2.18 Γραφική απεικόνιση της p_Γ με εξάρτηση του δείγματος και αθετήσεις

2.5 Περίοδος παρατήρησης άνω του ενός έτους

Στις περιπτώσεις που αναλύσαμε παραπάνω η πιθανότητα αθέτησης εκτιμήθηκε βάσει δεδομένων μόλις ενός έτους. Όταν ωστόσο, είναι διαθέσιμα δεδομένα μιας χρονικής περιόδου άνω του ενός έτους, υπάρχει η δυνατότητα να χωριστεί το χρονικό διάστημα σε περιόδους του ενός έτους και να εκτιμηθεί η πιθανότητα αθέτησης για κάθε μία από τις περιόδους. Οι εκτιμήσεις αυτές μπορούν να χρησιμοποιηθούν στη συνέχεια για να εκτιμήσουμε τη ζητούμενη πιθανότητα αθέτησης εντός της συνολικής περιόδου. Με τον τρόπο αυτό, γίνεται πιο αποτελεσματική χρήση των ιστορικών δεδομένων του πιστωτικού ιδρύματος.

Σύμφωνα με τη Βασιλεία II, οι τράπεζες πρέπει να χρησιμοποιούν δεδομένα τουλάχιστον των τελευταίων 5 ετών. Στο μοντέλο που θα παρουσιασθεί στη συνέχεια,

υποθέτουμε ότι για όλους τους οφειλέτες που ανήκουν στην ίδια βαθμίδα πιστοληπτικής ικανότητας, οι τιμές των ρ , ρ τα οποία συμβολίζουν την πιθανότητα αθέτησης και την συσχέτιση μεταξύ των στοιχείων του ενεργητικού αντίστοιχα είναι σταθερές. Συμβολίζουμε με $D_{i,t}$ την αξία των περιουσιακών στοιχείων του i οφειλέτη τη χρονική στιγμή t . Τότε το μοντέλο που πρότειναν οι Pluto and Tasche (2005) γράφεται ως:

$$D_{i,t} = \sqrt{\rho}X_t + Y_{i,t}\sqrt{1-\rho} \quad (2.23)$$

όπου $i = 1, \dots, n_t$, $t = 1, \dots, T$. Το X_t είναι ένας παράγοντας κοινός για όλους τους οφειλέτες για δεδομένο χρόνο t , T είναι τα χρόνια παρατήρησης και n_t είναι ο αριθμός των οφειλετών στο χαρτοφυλάκιο τη χρονική στιγμή t .

Επίσης, θεωρούμε ότι

$$\begin{aligned} Y_{i,t} &\sim N(0,1) \text{ για } i = 1, \dots, n_t, t = 1, \dots, T \\ X_t &\sim N(0,1) \text{ για } t = 1, \dots, T \end{aligned}$$

Επιπλέον, υποθέτουμε ότι τα $Y_{i,t}$ και X_t είναι ανεξάρτητα για κάθε i, t . Βάσει των υποθέσεων αυτών και της σχέσης (2.23) προκύπτει ότι τα $D_{i,t}$ ακολουθούν επίσης τυπική κανονική κατανομή για κάθε i, t . Θεωρούμε ότι ένας δανειολήπτης αθετεί την αποπληρωμή την χρονική στιγμή t αν η αξία των περιουσιακών του στοιχείων γίνει μικρότερη ή ίση από κάποιο όριο c για πρώτη φορά τη χρονική στιγμή t , δηλαδή:

$$D_{i,1} > c, \dots, D_{i,t-1} > c, D_{i,t} \leq c \quad (2.24)$$

Τότε η πιθανότητα που μας ενδιαφέρει να εκτιμήσουμε είναι ίση με

$$P[D_{i,t} \leq c] = p \quad (2.25)$$

και για την σταθερά c μπορούμε να γράψουμε,

$$c = \Phi^{-1}(p) \quad (2.26)$$

Ορίζουμε το διάνυσμα $X = (X_1, \dots, X_T)'$. Εφόσον οι συνιστώσες του διανύσματος είναι τυποποιημένες κανονικές, η από κοινού κατανομή μπορεί να προσδιορίζεται βάσει του πίνακα συσχετίσεων:

| | | | |
|-----------|-----------|----------|-----------|
| 1 | $r_{1,2}$ | ... | $r_{1,T}$ |
| $r_{2,1}$ | 1 | ... | $r_{2,T}$ |
| \vdots | \vdots | \ddots | \vdots |
| $r_{T,1}$ | $r_{T,2}$ | ... | 1 |

όπου

$$\text{Corr}(X_s, X_t) = r_{s,t} = \tau^{|s-t|}, s, t = 1, \dots, T, s \neq t \quad (2.27)$$

για κάποιο κατάλληλο $0 < \tau < 1$. Από την σχέση (2.27) λαμβάνουμε τη συσχέτιση των συστηματικών παραγόντων για τα χρόνια s, t .

Υποθέτουμε πάλι ότι έχουμε 3 βαθμίδες πιστοληπτικής ικανότητας. Εντός των T χρόνων παρατήρησης παρατηρήθηκαν συνολικά k_A, k_B και k_T αθετήσεις στις αντίστοιχες βαθμίδες. Για την εκτίμηση της p_A , όπως και στις προηγούμενες περιπτώσεις, θα χρησιμοποιήσουμε τον συνολικό αριθμό οφειλετών n και τον συνολικό αριθμό αθετήσεων

$$k = k_A + k_B + k_T . \quad (2.28)$$

Όπως και στις προηγούμενες περιπτώσεις, η εκτίμηση γίνεται χρησιμοποιώντας τη σχέση:

$$1 - \gamma \leq P[\text{να παρατηρηθούν το πολύ } k \text{ αθετήσεις}] .$$

Από τις σχέσεις (2.23),(2.24),(2.26) δοθεισών τιμών x_1, x_2, \dots, x_T για τους συστηματικούς παράγοντες, για τη δεσμευμένη πιθανότητα αθέτησης έχουμε :

$$\begin{aligned} \pi(x_1, \dots, x_T) &= P[\text{o οφειλέτης } i \text{ αθετεί} | x_1, \dots, x_T] \\ &= P[\min_{t=1, \dots, T} D_{i,t} \leq \Phi^{-1}(p) | x_1, \dots, x_T] \\ &= 1 - P[Y_{i,1} > \frac{\Phi^{-1}(p_A) - \sqrt{\rho}x_1}{\sqrt{1-\rho}}, \dots, Y_{i,T} > \frac{\Phi^{-1}(p_A) - \sqrt{\rho}x_T}{\sqrt{1-\rho}}] \\ &= 1 - \prod_{t=1}^T (1 - g(p_A, \rho, x_t)) . \end{aligned} \quad (2.29)$$

Για την εκτίμηση λοιπόν της αδέσμευτης p_A αρκεί να λυθεί η ανίσωση:

$$1 - \gamma \leq \int_{-\infty}^{\infty} \varphi(x) \sum_{i=0}^k \binom{n}{i} (\pi(x_1, \dots, x_T))^i (1 - \pi(x_1, \dots, x_T))^{n-i} dx . \quad (2.28)$$

Με την ίδια μεθοδολογία προκύπτουν και οι ανισώσεις για τις βαθμίδες B και T . Παρακάτω, παρατίθεται πίνακας με αριθμητικά αποτελέσματα στην περίπτωση που κατά την περίοδο παρατήρησης δεν έχει συμβεί καμία αθέτηση για τα χαρτοφυλάκια 1 και 5 της Ενότητας 2.1. Οι τιμές που ορίσαμε για τις συσχετίσεις και την περίοδο παρατήρησης είναι:

$$\tau = 0.40, T = 5, \rho = 0.20$$

| | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
| p_A | 0.05% | 0.12% | 0.26% | 0.39% | 0.80% | 1.61% |
| p_B | 0.06% | 0.15% | 0.31% | 0.47% | 0.94% | 1.88% |
| p_T | 0.14% | 0.35% | 0.71% | 1.03% | 1.97% | 3.71% |

Πίνακας 2.28 Εκτίμηση των p_A, p_B, p_T για το χαρτοφυλάκιο 1 στην πολυετή περίπτωση

| | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
| p_A | 0.03% | 0.08% | 0.19% | 0.28% | 0.59% | 1.19% |
| p_B | 0.04% | 0.09% | 0.20% | 0.30% | 0.62% | 1.25% |
| p_T | 0.08% | 0.20% | 0.40% | 0.59% | 1.17% | 2.26% |

Πίνακας 2.29 Εκτίμηση των p_A, p_B, p_T για το χαρτοφυλάκιο 5 στην πολυετή περίπτωση

Στην περίπτωση που κατά τη διάρκεια των 5 ετών έχει παρατηρηθεί 1 αθέτηση στην βαθμίδα A , 2 αθετήσεις στην βαθμίδα B και 3 στην βαθμίδα Γ , και θεωρώντας τις ίδιες τιμές για τα τ, ρ όπως στους προηγούμενους πίνακες, για το χαρτοφυλάκιο 4 έχουμε:

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.29% | 0.54% | 0.91% | 1.24% | 2.08% | 3.56% |
| p_B | 0.31% | 0.59% | 0.98% | 1.33% | 2.24% | 3.82% |
| p_Γ | 0.52% | 0.97% | 1.61% | 2.17% | 3.59% | 5.95% |

Πίνακας 2.30 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 4 στην πολυετή περίπτωση

Συγκρίνοντας τα αποτελέσματα των παραπάνω πινάκων με τα αντίστοιχα αποτελέσματα από τις Ενότητες 2.3 και 2.4, βλέπουμε πόσο σημαντικός παράγοντας για την εκτίμηση της πιθανότητας αθέτησης είναι η περίοδος παρατήρησης. Οι πιθανότητες που εκτιμήθηκαν με περίοδο παρατήρησης τα 5 έτη είναι δραματικά μικρότερες από αυτές που εκτιμήθηκαν για περίοδο παρατήρησης 1 έτους για όλες τις βαθμίδες, παρόλο που θεωρήσαμε υψηλότερη τιμή για την συσχέτιση ρ και αυξήσαμε τον αριθμό των αθετήσεων για το χαρτοφυλάκιο 4.

2.6 Η χρήση της γλώσσας R στην εκτίμηση των πιθανοτήτων αθέτησης

Όπως ήδη αναφέραμε η υπολογιστική δυσκολία για την εκτίμηση των πιθανοτήτων αθέτησης μέσα από κάποιους από τους τύπους που δόθηκαν στις Ενότητες 2.3, 2.4 και 2.5 είναι ιδιαίτερα υψηλή. Προκειμένου να εκτιμήσουμε τις ζητούμενες πιθανότητες, ενδείκνυται η χρήση της γλώσσας R και του πακέτου *LDPD* (Surzhko, 2015). Στο εν λόγω πακέτο μεταξύ άλλων, υπάρχουν οι ακόλουθες 2 εντολές:

- α) *PTOnePeriodPD* με ορίσματα *portf.uncond*, *portf.def*. και *conf.interval*
- β) *PTMultiPeriodPD* με ορίσματα *portf.uncond*, *portf.def.*, *rho*, *cor.St*, *kT*, *kNS* και *conf.interval*

Η εντολή *PTOnePeriodPD* χρησιμοποιείται για τις περιπτώσεις που αναλύθηκαν στις Ενότητες 2.1 και 2.2. ενώ η *PTMultiPeriodPD* για τις περιπτώσεις των Ενότητων 2.3, 2.4 και 2.5. Τα ορίσματα είναι τα εξής

portf.uncond : ένα διάνυσμα γραμμής που περιέχει τον αριθμό οφειλετών ανά βαθμίδα πιστοληπτικής ικανότητας

portf.def : ένα διάνυσμα γραμμής που περιέχει τον αριθμό αθετήσεων ανά βαθμίδα πιστοληπτικής ικανότητας

rho : ο συντελεστής ρ

cor.St : ο συντελεστής τ

kT : η χρονική περίοδος T

kNS : ο αριθμός των προσομοιώσεων για την εκτίμηση του ολοκληρώματος

conf.interval : ο συντελεστής γ

Χρειάζεται ιδιαίτερη προσοχή στα ορίσματα *portf.uncond* και *portf.def*, καθώς πρέπει να γράψουμε τα ζητούμενα διανύσματα ξεκινώντας από την χαμηλότερη βαθμίδα πιστοληπτικής ικανότητας και συνεχίζοντας προς την καλύτερη. Επίσης η εντολή *PTMultiPeriodPD* μπορεί να χρησιμοποιηθεί για όλες τις περιπτώσεις που αναφέρθηκαν στις Ενότητες 2.1 έως και 2.5.

Αν δώσουμε τις τιμές $\rho=0$, $kNS>1$, $kT=1$ και οποιαδήποτε τιμή στο *cor.ST* η εντολή *PTMultiPeriodPD* “λειτουργεί” ακριβώς όπως η *PTOnePeriodPD* και μπορούμε να εκτιμήσουμε πιθανότητες για τις περιπτώσεις των Ενοτήτων 2.1 και 2.2. Επίσης αν θέλουμε να υπολογίσουμε πιθανότητες αθέτησης με την μεθοδολογία των Ενοτήτων 2.3 και 2.4 αρκεί να θέσουμε $kT=1$ και οποιαδήποτε τιμή στο *cor.ST*. Γενικά επειδή το όρισμα *cor.ST* αντιστοιχεί στην συσχέτιση μεταξύ των παραγόντων X_1, \dots, X_T , στην περίπτωση που $kT=1$ δεν έχει καμία σημασία τι τιμή θα δώσουμε στο συγκεκριμένο όρισμα. Τέλος να σημειώσουμε πως για $\rho \neq 0$, ο υπολογισμός του ολοκληρώματος που περιέχεται στους τύπους για την εκτίμηση της πιθανότητας αθέτησης γίνεται μέσω του αριθμού των προσομοιώσεων kNS . Αυτό έχει σαν αποτέλεσμα κάθε φορά που τρέχουμε τον κώδικα να μην λαμβάνουμε ακριβώς τις ίδιες εκτιμήσεις. Για να επιλύσουμε το συγκεκριμένο ζήτημα επιλέξαμε να χρησιμοποιήσουμε επαναληπτικές εντολές.

2.7 Σχολιασμός της μεθόδου

Η μεθοδολογία που αναλύθηκε σε αυτό το κεφάλαιο απαντά στο κύριο ζήτημα που προκύπτει στα χαρτοφυλάκια υψηλής φερεγγυότητας, καθώς δύναται να εκτιμήσει την πιθανότητα αθέτησης του χαρτοφυλακίου, ακόμα και στην περίπτωση που δεν υπάρχουν αθετήσεις στα ιστορικά δεδομένα του χαρτοφυλακίου. Για αυτό το λόγο, αποτελεί μία μέθοδο η οποία εφαρμόζεται αρκετά συχνά στην πράξη. Ωστόσο από τη φύση της μεθόδου ανακύπτουν κάποια ζητήματα.

Αρχικά, η μεθοδολογία έχει επικριθεί για το γεγονός πως παράγει υπερβολικά συντηρητικά αποτελέσματα. Πράγματι, αν ανατρέξουμε στην Ενότητα 2.3, θα δούμε ότι στα παραδείγματά μας, παρότι έχουμε θεωρήσει μηδενικό αριθμό αθετήσεων για κάθε βαθμίδα κάθε χαρτοφυλακίου, οι εκτιμήσεις που λαμβάνουμε για $\gamma = 99,9\%$ κυμαίνονται από 8.22% έως 22.20%. Πρακτικά αυτό μας δείχνει ότι η εισαγωγή του συντελεστή ρ είναι μείζονος σημασίας και οδηγεί σε πολύ υψηλές εκτιμήσεις. Η χρήση δεδομένων για περίοδο μεγαλύτερη του 1 έτους αποτελεί ενδεχόμενη λύση στο θέμα της συντηρητικής εκτίμησης, όπως φαίνεται από τους πίνακες που δίνονται στην Ενότητα 2.5. Στην περίπτωση αυτή ωστόσο, για να μπορέσει το πιστωτικό ίδρυμα να εκτιμήσει την πιθανότητα αθέτησης, οφείλει να προσδιορίσει όχι μόνο τον συντελεστή γ και τη συσχέτιση ρ μεταξύ στοιχείων του ενεργητικού, αλλά και τη συσχέτιση μεταξύ των παραγόντων X_1, \dots, X_T .

3. Η προσέγγιση Bayes

Στη μέθοδο εκτίμησης που εισήγαγαν οι Pluto και Tasche (2005), είδαμε στο προηγούμενο κεφάλαιο να γίνεται σαφές ότι μπορούμε να εκτιμήσουμε την πιθανότητα αθέτησης σε χαρτοφυλάκια με πολύ μικρό ή και μηδενικό αριθμό αθετήσεων. Η προσέγγιση Bayes είναι ακόμα μία μέθοδος μέσω της οποίας επίσης μπορούμε να εκτιμήσουμε την πιθανότητα αθέτησης ενός χαρτοφυλακίου υψηλής φερεγγυότητας δίχως όμως τη χρήση διαστημάτων εμπιστοσύνης.

3.1 Το θεώρημα Bayes

Αν A και B είναι δύο ενδεχόμενα ενός δειγματικού χώρου Ω και $P(B) \geq 0$, η δεσμευμένη πιθανότητα του A δοθέντος B είναι

$$P(A|B) = \frac{P(A \cap B)}{P(B)} . \quad (3.1)$$

Ισοδύναμα, η παραπάνω σχέση γράφεται:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} . \quad (3.2)$$

Αν η X είναι μια τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x|\theta)$ και η θ ακολουθεί κατανομή με συνάρτηση πυκνότητας πιθανότητας $\pi(\theta)$, τότε η από κοινού συνάρτηση πυκνότητας πιθανότητας των X, θ δίνεται από τη σχέση:

$$f(x; \theta) = f(x|\theta) * \pi(\theta) . \quad (3.3)$$

Από το θεώρημα του Bayes και τη σχέση (3.3) προκύπτει

$$f(\theta|x) = \frac{f(x|\theta)*\pi(\theta)}{g(x)} . \quad (3.4)$$

όπου η $g(x)$ καλείται περιθώρια συνάρτηση κατανομής και υπολογίζεται ως :

$$g(x) = \int f(x|\theta) * \pi(\theta) d\theta . \quad (3.5)$$

Έστω τώρα ότι οι τυχαίες μεταβλητές X_1, X_2, \dots, X_n δοθέντος ότι $\theta = \theta$ είναι ανεξάρτητες και ισόνομες, με συνάρτηση πυκνότητας πιθανότητας $f(x|\theta)$ και η θ ακολουθεί την κατανομή $\pi(\theta)$, τότε η από κοινού συνάρτηση πυκνότητας πιθανότητας των $X_1, X_2, \dots, X_n, \theta$ δίνεται από τη σχέση:

$$f(x_1, x_2, \dots, x_n, \theta) = [\prod_{i=1}^n f(x_i|\theta)] * \pi(\theta) . \quad (3.6)$$

Και η περιθώρια κατανομή των X_1, X_2, \dots, X_n λαμβάνεται ως :

$$g(x_1, x_2, \dots, x_n) = \int [\prod_{i=1}^n f(x_i|\theta)] * \pi(\theta) d\theta . \quad (3.7)$$

Από τις σχέσεις (3.6),(3.7) και το θεώρημα του Bayes προκύπτει:

$$f(\theta/x_1, x_2, \dots, x_n) = \frac{= [\prod_{i=1}^n f(x_i|\theta)] * \pi(\theta)}{\int [\prod_{i=1}^n f(x_i|\theta)] * \pi(\theta) d\theta} . \quad (3.8)$$

Η $\pi(\theta)$ ονομάζεται εκ των προτέρων κατανομή της θ και αντιπροσωπεύει την εκ των προτέρων πληροφορία, δηλαδή τη γνώμη που έχουμε σχετικά με την θ . Η $f(\theta|x_1, x_2, \dots, x_n)$ καλείται εκ των υστέρων κατανομή της θ και αποτελεί την συνολική γνώση που έχουμε για την παράμετρο, αφού προσέξουμε την πληροφορία που λαμβάνουμε για την τυχαία μεταβλητή δοθέντος του διανύσματος (X_1, X_2, \dots, X_n) .

Όπως παρατηρούμε από τη σχέση (3.8), ο παρονομαστής είναι ανεξάρτητος του θ . Για αυτό το λόγο, η σχέση μπορεί πιο εύκολα να γραφεί ως :

$$f(\theta|x_1, x_2, \dots, x_n) \propto \prod_{i=1}^n f(x_i|\theta) * \pi(\theta) . \quad (3.9)$$

Από την παραπάνω σχέση βλέπουμε ότι προκειμένου να προσδιορίσουμε την εκ των υστέρων κατανομή της τ.μ. θ χρειαζόμαστε την συνάρτηση πιθανοφάνειας των τ.μ. X_1, X_2, \dots, X_n και την εκ των προτέρων κατανομή της θ .

3.2 Η κατανομή του δείγματος

Υποθέτουμε την ύπαρξη ενός χαρτοφυλακίου με n τον αριθμό οφειλετών και συμβολίζουμε το εν λόγω χαρτοφυλάκιο με

$$D = \{D_i, i = 1, \dots, n\} .$$

Η τυχαία μεταβλητή D_i παριστά το ενδεχόμενο της αθέτησης από τον i δανειολήπτη.

Προφανώς τότε:

$$D_i = \begin{cases} 1, & \text{αν ο } i \text{ δανειολήπτης αθετήσει} \\ 0, & \text{αν ο } i \text{ δανειολήπτης δεν αθετήσει} \end{cases} . \quad (3.10)$$

Εφόσον η τ.μ. D_i ακολουθεί κατανομή Bernoulli με παράμετρο θ_i , η συνάρτηση πυκνότητας πιθανότητας της τ.μ. δίνεται ως:

$$f_{D_i}(d_i) = \theta_i^{d_i} (1 - \theta_i)^{1-d_i} . \quad (3.11)$$

Υποθέτουμε ότι υπάρχει ανεξαρτησία μεταξύ των αθετήσεων και πως όλοι οι οφειλέτες του χαρτοφυλακίου έχουν την ίδια πιθανότητα αθέτησης $\theta = \theta_1 = \dots = \theta_n$.

Σε αυτή την περίπτωση η συνάρτηση πιθανοφάνειας για το δείγμα μας είναι:

$$f(d_1, d_2, \dots, d_n/\theta) = \prod_{i=1}^n \theta^{d_i} (1 - \theta)^{1-d_i} \quad (3.12)$$

$$\Leftrightarrow f(d_1, d_2, \dots, d_n/\theta) = \theta^{\sum_{i=1}^n d_i} (1 - \theta)^{n - \sum_{i=1}^n d_i} . \quad (3.13)$$

Θέτουμε $r = \sum_{i=1}^n d_i$ τον αριθμό των συνολικών αθετήσεων. Τότε,

$$f(d_1, d_2, \dots, d_n/\theta) = \theta^r (1 - \theta)^{n-r} . \quad (3.14)$$

Προφανώς, η κατανομή του συνολικού αριθμού αθετήσεων είναι Binomial (n, θ) . Η συγκεκριμένη κατανομή απλοποιεί σε αρκετά μεγάλο βαθμό το μοντέλο. Οι

αθετήσεις που συμβαίνουν στο χαρτοφυλάκιο, δεν είναι στην πραγματικότητα ανεξάρτητες. Επίσης, η υπόθεση της ομοιογένειας είναι και αυτή αρκετά σημαντική, καθώς αν το χαρτοφυλάκιο περιέχει ανόμοια περιουσιακά στοιχεία και δεν μπορεί να θεωρηθεί ομοιογενές, τα αποτελέσματα που θα λάβουμε από τη συγκεκριμένη μεθοδολογία δεν θα είναι απολύτως αξιόπιστα.

3.3 Επιλογή της εκ των προτέρων κατανομής

Εφόσον η συνάρτηση πιθανοφάνειας είναι γνωστή, για να εκτιμήσουμε την πιθανότητα αθέτησης, πρέπει να προσδιορίσουμε την εκ των προτέρων κατανομή. Η κατανομή αυτή βασίζεται στις προσδοκίες μας για την κατανομή της θ , πριν μπορέσουμε να παρατηρήσουμε τα αποτελέσματα του χαρτοφυλακίου. Η επιλογή της εκ των προτέρων κατανομής είναι σημείο συζήτησης, καθώς υπάρχει υποκειμενικότητα.

Ακόμα και όταν δεν θέλουμε να εκφράσουμε καμία προσδοκία για την τυχαία μεταβλητή, μπορούμε να ορίσουμε μια κατάλληλη εκ των προτέρων κατανομή. Χαρακτηριστικά μπορούμε να χρησιμοποιήσουμε την ομοιόμορφη στο διάστημα $(0,1)$ με $\pi(\theta) = 1, \theta \in [0,1]$.

Χρησιμοποιώντας τη συγκεκριμένη κατανομή, η οποία αναφέρεται ως μη πληροφοριακή εκ των προτέρων κατανομή, αντιστοιχούμε ίσες πιθανότητες σε όλες τις ενδεχόμενες τιμές της παραμέτρου. Άμεση συνέπεια της συγκεκριμένης επιλογής είναι ότι η εκ των υστέρων κατανομή διαφέρει ελάχιστα από την πιθανοφάνεια.

Στα χαρτοφυλάκια υψηλής φερεγγυότητας ωστόσο, υπάρχουν σημαντικές πληροφορίες για την μεταβλητή θ , οι οποίες θα μας βοηθήσουν να καθορίσουμε όσο το δυνατόν πιο αντικειμενικά την εκ των προτέρων κατανομή. Συγκεκριμένα, στα πλαίσια της παρούσας Δ.Ε, ενδιαφερόμαστε για την περίπτωση που το χαρτοφυλάκιο το οποίο μελετάμε χαρακτηρίζεται ως χαρτοφυλάκιο υψηλής φερεγγυότητας. Ως εκ τούτου, προσδοκούμε μικρό αριθμό αθετήσεων, άρα είναι λογικό να αναμένουμε μία μικρή τιμή ως εκτίμηση της πιθανότητας αθέτησης.

Βάσει αυτού, μπορούμε με σχετικά μεγάλη βεβαιότητα να απορρίψουμε τη μη πληροφοριακή κατανομή, καθώς εύκολα προκύπτει ότι σε αυτήν την περίπτωση η μέση τιμή της πιθανότητας αθέτησης είναι 0.5. Δεδομένου ότι μελετάμε χαρτοφυλάκιο υψηλής φερεγγυότητας, η ομοιόμορφη κατανομή δε θεωρείται κατάλληλη για το χαρτοφυλάκιο.

Τυπικά, οποιαδήποτε κατανομή μπορεί να θεωρηθεί ως εκ των προτέρων κατανομή της τυχαίας μεταβλητής θ . Ωστόσο, στους περισσότερους συνδυασμούς θα προκύψουν υπολογιστικές δυσκολίες. Για τον λόγο αυτό, εΐθισται να επιλέγεται μια συζυγής κατανομή ως εκ των προτέρων κατανομή της πιθανότητας αθέτησης. Συζυγείς ονομάζονται οι οικογένειες κατανομών οι οποίες, όταν επιλεγούν ως εκ των προτέρων κατανομές, τότε και οι εκ των υστέρων κατανομές θα ανήκουν στην ίδια οικογένεια κατανομών. Παρότι η επιλογή μιας συζυγούς εκ των προτέρων κατανομής θα απλοποιήσει σε μεγάλο βαθμό τους υπολογισμούς, η επιλογή της οφείλει να γίνεται με

μεγάλη προσοχή και μόνο εφόσον συνάδει με τις πεποιθήσεις μας για την κατανομή της τυχαίας μεταβλητής.

Συνοπτικά τονίζουμε ότι:

- α) Όταν δεν υπάρχει καμία πληροφορία, επιλέγεται μια μη πληροφοριακή κατανομή.
- β) Η κατανομή επιλέγεται κυρίως βάσει υποκειμενικής ανάλυσης-γνώσης του ειδικού.
- γ) Υπό κατάλληλες συνθήκες επιλέγεται μια κατανομή τέτοια, ώστε να υπάρχει υπολογιστική ευκολία.
- δ) Το μέγεθος των δεδομένων παίζει καθοριστικό ρόλο στην εκτίμηση της εκ των υστέρων κατανομής. Όταν υπάρχουν πολλά δεδομένα, η εκ των υστέρων κατανομή είναι σταθμισμένη προς την πιθανοφάνεια.

3.4 Η εκτίμηση της πιθανότητας αθέτησης

Στις επόμενες ενότητες, θα εκτιμήσουμε την πιθανότητα αθέτησης χρησιμοποιώντας ως εκ των προτέρων κατανομές τις:

- α) $Beta(\alpha, \beta)$
- β) Συντηρητική (conservative) κατανομή
- γ) Ομοιόμορφη στο $[0, u]$

3.4.1 Η κατανομή Beta με δύο παραμέτρους

Οι περιπτώσεις που θα αναλυθούν στην παρούσα ενότητα καθώς και στις Ενότητες 3.4.2 και 3.4.3 αφορούν μονοετή περίοδο παρατήρηση και υποθέτουν ανεξαρτησία των αθετήσεων.

Ο Kiefer (2007) πρότεινε ως εκ των προτέρων κατανομή για την παράμετρο θ την κατανομή $Beta(\alpha, \beta)$:

$$\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}, 0 < \theta < 1. \quad (3.15)$$

Ένας επιπλέον λόγος που συνηγορεί στη χρήση της $Beta$ ως εκ των προτέρων κατανομή είναι η κατανομή των παρατηρήσεων που έχουμε. Όπως είδαμε προηγουμένως, οι τ.μ. $D_i, i = 1, \dots, n$ ακολουθούν την κατανομή $Bernoulli(\theta)$ και η συνάρτηση πιθανοφάνειας αυτών είναι Binomial (n, θ) . Σε αυτήν την περίπτωση, με χρήση της $Beta(a, b)$ ως εκ των προτέρων κατανομή της τ.μ. θ έχουμε:

$$\begin{aligned} f(\theta/d_1, d_2, \dots, d_n) &\propto \left[\prod_{i=1}^n f(d_i/\theta) \right] * \pi(\theta) \\ &= \theta^r (1-\theta)^{n-r} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \propto \\ &\theta^{(\alpha+r)-1} (1-\theta)^{(n+\beta-r)-1}. \end{aligned} \quad (3.16)$$

Από την (3.16), καταλήγουμε στο συμπέρασμα ότι η εκ των υστέρων κατανομή της πιθανότητας αθέτησης είναι η Beta με παραμέτρους $a' = a + r$ και $\beta' = n + \beta - r$. Γνωρίζοντας την εκ των υστέρων κατανομή, μπορούμε να πάρουμε ως εκτίμηση για την πιθανότητα αθέτησης την μέση τιμή.

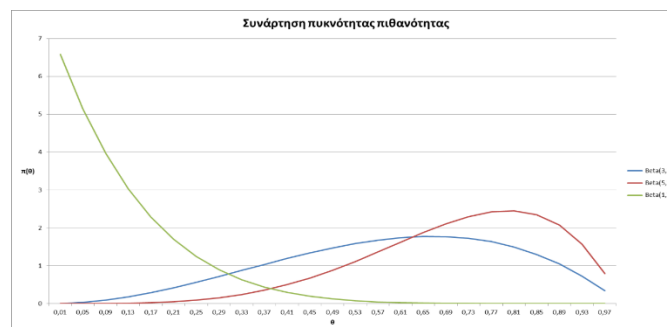
$$\hat{\theta} = E(\theta) = \frac{a+r}{a+n+b} . \quad (3.17)$$

Η εκ των προτέρων κατανομή έχει μέση τιμή $\frac{\alpha}{\alpha+\beta}$ η οποία θα ήταν η καλύτερη εκτίμηση για την θ , δίχως να παρατηρήσουμε το χαρτοφυλάκιο. Αν αντί να λάβουμε μια εκ των προτέρων κατανομή και να χρησιμοποιήσουμε την προσέγγιση Bayes, περιμέναμε να δούμε πόσοι από τους δανειολήπτες αθέτησαν τις υποχρεώσεις τους, τότε θα λαμβάναμε τον ιστορικό μέσο $\frac{r}{n}$ ως εκτίμηση για την πιθανότητα αθέτησης, υπό την προϋπόθεση να είχαν υπάρξει αθετήσεις. Μέσω της προσέγγισης Bayes και οι δύο αυτές πληροφορίες συνδυάστηκαν. Αυτό γίνεται πιο εμφανές, αν γράψουμε την (3.17) ως:

$$E(\theta) = \left(\frac{n}{\alpha+\beta+n}\right)\frac{r}{n} + \left(\frac{\alpha+\beta}{\alpha+\beta+n}\right)\frac{\alpha}{\alpha+\beta} . \quad (3.18)$$

Από την (3.19) βλέπουμε ότι η εκ των υστέρων μέση τιμή της θ είναι γραμμικός συνδυασμός του ιστορικού μέσου και της εκ των προτέρων μέσης τιμής με τα βάρη να καθορίζονται από τα a, b, n .

Δεν είναι απαραίτητο, όταν θέλουμε να εκτιμήσουμε μια διωνυμική παράμετρο να χρησιμοποιήσουμε μια Beta ως εκ των προτέρων κατανομή. Παρόλα αυτά, υπάρχουν σημαντικά πλεονεκτήματα για να επιλέξουμε την οικογένεια των Beta κατανομών. Το κυριότερο εξ αυτών είναι πως η Beta αποτελεί συζυγή κατανομή, με αποτέλεσμα να λαμβάνουμε κλειστή μορφή υπολογισμού για τον εκτιμητή και η ενημέρωση της εκ των προτέρων κατανομής γίνεται με την αλλαγή των παραμέτρων. Μαθηματικά αυτό είναι αρκετά βολικό, καθώς κάνει τους υπολογισμούς αρκετά εύκολους. Παρακάτω παρατίθεται διάγραμμα με τις συναρτήσεις πυκνότητας πιθανότητας των κατανομών $Beta(3,2)$, $Beta(5,2)$ και $Beta(1,7)$. Όπως είναι εμφανές από το εν λόγω διάγραμμα, η Beta είναι μια κατανομή που μπορεί να πάρει αρκετά διαφορετικά σχήματα.



Σχήμα 3.1 Συνάρτηση πυκνότητας πιθανότητας κατανομών Beta

Προκειμένου να επιλεγούν οι παράμετροι α , β της εκ των προτέρων κατανομής *Beta* ο ειδικός του εκάστοτε πιστωτικού ιδρύματος βάσει των γνώσεων και της εμπειρίας του εκφράζει την γνώμη του για 2 μέτρα της κατανομής, όπως η μέση τιμή, η διακύμανση ή η ασυμμετρία. Για τα παρακάτω αριθμητικά παραδείγματα χρησιμοποιήσαμε τα χαρτοφυλάκια της Ενότητας 2.2. Προφανώς αν στο κάθε χαρτοφυλάκιο θεωρήσουμε ότι υπάρχουν πάνω από μία βαθμίδες πιστοληπτικής ικανότητας πρέπει να επιλέξουμε διαφορετική εκ των προτέρων κατανομή για την κάθε βαθμίδα. Χάριν ευκολίας, θεωρήσαμε πως όλοι οι δανειολήπτες ανήκουν σε μία μοναδική βαθμίδα πιστοληπτικής ικανότητας. Για να καθορίσουμε τις παραμέτρους των κατανομών βασιστήκαμε στις εκτιμήσεις της Ενότητας 2.2 για $\gamma = 90\%$. Παρατηρούμε ότι η μικρότερη εκτίμηση που λάβαμε ήταν 0.53% και η μεγαλύτερη 2.64%. Βάσει αυτών, για την εκτίμηση θ_1 επιλέξαμε ως εκ των προτέρων κατανομή την *Beta* (2.1,300) η οποία έχει μέση τιμή 0.7%, για την θ_2 την *Beta* (2.1,418) με μέση τιμή 0.5% και για την θ_3 την *Beta* (4.2,300) με μέση τιμή 1.4% ,ενώ στην τελευταία στήλη του πίνακα έχουμε την εκτίμηση της πιθανότητας αθέτησης μέσω του εκτιμητή μέγιστης πιθανοφάνειας.

| n | r | θ_1 | θ_2 | θ_3 | MLE |
|------|---|------------|------------|------------|-------|
| 500 | 2 | 0.51% | 0.45% | 0.77% | 0.4% |
| 600 | 1 | 0.34% | 0.30% | 0.58% | 0.17% |
| 650 | 2 | 0.43% | 0.38% | 0.65% | 0.31% |
| 700 | 3 | 0.51% | 0.46% | 0.72% | 0.43% |
| 750 | 2 | 0.39% | 0.35% | 0.59% | 0.27% |
| 1000 | 2 | 0.31% | 0.29% | 0.48% | 0.20% |

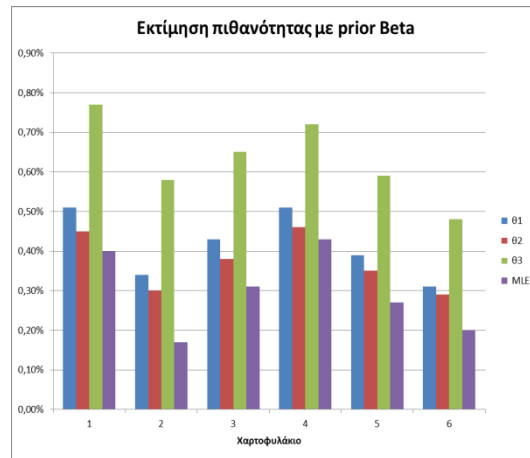
Πίνακας 3.1 Εκτίμηση της πιθανότητας αθέτησης με χρήση *Beta* εκ των προτέρων κατανομής

Από τα αποτελέσματα του παραπάνω πίνακα και από το διάγραμμα 3.2 συγκρίνοντας τις εκτιμήσεις θ_1 και θ_2 , μπορούμε να συμπεράνουμε πως, για δεδομένη τιμή της παραμέτρου α , όσο μεγαλύτερη τιμή έχει η παράμετρος β τόσο μικρότερη θα είναι η εκτίμηση της πιθανότητας αθέτησης. Αντιστρόφως, μέσω σύγκρισης των τιμών θ_2 και θ_3 προκύπτει πως, αν δώσουμε μεγάλη τιμή στην παράμετρο α , αυτό θα οδηγήσει στην εκτίμηση μεγαλύτερης τιμής για την πιθανότητα αθέτησης.

Επιπλέον, παρατηρούμε ότι μέσω του εκτιμητή μέγιστης πιθανοφάνειας λαμβάνουμε τις πιο μικρές εκτιμήσεις. Γενικά ο εκτιμητής μέγιστης πιθανοφάνειας μπορεί να θεωρηθεί ένας καλός τρόπος εκτίμησης, κυρίως σε μεγάλα δείγματα, αλλά απαραίτητη προϋπόθεση είναι να έχουν υπάρξει αθετήσεις στο δείγμα. Επίσης, όπως ήταν αναμενόμενο, εύκολα μπορούμε να παρατηρήσουμε ότι όσο μεγαλύτερος είναι ο αριθμός του δείγματος, οι εκ των υστέρων μέσες τιμές είναι πιο κοντά στον εκτιμητή μέγιστης πιθανοφάνειας από ό, τι στην εκ των προτέρων μέση τιμή.

Τέλος, να επισημάνουμε ότι εφόσον για το κάθε χαρτοφυλάκιο χρησιμοποιήσαμε το σύνολο των οφειλετών και τον συνολικό αριθμό των αθετήσεων, τα αποτελέσματα που παραθέτουμε σε αυτήν την Ενότητα μπορούν να συγκριθούν με τα αποτελέσματα της Ενότητας 2.2 για τη βαθμίδα Α. Κάνοντας αυτήν την σύγκριση,

βλέπουμε πως οι εκτιμήσεις μας δεν ξεπερνούν σε καμία περίπτωση τις τιμές που λάβαμε μέσω της προσέγγισης των Pluto και Tasche (2005) για $\gamma=90\%$.



Σχήμα 3.2 Πιθανότητα αθέτησης με χρήση Beta εκ των προτέρων κατανομής

Μία εκ των προτέρων κατανομή που έχει προταθεί από τον Kiefer (2009) και αποτελεί μία τροποποίηση της κλασσικής $Beta(\alpha, \beta)$ είναι η $Beta(\alpha, \beta)$ με $p_{min} < \theta < p_{max}$.

Υποθέτουμε ότι :

$$T \sim Beta(\alpha, \beta)$$

και

$$\theta = p_{min} + (p_{max} - p_{min})T . \quad (3.19)$$

Από την (3.19) προκύπτει ότι :

$$T = \frac{\theta - p_{min}}{p_{max} - p_{min}} . \quad (3.20)$$

Τότε, η συνάρτηση πυκνότητας πιθανότητας της Θ δίνεται από τη σχέση:

$$\pi(\theta) = f_T(t(\theta))|t'(\theta)| \quad (3.21)$$

όπου

$$t(\theta) = \frac{\theta - p_{min}}{p_{max} - p_{min}} .$$

Επομένως

$$\begin{aligned} \pi(\theta|\alpha, \beta, p_{min}, p_{max}) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)(p_{max}-p_{min})} \left(\frac{\theta-p_{min}}{p_{max}-p_{min}}\right)^{\alpha-1} \left(1 - \frac{\theta-p_{min}}{p_{max}-p_{min}}\right)^{\beta-1} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)(p_{max}-p_{min})} \left(\frac{p_{min}-\theta}{p_{min}-p_{max}}\right)^{\alpha-1} \left(\frac{\theta-p_{max}}{p_{min}-p_{max}}\right)^{\beta-1} . \end{aligned} \quad (3.22)$$

Μία ακόμα γενίκευση της παραπάνω περίπτωσης που προτάθηκε από τον Kiefer (2009) είναι η μίξη δύο κατανομών $Beta$ με $p_{min} < \theta < p_{max}$. Σε αυτήν την περίπτωση, η εκ των προτέρων κατανομή της θ εκφράζεται μέσω μίας συνάρτησης με 7 παραμέτρους. Θεωρούμε ότι οι συναρτήσεις $Beta$ έχουν παραμέτρους a_1, b_1 και a_2, b_2 και βάρη μίξης λ και $1 - \lambda$ αντίστοιχα. Σε αυτό το σημείο, η εκ των προτέρων συνάρτηση πυκνότητας πιθανότητας γράφεται :

$$\begin{aligned} \pi(\theta | \alpha_1, \beta_1, \alpha_2, \beta_2, p_{min}, p_{max}) &= \\ &= \lambda \frac{\Gamma(\alpha_1 + \beta_1)}{\Gamma(\alpha_1)\Gamma(\beta_1)(p_{max} - p_{min})} \left(\frac{p_{min} - \theta}{p_{min} - p_{max}}\right)^{\alpha_1 - 1} \left(\frac{\theta - p_{max}}{p_{min} - p_{max}}\right)^{\beta_1 - 1} \\ &+ (1 - \lambda) \frac{\Gamma(\alpha_2 + \beta_2)}{\Gamma(\alpha_2)\Gamma(\beta_2)(p_{max} - p_{min})} \left(\frac{p_{min} - \theta}{p_{min} - p_{max}}\right)^{\alpha_2 - 1} \left(\frac{\theta - p_{max}}{p_{min} - p_{max}}\right)^{\beta_2 - 1}. \end{aligned} \quad (3.23)$$

Από τις σχέσεις (3.9) και (3.23) μπορεί να υπολογιστεί η εκ των υστέρων κατανομή του δείγματος και μέσω αυτής να εκτιμηθεί η μέση τιμή της τυχαίας μεταβλητής θ .

Προφανώς, η διαδικασία που περιγράφηκε παραπάνω, μπορεί να επεκταθεί, προκειμένου η εκ των προτέρων κατανομή της θ να εκφραστεί ως μίξη των παραπάνω κατανομών $Beta$. Επίσης, όσες πιο πολλές κατανομές $Beta$ χρησιμοποιηθούν τόσο πιο ακριβής θα είναι η εκ των προτέρων κατανομή. Ωστόσο, πρέπει να επισημάνουμε πως ακόμα και με μίξη 2 κατανομών είναι πρακτικά αδύνατο να έχουμε αρκετές πληροφορίες, ώστε να μπορέσουμε να εκφράσουμε την εκ των προτέρων κατανομή σε μία μορφή σαν την (3.23).

3.4.2 Η περίπτωση της συντηρητικής εκ των προτέρων κατανομής

Ο Tasche (2013) πρότεινε ως αθροιστική συνάρτηση της εκ των προτέρων κατανομής της πιθανότητας αθέτησης θ την ακόλουθη:

$$P(\theta \leq \theta) = \int_0^\theta \frac{du}{1-u} = -\log(1 - \theta), 0 < \theta < 1 \quad (3.24)$$

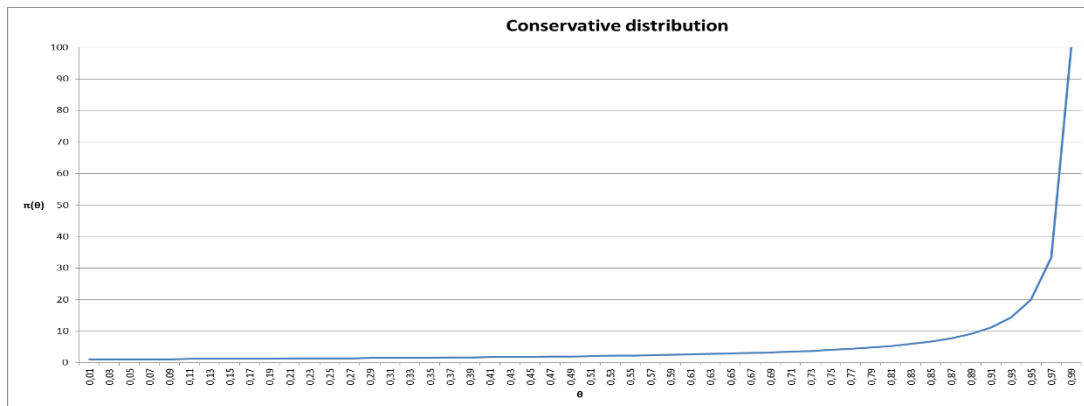
έτσι ώστε, η συνάρτηση πυκνότητας πιθανότητας να είναι :

$$\pi(\theta) = \frac{1}{1-\theta}, 0 < \theta < 1. \quad (3.25)$$

Εδώ, να τονίσουμε ότι κακώς η (3.25) αναφέρεται ως συνάρτηση πυκνότητας πιθανότητας, καθώς $\int_0^1 \pi(\theta) d\theta = \infty$. Παρότι όμως, δεν αποτελεί συνάρτηση πυκνότητας πιθανότητας δύναται να χρησιμοποιηθεί ως τέτοια, διότι στην προσέγγισή μας θα χρησιμοποιηθεί μόνο για να καθορίσει τις παραμέτρους της εκ των υστέρων κατανομής της τυχαίας μεταβλητής που μας αφορά.

Η εν λόγω κατανομή καλείται συντηρητική (conservative) εκ των προτέρων κατανομή, αφού όπως διακρίνεται και στο διάγραμμα για μεγάλες τιμές του θ αντιστοιχούν μεγάλες τιμές στην $\pi(\theta)$. Οι τιμές αυτές, όπως προαναφέραμε δεν είναι

πιθανότητες, αλλά είναι ένα μέτρο που μας υποδεικνύει ότι είναι πιο πιθανό να <<βρεθούμε>> σε μεγάλες τιμές του θ .



Σχήμα 3.3 Γραφική παράσταση της συντηρητικής εκ των προτέρων κατανομής

Η εκ των υστέρων κατανομή της πιθανότητας αθέτησης υπολογίζεται βάσει των σχέσεων (3.9),(3.14) και (3.25) :

$$f(\theta|d_1, \dots, d_n) \propto \left[\prod_{i=1}^n f(d_i/\theta) \right] \pi(\theta) = \theta^r (1 - \theta)^{n-r} * \frac{1}{1 - \theta} =$$

$$= \theta^{(r+1)-1} (1 - \theta)^{n-r-1} . \quad (3.26)$$

Από την (3.26) είναι εμφανές ότι η εκ των υστέρων κατανομή της πιθανότητας αθέτησης είναι η *Beta* με παραμέτρους

$$\alpha = r + 1 \text{ και } \beta = n - r .$$

Λαμβάνοντας τη μέση τιμή ως εκτίμηση της πιθανότητας αθέτησης έχουμε:

$$\hat{\theta} = E(\theta) = \frac{r+1}{n+1} . \quad (3.27)$$

Στον παρακάτω πίνακα θα παραθέσουμε τους εκτιμητές $\hat{\theta}$ της συντηρητικής εκ των προτέρων κατανομής, τον εκτιμητή θ_1 της Ενότητας 3.4.1, τη συντηρητική εκτίμηση για $\gamma=50\%$, έστω $\theta_{0,5}$ της Ενότητας 2.2 και τον εκτιμητή μέγιστης πιθανοφάνειας.

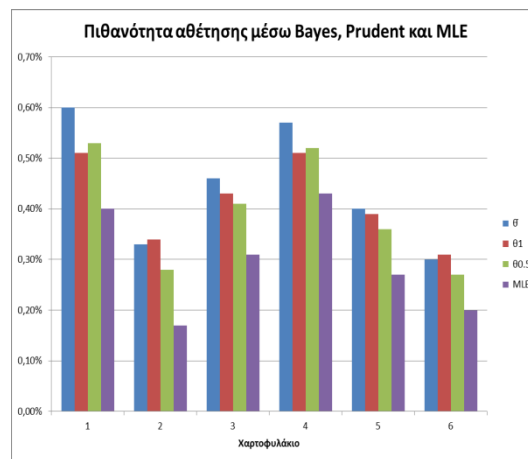
| n | r | $\hat{\theta}$ | θ_1 | $\theta_{0,5}$ | MLE |
|------|---|----------------|------------|----------------|-------|
| 500 | 2 | 0.60% | 0.51% | 0.53% | 0.40% |
| 600 | 1 | 0.33% | 0.34% | 0.28% | 0.17% |
| 650 | 2 | 0.46% | 0.43% | 0.41% | 0.31% |
| 700 | 3 | 0.57% | 0.51% | 0.52% | 0.43% |
| 750 | 2 | 0.40% | 0.39% | 0.36% | 0.27% |
| 1000 | 2 | 0.30% | 0.31% | 0.27% | 0.20% |

Πίνακας 3.2 Εκτίμηση της πιθανότητας αθέτησης μέσω της συντηρητικής εκ των προτέρων κατανομής, της κατανομής *Beta*, της συντηρητικής αρχής εκτίμησης και του εκτιμητή μέγιστης πιθανοφάνειας

Βάσει του παραπάνω πίνακα, παρατηρούμε ότι οι εκτιμήσεις που λάβαμε με τη χρήση της $\pi(\theta) = \frac{1}{1-\theta}$ είναι πολύ κοντά με τις τιμές που υπολογίσαμε, υποθέτοντας

ότι η εκ των προτέρων κατανομή είναι η $Beta(2.1,300)$. Σε σύγκριση με τον εκτιμητή μέγιστης πιθανοφάνειας, οι τιμές $\hat{\theta}$ είναι ελαφρώς μεγαλύτερες όπως και ήταν αναμενόμενο. Επιπλέον, η εκτίμηση $\hat{\theta}$ παρουσιάζει επίσης “ευαισθησία” στον αριθμό των αθετήσεων, ενώ αντιθέτως με τον MLE δύναται να εκτιμήσει την πιθανότητα αθέτησης, ακόμα και στην περίπτωση που δεν έχει υπάρξει καμία αθέτηση στο χαρτοφυλάκιο.

Τέλος συγκρίνοντας τα αποτελέσματα με τον εκτιμητή $\theta_{0,5}$ κατανοούμε ότι οι τιμές που λάβαμε είναι αρκετά κοντά. Στο Σχήμα 3.4 απεικονίζονται γραφικά για κάθε χαρτοφυλάκιο τα αποτελέσματα του Πίνακα 3.2.



Σχήμα 3.4 Πιθανότητα αθέτησης με conservative prior, Beta prior, prudent estimation και MLE

3.4.3 Η ομοιόμορφη εκ των προτέρων κατανομή

Όπως ήδη αναφέραμε, η ομοιόμορφη στο διάστημα $(0,1)$ αποτελεί μία μη πληροφοριακή εκ των προτέρων κατανομή. Αντί λοιπόν αυτής, μπορεί να χρησιμοποιηθεί η ομοιόμορφη στο διάστημα (u,l) :

$$\pi(\theta) = \frac{1}{u-l}, 0 \leq l \leq \theta \leq u \leq 1. \quad (3.28)$$

Ο Tasche (2013) χρησιμοποιεί τη συγκεκριμένη κατανομή με $l = 0$. Πρόκειται για έναν αρκετά λογικό περιορισμό της εκ των προτέρων κατανομής, καθώς αναμένουμε αρκετά μικρή εκτιμώμενη τιμή για την πιθανότητα αθέτησης. Σε αυτή την περίπτωση η τυχαία μεταβλητή Θ ακολουθεί την Ομοιόμορφη $(0,u)$. Μία ακόμα λογική υπόθεση είναι ότι το u είναι “μικρό”. Σε αυτή την περίπτωση η (3.28) γράφεται :

$$\pi(\theta) = \frac{1}{u}, 0 \leq \theta \leq u. \quad (3.29)$$

Στην παρούσα φάση, μπορούμε να υπολογίσουμε την εκτίμηση της πιθανότητας αθέτησης μέσω της εκ των υστέρων μέσης τιμής της τυχαίας μεταβλητής Θ . Από τη θεωρία γνωρίζουμε ότι η εκ των ζητούμενη μέση τιμή δίνεται από τη σχέση:

$$E(\Theta) = \frac{\int_0^u \theta f(D/\theta) \pi(\theta) d\theta}{\int_0^u f(D/\theta) \pi(\theta) d\theta}. \quad (3.30)$$

Από τις σχέσεις (3.14),(3.29) και (3.30) :

$$\frac{\int_0^u \theta^{r+1}(1-\theta)^{n-r} d\theta}{\int_0^u \theta^r(1-\theta)^{n-r} d\theta} = \frac{\Gamma(n+2)\Gamma(r+2)}{\Gamma(n+3)\Gamma(r+1)} * \frac{F_{B_1}(u)}{F_{B_2}(u)} = \frac{r+1}{n+2} * \frac{F_{B_1}(u)}{F_{B_2}(u)} \quad (3.31)$$

όπου $B_1 \sim \text{Beta}(r+2, n-r+1)$, $B_2 \sim \text{Beta}(r+1, n-r+1)$. Για την ειδική περίπτωση $u = 1$, η εκ των προτέρων κατανομή είναι η Ομοιόμορφη(0,1). Τότε, εύκολα προκύπτει από την (3.31) πως η εκ των υστέρων μέση τιμή είναι $\frac{r+1}{n+2}$. Στον παρακάτω πίνακα παρατίθενται οι εκτιμήσεις μέσω της σχέσης (3.31) για $u = 0.025, u = 0.05, u = 0.1$ και $u = 1$, καθώς και οι $\hat{\theta}, \theta_1, \theta_{0,5}$ και MLE που είχαν αναφερθεί και στις προηγούμενες ενότητες.

| Εκτιμητής | $n=500, r=2$ | $n=600, r=1$ | $n=650, r=2$ | $n=700, r=3$ | $n=750, r=2$ | $n=1000, r=2$ |
|----------------|--------------|--------------|--------------|--------------|--------------|---------------|
| (0,0.025) | 0.5969% | 0.3322% | 0.4601% | 0.5697% | 0.3989% | 0.2994% |
| (0,0.05) | 0.5976% | 0.3322% | 0.4601% | 0.5698% | 0.3989% | 0.2994% |
| (0,0.1) | 0.5976% | 0.3322% | 0.4601% | 0.5698% | 0.3989% | 0.2994% |
| (0,1) | 0.5976% | 0.3322% | 0.4601% | 0.5698% | 0.3989% | 0.2994% |
| $\hat{\theta}$ | 0.60% | 0.33% | 0.46% | 0.57% | 0.40% | 0.30% |
| θ_1 | 0.51% | 0.34% | 0.43% | 0.51% | 0.39% | 0.31% |
| $\theta_{0,5}$ | 0.53% | 0.28% | 0.41% | 0.52% | 0.36% | 0.27% |
| MLE | 0.40% | 0.17% | 0.31% | 0.43% | 0.27% | 0.20% |

Πίνακας 3.3 Εκτίμηση της πιθανότητας αθέτησης με την ομοιόμορφη, συντηρητική, $\text{Beta}(2.1,300)$, $\theta_{0,5}$, MLE

Όπως είναι εμφανές, η εκτίμηση με χρήση της ομοιόμορφης για τα επιλεγμένα u σαν εκ των προτέρων κατανομή είναι πανομοιότυπη με την εκτίμηση της ενότητας 3.4.2 και ως εκ τούτου παρουσιάζει τα ίδια χαρακτηριστικά με την $\hat{\theta}$ ως προς τις εκτιμήσεις $\theta_1, \theta_{0,5}$ και MLE.

3.5 Εξάρτηση του δείγματος και περίοδος παρατήρησης άνω του ενός έτους

Όπως αναλύθηκε και στο 2^ο Κεφάλαιο η εισαγωγή ενός συστηματικού παράγοντα X μας δίνει την δυνατότητα να δημιουργήσουμε εξάρτηση μεταξύ των αθετίσεων. Σε αυτή την περίπτωση η πιθανότητα αθέτησης δοθείσας μιας τιμής x για τον συστηματικό παράγοντα X δίνεται από την σχέση (2.15) :

$$P[D_i \leq \Phi^{-1}(p)] = P \left[Y_i \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}x}{\sqrt{1-\rho}} \right] = \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}x}{\sqrt{1-\rho}} \right).$$

Από την παραπάνω σχέση προκύπτει ότι η συνάρτηση πιθανοφάνειας είναι:

$$f(D/\theta) = g(\rho, p_A, x)^r \left(1 - g(\rho, p_A, x) \right)^{n-r} \quad (3.32)$$

όπου, $g(\rho, p_A, x) = \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}x}{\sqrt{1-\rho}} \right)$. Συνδυάζοντας τις σχέσεις (3.9) και (3.32) προκύπτει πως θεωρώντας ως εκ των προτέρων κατανομή την $\text{Beta}(\alpha, \beta)$ η εκ των

υστέρων κατανομή της πιθανότητας αθέτησης δοθείσης μια τιμής x για τον συστηματικό παράγοντα X δίνεται από τη σχέση,

$$g(\rho, p, X)^r (1 - g(\rho, p, x))^{n-r} * p^{\alpha-1} * (1 - p)^{\beta-1}. \quad (3.33)$$

Αντίστοιχα, χρησιμοποιώντας την συντηρητική και την ομοιόμορφη εκ των προτέρων κατανομή λαμβάνουμε αντίστοιχα

$$g(\rho, p, x)^r (1 - g(\rho, p, x))^{n-r} * \left(\frac{1}{1-p}\right) \quad (3.34)$$

$$g(\rho, p, x)^r (1 - g(\rho, p, x))^{n-r} * \frac{1}{u}. \quad (3.35)$$

Όπως και στο Κεφάλαιο 2, κάνοντας τις ίδιες υποθέσεις μπορούμε να επεκτείνουμε το μοντέλο στην περίπτωση της πολυετούς περιόδου παρατήρησης. Σε αυτή την περίπτωση, θεωρώντας T χρόνια παρατήρησης και μια τιμή x για τον συστηματικό παράγοντα X , η εκ των υστέρων κατανομή της πιθανότητας αθέτησης στην περίπτωση της $Beta(\alpha, \beta)$, συντηρητικής κατανομής και της ομοιόμορφης κατανομής δίνεται από τις επόμενες σχέσεις

$$\prod_{t=1}^T g_t(\rho, p, x_t)^{r_t} (1 - g(\rho, p, x_t))^{n_t-r_t} * p^{\alpha-1} * (1 - p)^{\beta-1} \quad (3.36)$$

$$\prod_{t=1}^T g_t(\rho, p, x_t)^{r_t} (1 - g(\rho, p, x_t))^{n_t-r_t} * \left(\frac{1}{1-p}\right) \quad (3.37)$$

$$\prod_{t=1}^T g_t(\rho, p, x_t)^{r_t} (1 - g(\rho, p, x_t))^{n_t-r_t} * \left(\frac{1}{u}\right). \quad (3.38)$$

όπου n_t και r_t είναι ο αριθμός των οφειλετών και ο αριθμός των αθετήσεων κατά την χρονική περίοδο t αντίστοιχα.

3.6 Σχολιασμός της μεθόδου

Όπως γίνεται σαφές η εκτίμηση της πιθανότητας αθέτησης μέσω της προσέγγισης Bayes προσφέρει το πλεονέκτημα της χρήσης διάφορων εκ των προτέρων κατανομών. Με αυτό τον τρόπο ο ειδικός δύναται να “εκφράσει” τις προσδοκίες του για την πιθανότητα αθέτησης με την επιλογή κάποιας συγκεκριμένης κατανομής. Επίσης όπως είναι εμφανές από του Πίνακες 3.1, 3.2 και 3.3 οι εκτιμήσεις που λαμβάνουμε με αυτή την μέθοδο είναι σαφώς λιγότερο συντηρητικές από ότι με την μέθοδο που αναλύθηκε στο Κεφάλαιο 2 ενώ και με τις τρεις κατανομές που εξετάσαμε μπορεί να εκτιμηθεί η πιθανότητα αθέτησης ακόμα και στην περίπτωση που δεν έχουν υπάρξει αθετήσεις στο δείγμα. Ωστόσο τα αποτελέσματα που θα λάβουμε μπορεί να είναι αναξιόπιστα στην περίπτωση που ο ειδικός δεν είναι αμερόληπτος στην επιλογή της εκ των προτέρων κατανομής. Ένα επιπλέον μειονέκτημα της συγκεκριμένης μεθόδου είναι πως δεν χρησιμοποιούνται διαστήματα εμπιστοσύνης για την εκτίμηση της πιθανότητας αθέτησης. Τέλος όπως και στο Κεφάλαιο 2 αν θεωρήσουμε εξάρτηση του δείγματος και έχουμε πολυετή περίοδο παρατήρησης οι υπολογισμοί γίνονται σαφώς πιο δύσκολοι ενώ το πιστωτικό ίδρυμα οφείλει να εκτιμήσει την συσχέτιση μεταξύ των συστηματικών παραγόντων X_1, \dots, X_T και την συσχέτιση μεταξύ των οφειλετών ρ .

4. Η καμπύλη ισχύος

Σε αυτό το Κεφάλαιο θα αναφερθούμε στην εκτίμηση πιθανοτήτων αθέτησης μέσω της καμπύλης CAP (Cumulative Accuracy Profile), η οποία είναι γνωστή και ως καμπύλη ισχύος (power curve). Γενικά μέσω της καμπύλης CAP γίνεται αξιολόγηση της διακριτικής ισχύος ενός μοντέλου πιστοληπτικής αξιολόγησης. Παρόλα αυτά, με την μεθοδολογία που προτείνει ο Burgt(2008) μπορούν να προκύψουν μέσω της καμπύλης CAP και οι πιθανότητες αθέτησης για ένα χαρτοφυλάκιο.

4.1 Η κατασκευή της καμπύλης CAP

Θεωρούμε ένα μοντέλο αξιολόγησης το οποίο αποδίδει σε κάθε οφειλέτη μία βαθμολογία s_i από k πιθανές τιμές $\{s_1, s_2, \dots, s_k\}$ με $s_1 < s_2 < \dots < s_k$. Μία υψηλή βαθμολογία υποδεικνύει χαμηλή πιθανότητα αθέτησης. Εκχωρούμε λοιπόν βαθμολογίες στους οφειλέτες ενός χαρτοφυλακίου και ελέγχουμε αν θα αθετήσουν τις υποχρεώσεις τους ή θα παραμείνουν φερέγγυοι κατά την επόμενη χρονική περίοδο. Για το λόγο αυτό εισάγουμε τις τυχαίες μεταβλητές S_T, S_D και S_{ND} οι οποίες αφορούν τις βαθμολογίες όλων των οφειλετών, των οφειλετών οι οποίοι δεν πλήρωσαν τις υποχρεώσεις τους και των οφειλετών που παρέμειναν φερέγγυοι, αντίστοιχα. Συμβολίζουμε με $p_D^{(i)}$ την πιθανότητα ένας οφειλέτης που αθέτησε να έχει βαθμολογία s_i . Αντίστοιχα ορίζουμε τις $p_T^{(i)}$ και $p_{ND}^{(i)}$.

Επίσης ορίζουμε,

$$CD_T^{(i)} = \sum_{j=1}^i p_T^{(j)}, i = 1, \dots, k \quad (4.1)$$

$$CD_D^{(i)} = \sum_{j=1}^i p_D^{(j)}, i = 1, \dots, k \quad (4.2)$$

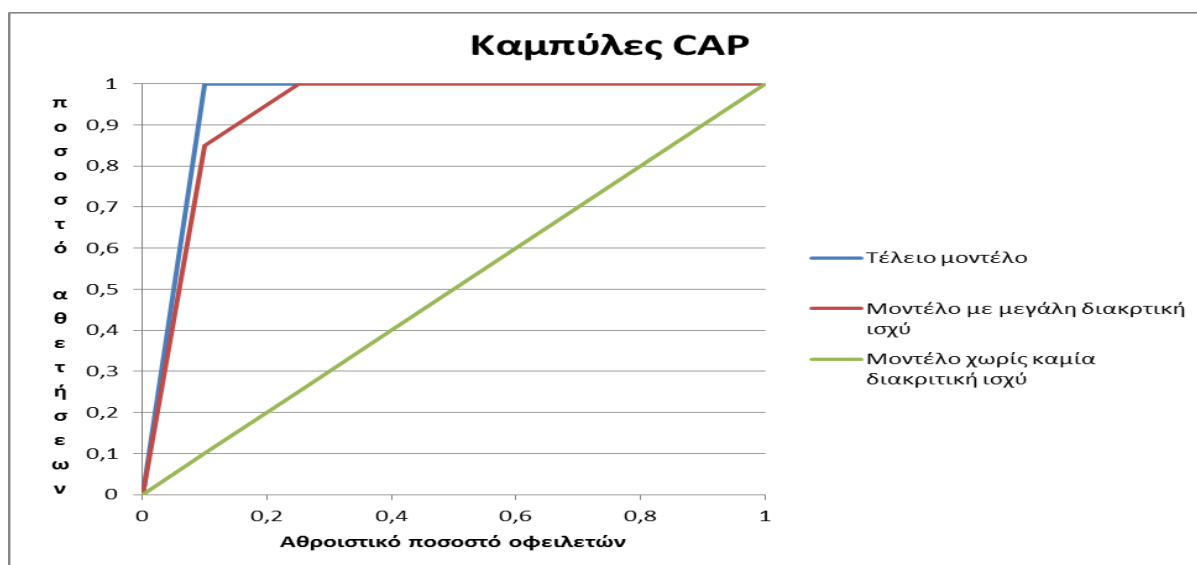
$$p_{ND}^{(i)} = \sum_{j=1}^i p_{ND}^{(j)}, i = 1, \dots, k. \quad (4.3)$$

Προφανώς τα CD_T, CD_D και CD_{ND} εκφράζουν τις αθροιστικές συναρτήσεις κατανομής των βαθμολογιών για όλους τους οφειλέτες, τους οφειλέτες που αθέτησαν τις υποχρεώσεις τους και τους οφειλέτες που παρέμειναν φερέγγυοι, αντίστοιχα. Για παράδειγμα το CD_D^i δηλώνει την πιθανότητα ένας οφειλέτης που αθέτησε τις υποχρεώσεις του να έχει βαθμολογία το πολύ s_i . Ορίζουμε επίσης,

$$CD_T^0 = CD_D^0 = CD_{ND}^0 = 0. \quad (4.4)$$

Η καμπύλη CAP αποτελεί τη γραφική παράσταση όλων των σημείων $(CD_T^{(i)}, CD_D^{(i)})$, $i = 0, \dots, k$. Όταν χρησιμοποιείται ένα τέλειο μοντέλο αξιολόγησης όλες οι αθετήσεις θα έχουν συμβεί στην κατηγορία με την μικρότερη βαθμολογία ενώ σε ένα τυχαίο μοντέλο το οποίο δεν έχει καμία διακριτική ισχύ το σωρευτικό ποσοστό αθετήσεων αυξάνεται αναλογικά με το σωρευτικό ποσοστό των οφειλετών. Στο παρακάτω σχήμα μπορούμε να δούμε τις καμπύλες CAP για ένα τέλειο μοντέλο

αξιολόγησης, ένα μοντέλο αξιολόγησης με μεγάλη διακριτική ισχύ και ένα τυχαίο μοντέλο αξιολόγησης.



Σχήμα 4.1 Καμπύλες CAP για 3 διαφορετικά μοντέλα αξιολόγησης

4.2 Μοντελοποίηση της καμπύλης

Όπως μπορούμε εύκολα να διαπιστώσουμε από το Σχήμα 4.1 για ένα μοντέλο αξιολόγησης το οποίο έχει έστω και μια μικρή διακριτική ισχύ, η καμπύλη CAP έχει κοίλο σχήμα. Συμβολίζοντας με y το σωρευτικό ποσοστό αθετήσεων που έχουν συμβεί στο χαρτοφυλάκιο και με x το αθροιστικό ποσοστό οφειλετών ο Burgt(2008) προτείνει να μοντελοποιηθεί η καμπύλη CAP με την ακόλουθη συνάρτηση

$$y(x) = \frac{1-e^{-kx}}{1-e^{-k}} \quad (4.5)$$

Η παράμετρος k ονομάζεται κοιλότητα και αποτελεί ένα μέτρο διακριτικής ισχύος. Από την (4.5) προκύπτουν τα εξής,

- α) $x = 0 \Rightarrow y = 0$ δηλαδή όταν το σωρευτικό ποσοστό οφειλετών είναι 0% τότε το σωρευτικό ποσοστό των αθετήσεων είναι επίσης 0%
- β) $x = 1 \Rightarrow y = 1$. Όταν το σωρευτικό ποσοστό οφειλετών είναι 100% τότε και το σωρευτικό ποσοστό των αθετήσεων είναι 100%
- γ) $y(x) \rightarrow 1$ όταν $k \rightarrow \infty$ και η CAP αντιστοιχεί σε ένα τέλειο σύστημα αξιολόγησης.
- δ) $y(x) \rightarrow x$ όταν $k \rightarrow 0$ και η CAP αντιστοιχεί σε ένα τυχαίο σύστημα αξιολόγησης.

Για να αποδείξουμε το (δ) αρκεί να εφαρμόσουμε τον κανόνα de L' Hospital.

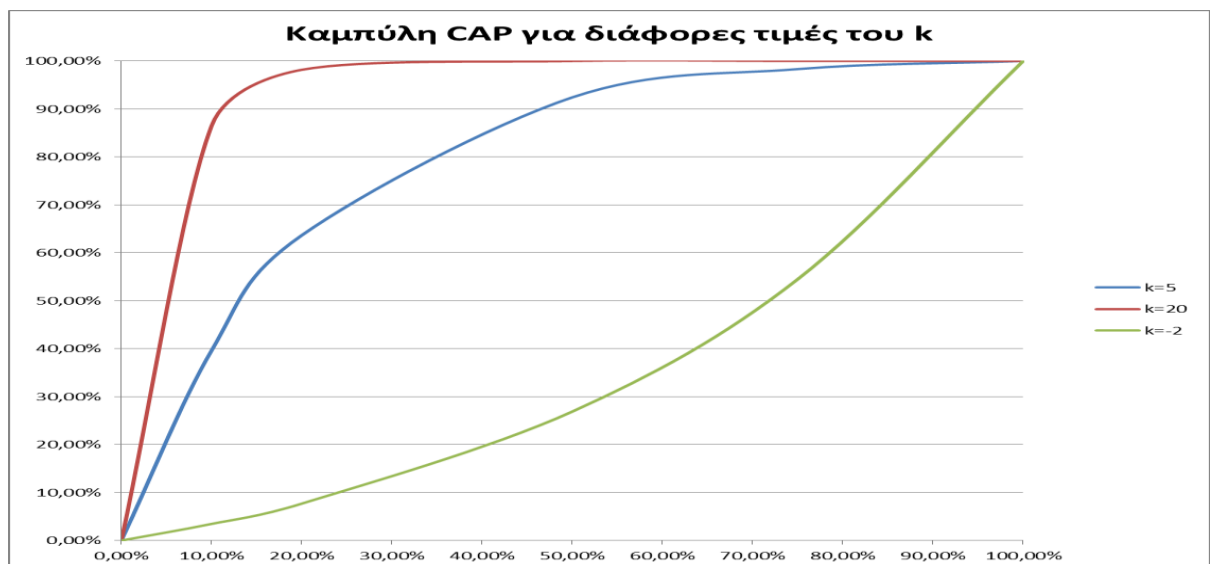
Τότε,

$$\lim_{k \rightarrow 0} \frac{(1-e^{-kx})'}{(1-e^{-k})'} = \lim_{k \rightarrow 0} \frac{x e^{-kx}}{e^{-k}} = x \quad (4.6)$$

Από το πρόσημο της 2^{ης} παραγώγου προκύπτει πως για $k > 0$ η καμπύλη CAP έχει κοίλο σχήμα και το μοντέλο αξιολόγησης έχει κάποια διακριτική ισχύ. Στην περίπτωση που $k < 0$ το μεγαλύτερο ποσοστό των αθετήσεων συμβαίνει στις καλύτερες κατηγορίες. Πρακτικά αυτό σημαίνει πως για μια αρνητική τιμή του k το σύστημα αξιολόγησης εκχωρεί υψηλές βαθμολογίες σε οφειλέτες που αποδεικνύονται “κακοί” και χαμηλές βαθμολογίες σε πελάτες που είναι “καλοί”. Στο ακόλουθο σχήμα βλέπουμε τη μορφή που παίρνει η καμπύλη CAP για

$$k = 5, k = 20 \text{ και } k = -2 .$$

Όπως είναι εμφανές από το Σχήμα 4.2 όσο μεγαλύτερη είναι η τιμή k τόσο μεγαλύτερη διακριτική ισχύ έχει το μοντέλο αξιολόγησης ενώ για αρνητική τιμή του k το χαρτοφυλάκιο αντιμετωπίζει τον κίνδυνο της αντεπιλογής (adverse selection) .



Σχήμα 4.2 Καμπύλη CAP και $k=5, k=20$ και $k=-2$

Πριν προχωρήσουμε σε περαιτέρω ανάλυση του μοντέλου να επισημάνουμε πως επειδή η καμπύλη CAP μπορεί να χρησιμοποιηθεί για την εκτίμηση της πιθανότητας αθέτησης σε οποιοδήποτε χαρτοφυλάκιο, η μέθοδος μπορεί να επεκταθεί και σε μια πολυπαραμετρική μορφή. Σε αυτή την περίπτωση η καμπύλη CAP μοντελοποιείται με την εξής μορφή

$$y(x) = m \frac{1-e^{-kx}}{1-e^{-k}} + (1-m) \frac{1-e^{-lx}}{1-e^{-l}}, \quad 0 \leq m \leq 1. \quad (4.7)$$

Τα k και l αποτελούν παραμέτρους κοιλότητας ενώ τα m και $1-m$ είναι βάρη μίξης. Η παραπάνω μορφή είναι κατάλληλη μόνο αν έχουμε αρκετά στοιχεία για να καθορίσουμε και τις 3 παραμέτρους. Πρακτικά αυτό σημαίνει πως πρέπει να έχει υπάρξει ένας υψηλός αριθμός αθετήσεων στο χαρτοφυλάκιο που μελετάμε. Στο επόμενο σχήμα βλέπουμε την γραφική παράσταση της $y(x)$ που δίνεται από τη σχέση (4.7).



Σχήμα 4.3 Πολυπαραμετρική καμπύλη CAP

Εφόσον στην παρούσα Δ.Ε αντιμετωπίζουμε το πρόβλημα της εκτίμησης της πιθανότητας αθέτησης σε χαρτοφυλάκια υψηλής φερεγγυότητας, όπου συνήθως ο αριθμός των αθετήσεων είναι ελάχιστος δεν θα ασχοληθούμε περαιτέρω με την παραπάνω περίπτωση. Η διαδικασία όμως που θα ακολουθούσαμε είναι ανάλογη αυτής που θα αναφέρουμε παρακάτω.

4.3 Δείκτες αξιολόγησης και η κοιλότητα της καμπύλης

Όπως αναφέραμε η τιμή της παραμέτρου k αποτελεί ένα μέτρο αξιολόγησης της διακριτικής ισχύος του μοντέλου αξιολόγησης, χωρίς ωστόσο να είναι το μόνο. Το κυριότερο μέτρο που χρησιμοποιούμε για να δούμε την ποιότητα του μοντέλου είναι ο Δείκτης Ακριβείας (Accuracy Ratio) ή AR . Ο δείκτης αυτός ορίζεται ως

$$AR = \frac{a_R}{a_P} \quad (4.8)$$

όπου, το a_R είναι το εμβαδόν του χωρίου που περικλείεται μεταξύ του πραγματικού μοντέλου αξιολόγησης και του μοντέλου χωρίς διακριτική ισχύ ενώ a_P είναι το εμβαδόν του χωρίου που περικλείεται μεταξύ του τέλειου μοντέλου και του τυχαίου μοντέλου. Κατά αυτόν τον τρόπο, όσο πιο κοντά στον αριθμό 1 είναι το AR τόσο καλύτερο είναι το μοντέλο που χρησιμοποιούμε.

Ένα ακόμα μέτρο ποιότητας του μοντέλου αξιολόγησης είναι το εμβαδόν της περιοχής που βρίσκεται κάτω από την καμπύλη. Το μέτρο αυτό συμβολίζεται με AUC (Area Under Curve). Προφανώς, όσο μεγαλύτερη είναι η τιμή του AUC τόσο καλύτερο είναι το μοντέλο αξιολόγησης. Στην πράξη, αν θεωρήσουμε πως το σύστημα αξιολόγησης δεν μας οδηγεί στην επιλογή “κακών” κινδύνων ισχύει

$$0.5 \leq AUC \leq 1 . \quad (4.9)$$

Η παραπάνω σχέση ισχύει καθώς για το τέλειο μοντέλο έχουμε $AUC = 1$ ενώ για το τυχαίο ισχύει $AUC = 0.5$.

Εφόσον τα AR , AUC , k είναι μέτρα ποιότητας για το μοντέλο αξιολόγησης είναι αναμενόμενο να υπάρχουν μαθηματικές σχέσεις μεταξύ τους. Όπως αποδεικνύεται στους Engelman et al (2003), το AR συνδέεται με το AUC με την παρακάτω σχέση,

$$AR = 2AUC - 1 . \quad (4.10)$$

Προκειμένου να βρούμε εκτίμηση για το k θα υπολογίσουμε το AUC . Από τον ορισμό που δώσαμε προηγουμένως είναι προφανές πως,

$$AUC = \int_0^1 y(x) dx . \quad (4.11)$$

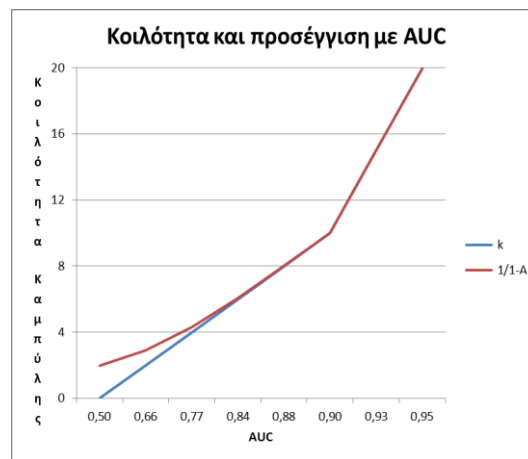
Υπολογίζοντας το δεξί μέλος της (4.11) παίρνουμε την παρακάτω σχέση:

$$AUC = \frac{1}{1-e^{-k}} - \frac{1}{k} , k \neq 0 . \quad (4.12)$$

Προφανώς, αν $k = 0 \Rightarrow y(x) = x$ και $AUC = 0.5$. Όπως παρατηρούμε από την (4.12) ο όρος $\frac{1}{1-e^{-k}}$ συγκλίνει στην μονάδα όταν το k παίρνει μεγάλες τιμές, διότι το e^{-k} συγκλίνει γρήγορα στο 0. Βάσει αυτής της παρατήρησης ο Burgt(2008) προτείνει την εκτίμηση του k από την σχέση,

$$AUC \approx 1 - \frac{1}{k} \Leftrightarrow k \approx \frac{1}{1-AUC} . \quad (4.13)$$

Στο επόμενο διάγραμμα βλέπουμε πως όταν $AUC \geq 0.8$ η σχέση (4.13) δίνει μια πολύ καλή προσέγγιση για το k .



Σχήμα 4.4 Προσέγγιση της κοιλότητας για διάφορες τιμές του AUC

Συνδυάζοντας τις σχέσεις (4.10) και (4.12) παίρνουμε την ακόλουθη σχέση για τα AR και k ,

$$AR = 2 \left(\frac{1}{1-e^{-k}} - \frac{1}{k} \right) - 1 , k \neq 0 . \quad (4.14)$$

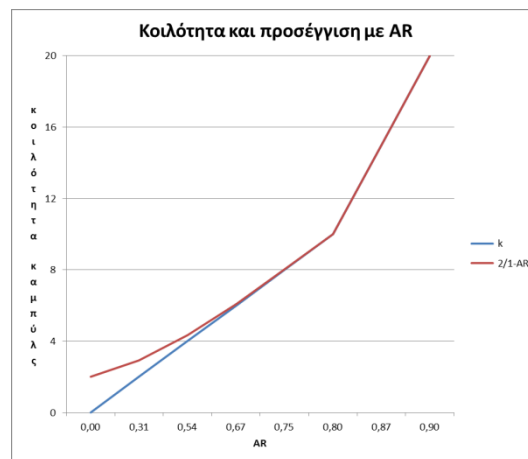
Θεωρώντας πάλι πως το $\frac{1}{1-e^{-k}}$ συγκλίνει πολύ γρήγορα στην μονάδα, παίρνουμε την ακόλουθη προσέγγιση,

$$AR \approx 2 \left(1 - \frac{1}{k}\right) - 1 \Leftrightarrow AR \approx 1 - \frac{2}{k}. \quad (4.15)$$

Η τελευταία σχέση μπορεί επίσης να γραφεί,

$$k \approx \frac{2}{1-AR}. \quad (4.16)$$

Οι τελευταίες σχέσεις δίνουν μια καλή προσέγγιση των AR και k όταν $AUC \geq 0.8$. Εφόσον ωστόσο στις σχέσεις αυτές δεν εμφανίζεται το AUC θα ήταν προτιμότερο να βρούμε μία τιμή του AR , πάνω από την οποία η σχέση (4.16) θα δίνει μια καλή προσέγγιση για την παράμετρο κοιλότητας. Από την σχέση (4.10) και γνωρίζοντας ότι έχουμε καλή προσέγγιση όταν $AUC \geq 0.8$ υπολογίζουμε εύκολα πως η σχέση (4.16) ενδείκνυται να χρησιμοποιείται όταν $AR \geq 0.6$. Στο παραπάνω αποτέλεσμα μπορούμε να καταλήξουμε και από το ακόλουθο διάγραμμα.



Σχήμα 4.5 Προσέγγιση της κοιλότητας για διάφορες τιμές του AR

4.4 Υπολογισμός της κοιλότητας και εκτίμηση της πιθανότητας αθέτησης

Προκειμένου να εκτιμήσουμε την πιθανότητα αθέτησης για ένα χαρτοφυλάκιο πρέπει πρώτα να εκτιμήσουμε την παράμετρο k . Για να το πετύχουμε αυτό αρκεί να βρούμε την τιμή του k η οποία ελαχιστοποιεί την τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος $RMSE$ (Root Mean Square Error). Για το $RMSE$ ισχύει,

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(y_i - \frac{1-e^{-kx_i}}{1-e^{-k}} \right)^2} \quad (4.17)$$

όπου, N είναι ο αριθμός των βαθμίδων πιστοληπτικής ικανότητας, τα x_i, y_i είναι τα παρατηρούμενα (observed) αθροιστικά ποσοστά οφειλετών και αθετήσεων στην βαθμίδα i αντίστοιχα και το $\frac{1-e^{-kx_i}}{1-e^{-k}}$ είναι το εκτιμώμενο (predicted) αθροιστικό ποσοστό αθετήσεων για την βαθμίδα i μέσω της (4.5). Ένας ακόμα τρόπος για να εκτιμήσουμε το k από την σχέση (4.15) είναι χρήση της μεθόδου Newton-Raphson όπου η τιμή για το k υπολογίζεται από την παρακάτω σχέση,

$$k_{i+1} = k_i - \left[\frac{\theta RMSE / \theta k}{\theta^2 RMSE / \theta^2 k} \right]_{k=k_i}, k_0 = 1. \quad (4.18)$$

Εφόσον έχει εκτιμηθεί μια τιμή για την κοιλότητα, η πιθανότητα αθέτησης για μία βαθμίδα πιστοληπτικής ικανότητας R , δίνεται από την σχέση:

$$PD_R = CT * \frac{dy}{dx} = CT \frac{ke^{-kx_R}}{1-e^{-k}}, R = 1, \dots, N. \quad (4.19)$$

Το CT είναι η παρατηρούμενη πιθανότητα αθέτησης του χαρτοφυλακίου ενώ το x_R υπολογίζεται ως,

$$x_R = \frac{\sum_{i=1}^{R-1} n_i + \frac{n_R}{2}}{n}, R = 1, \dots, N \quad (4.20)$$

Το n_i είναι ο αριθμός των οφειλετών που βρίσκονται στην βαθμίδα αξιολόγησης i , n είναι ο συνολικός αριθμός των οφειλετών που βρίσκονται στο χαρτοφυλάκιο ενώ θεωρούμε ότι για $R = 1$ έχουμε την βαθμίδα αξιολόγησης με τους πιο αφερέγγυους οφειλέτες.

Προκειμένου να παραθέσουμε αριθμητικά αποτελέσματα χρησιμοποιούμε το παρακάτω χαρτοφυλάκιο. Η βαθμίδα A είναι αυτή με την μικρότερη πιστοληπτική ικανότητα και η E αυτή με τη μεγαλύτερη.

| Χαρτοφυλάκιο | Οφειλέτες A βαθμίδας | Οφειλέτες B βαθμίδας | Οφειλέτες Γ βαθμίδας | Οφειλέτες Δ βαθμίδας | Οφειλέτες E βαθμίδας | Συνολικοί οφειλέτες |
|--------------|------------------------------|------------------------------|-----------------------------------|-----------------------------------|------------------------------|------------------------|
| 7 | 100 | 175 | 225 | 150 | 50 | 700 |

Στον επόμενο πίνακα δίνονται οι αθετήσεις που έχουν συμβεί για την εκάστοτε βαθμίδα του χαρτοφυλακίου.

| Χαρτοφυλάκιο | Αθετήσεις A βαθμίδας | Αθετήσεις B βαθμίδας | Αθετήσεις Γ βαθμίδας | Αθετήσεις Δ βαθμίδας | Αθετήσεις E βαθμίδας | Συνολικές αθετήσεις |
|--------------|------------------------------|------------------------------|-----------------------------------|-----------------------------------|------------------------------|------------------------|
| 7 | 3 | 2 | 0 | 0 | 0 | 5 |

Για το εν λόγω χαρτοφυλάκιο έχουμε $CT = \frac{5}{700}$ ενώ υπολογίζοντας το k μέσω της σχέσης (4.17) λαμβάνουμε τις τιμές $k = 6.96$ και $RMSE = 0.032$. Στο παρακάτω διάγραμμα μπορούμε να δούμε πως για $k = 6.96$ η καμπύλη $y(x)$ από την σχέση (4.5) ταιριάζει αρκετά καλά στην καμπύλη CAP που προκύπτει από τα δεδομένα μας.

Προτού εκτιμήσουμε τις πιθανότητες αθέτησης για τις βαθμίδες του χαρτοφυλακίου 7 θα υπολογίσουμε τους δείκτες AUC , AR και θα εκτιμήσουμε το k .

Από την σχέση (4.12) έχουμε $AUC = 0.86$ και από την (4.10) $AR = 0.71$.



Σχήμα 4.6 Προσαρμογή της καμπύλης $\gamma(x)$ για $k = 6.96$ στα δεδομένα του χαρτοφυλακίου 7

Αξίζει να αναφέρουμε ότι μπορούμε να εκτιμήσουμε το AUC χωρίς να έχουμε εκτιμήσει πρώτα το k . Για να γίνει αυτό αρκεί να χρησιμοποιήσουμε αριθμητική ολοκλήρωση και πιο συγκεκριμένα τον σύνθετο κανόνα του τραπεζίου. Σύμφωνα με τον κανόνα όταν για μια συνάρτηση $f(x)$ συνεχής σε ένα διάστημα $[a, b]$ θέλουμε να υπολογίσουμε το $\int_a^b f(x) dx$ και το διάστημα $[a, b]$ υποδιαιρείται σε N το πλήθος υποδιαστημάτων πλάτους $h = \frac{b-a}{N}$ από τα $N + 1$ το πλήθος σημεία,

$$x_i = a + ih \text{ με } i = 0, 1, \dots, N$$

ισχύει η παρακάτω σχέση,

$$\int_a^b f(x) dx \approx \frac{h}{2} [f(x_0) + f(x_1)] + \dots + \frac{h}{2} [f(x_{N-1}) + f(x_N)]. \quad (4.21)$$

Εφόσον θέλουμε να χρησιμοποιήσουμε την σχέση (4.21) για να εκτιμήσουμε το AUC για το χαρτοφυλάκιο 7 παίρνουμε:

$$AUC = \sum_{i=0}^4 \frac{1}{2} (x_{i+1} - x_i)(y_{i+1} + y_i), x_0 = y_0 = 0. \quad (4.22)$$

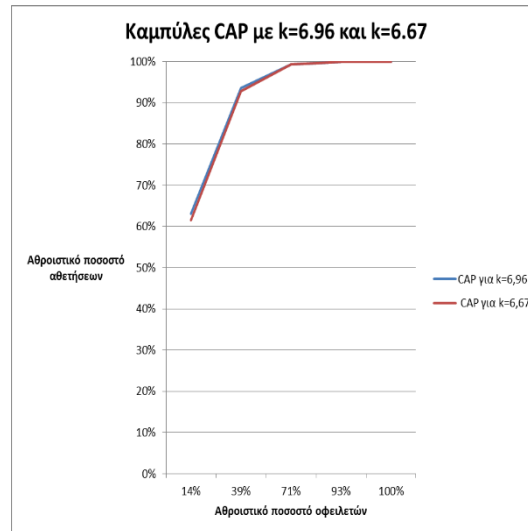
Στον επόμενο πίνακα παραθέτουμε τα x_i και y_i για το χαρτοφυλάκιο 7.

| Βαθμίδες | Οφειλέτες | Αθετήσεις | x_i | y_i |
|----------|-----------|-----------|---------|-------|
| A | 100 | 3 | 100/700 | 3/5 |
| B | 175 | 2 | 275/700 | 1 |
| Γ | 225 | 0 | 500/700 | 1 |
| Δ | 150 | 0 | 650/700 | 1 |
| E | 50 | 0 | 1 | 1 |
| Σύνολο | 700 | 5 | - | - |

Πίνακας 4.1 Αθροιστικά ποσοστά οφειλετών και αθετήσεων για το χαρτοφυλάκιο 7

Χρησιμοποιώντας την σχέση (4.22) παίρνουμε τις τιμές $AUC = 0.85$ και $AR = 0.7$. Επίσης, εφόσον $AUC > 0.8$ μπορούμε να χρησιμοποιήσουμε την σχέση (4.13) για

την εκτίμηση του k . Σε αυτή την περίπτωση θα λαμβάναμε $k = 6.67$. Για αυτήν την τιμή της κοιλότητας το $RMSE$ είναι 0.033. Όπως βλέπουμε οι διαφορές στις τιμές των $AR, AUC, RMSE$ είναι πολύ μικρές ενώ και οι εκτιμώμενες τιμές της κοιλότητας δεν έχουν μεγάλη διαφορά. Όπως ήταν αναμενόμενο από το επόμενο διάγραμμα διαπιστώνουμε ότι οι καμπύλες CAP που δίνονται από την σχέση (4.5) για $k = 6.96$ και $k = 6.67$ διαφέρουν ελάχιστα.

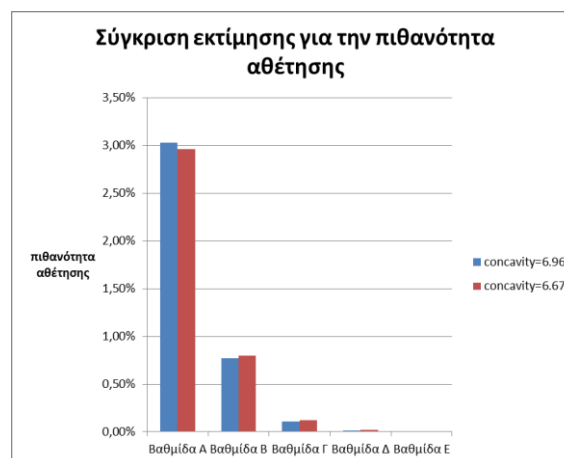


Σχήμα 4.7 Καμπύλες CAP για $k = 6.96$ και $k = 6.67$ στο χαρτοφυλάκιο 7

Στον ακόλουθο πίνακα θα παρουσιάσουμε τις εκτιμώμενες πιθανότητες αθέτησης για κάθε βαθμίδα του χαρτοφυλακίου 7 και για τις δύο τιμές του k . Στην 2^η στήλη έχουμε τις εκτιμήσεις για $k = 6.96$ και στην 3^η για $k = 6.67$.

| | | |
|------------|--------|--------|
| p_A | 3.03% | 2.96% |
| p_B | 0.77% | 0.80% |
| p_Γ | 0.11% | 0.12% |
| p_Δ | 0.016% | 0.02% |
| p_E | 0.006% | 0.008% |

Πίνακας 4.2 Εκτίμηση των πιθανοτήτων αθέτησης για $k = 6.96$ και $k = 6.67$



Σχήμα 4.8 Εκτίμηση των $p_A, p_B, p_\Gamma, p_\Delta, p_E$ για 2 διαφορετικές τιμές του concavity

Όπως διαπιστώνουμε από τον Πίνακα 5.1 και το Σχήμα 5.7 οι εκτιμήσεις που λαμβάνουμε για τις 2 επιλεγμένες τιμές του k είναι πολύ κοντά. Παρατηρούμε επίσης πως για $k = 6.67$ οι εκτιμώμενες πιθανότητες για τις βαθμίδες B έως και E είναι μεγαλύτερες από τις αντίστοιχες εκτιμήσεις για $k = 6.96$ ενώ για την βαθμίδα A λάβαμε μεγαλύτερη εκτίμηση για $k = 6.96$. Τέλος, να σημειώσουμε πως και για τις δύο τιμές της παραμέτρου κοιλότητας οι εκτιμήσεις για τις βαθμίδες A και E είναι τόσο μικρές που σχεδόν δεν φαίνονται στο Σχήμα 5.7.

Προκειμένου να αναλύσουμε περαιτέρω το μοντέλο θα παραθέσουμε αριθμητικά αποτελέσματα για το χαρτοφυλάκιο 7 αλλάζοντας τον αριθμό των αθετήσεων. Συγκεκριμένα, θα θεωρήσουμε πως έχουν συμβεί από 2 αθετήσεις στις βαθμίδες A και B , 1 στην βαθμίδα Γ και καμία στις βαθμίδες Δ και E . Σε αυτή την περίπτωση το CT έχει παραμείνει $\frac{5}{700}$ αλλά η τιμή του k προκειμένου να ελαχιστοποιείται το $RMSE$ είναι πλέον 3.77. Επίσης, από τις σχέσεις (4.12) και (4.14) τα AUC και AR είναι 0.71 και 0.42 αντίστοιχα. Αν χρησιμοποιούσαμε πάλι τον κανόνα του τραπεζίου για να εκτιμήσουμε τα AUC και AR θα παίρναμε τις τιμές 0.75 και 0.51. Όπως είδαμε στην προηγούμενη Ενότητα για αυτές τις τιμές δεν ενδείκνυται να χρησιμοποιήσουμε τους τύπους (4.13) και (4.16) προκειμένου να λάβουμε μια δεύτερη εκτίμηση για το k . Στον επόμενο πίνακα παρουσιάζονται οι εκτιμήσεις της πιθανότητας αθέτησης για κάθε βαθμίδα του χαρτοφυλακίου 7 για $k = 3.77$.

| | |
|------------|-------|
| p_A | 2.11% |
| p_B | 1.00% |
| p_Γ | 0.34% |
| p_Δ | 0.12% |
| p_E | 0.07% |

Πίνακας 4.3 Εκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 7 με $k = 3.77$

Συγκρίνοντας τα αποτελέσματα του Πίνακα 5.2 με τα αποτελέσματα του Πίνακα 5.1 παρατηρούμε πως παρότι ο αριθμός των αθετήσεων στις βαθμίδες A , E δεν άλλαξε οι εκτιμήσεις που λάβαμε στην 2^η περίπτωση είναι δραματικά μεγαλύτερες. Το ίδιο συνέβη και στην βαθμίδα B σε μικρότερο ωστόσο βαθμό. Για τις βαθμίδες A και Γ η εκτίμηση μικρότερης και μεγαλύτερης, αντίστοιχα, πιθανότητας αθέτησης ήταν αναμενόμενη.

Από τα αποτελέσματα των Πινάκων 5.1 και 5.2 παρατηρούμε πως οι εκτιμήσεις των πιθανοτήτων αθέτησης παρουσιάζουν μεγάλες διαφορές από βαθμίδα σε βαθμίδα. Χαρακτηριστικό παράδειγμα αποτελούν οι εκτιμήσεις για τις βαθμίδες Γ και E του Πίνακα 5.1. Παρατηρούμε πως παρότι και στις 2 βαθμίδες δεν έχουν παρατηρηθεί αθετήσεις και στην βαθμίδα Γ υπάρχουν σχεδόν 4 φορές παραπάνω οφειλέτες, η εκτίμηση της p_E είναι περίπου 18 φορές μικρότερη από την εκτίμηση για την p_Γ . Παρακάτω θα δείξουμε τον λόγο που οδηγεί σε μια τόσο μεγάλη διαφορά.

Γνωρίζουμε από την σχέση (4.20) ότι η εκτίμηση της πιθανότητας αθέτησης για μια βαθμίδα αξιολόγησης R δίνεται από την παρακάτω σχέση,

$$PD_R = CT * \frac{dy}{dx} = CT \frac{ke^{-kx_R}}{1-e^{-k}}, R = 1, \dots, N.$$

Θα υπολογίσουμε τον λόγο $\frac{PD_R}{PD_{R+1}}$ θεωρώντας ότι η βαθμίδα R είναι χαμηλότερης πιστοληπτικής ικανότητας από την βαθμίδα $R + 1$. Από την (4.20) έχουμε,

$$\frac{PD_R}{PD_{R+1}} = \frac{e^{-kx_R}}{e^{-kx_{R+1}}} = e^{-kx_R+kx_{R+1}} = e^{k(x_{R+1}-x_R)}. \quad (4.24)$$

Χρησιμοποιώντας την σχέση (4.21) προκύπτει ότι,

$$x_R - x_{R+1} = \frac{n_{R+1}+n_R}{2n} \quad (4.25)$$

και αντικαθιστώντας την (4.25) στην (4.24) παίρνουμε,

$$\frac{PD_R}{PD_{R+1}} = e^{\frac{k(n_{R+1}+n_R)}{2n}}. \quad (4.26)$$

Ο παραπάνω λόγος μπορεί να υπολογιστεί και προσεγγιστικά αν έχουμε ένα μοντέλο με μεγάλη διακριτική ισχύ, όπως στην περίπτωση που αντιστοιχεί στον Πίνακα 5.1, με $AUC \geq 0.8$. Σε αυτή την περίπτωση χρησιμοποιώντας τις σχέσεις (4.13) και (4.16) παίρνουμε τους παρακάτω προσεγγιστικούς τύπους, αντίστοιχα.

$$\frac{PD_R}{PD_{R+1}} \approx e^{\frac{(n_{R+1}+n_R)}{2n(1-AUC)}} \quad (4.27)$$

$$\frac{PD_R}{PD_{R+1}} \approx e^{\frac{(n_{R+1}+n_R)}{n(1-AR)}} \quad (4.28)$$

Από τις σχέσεις (4.26), (4.27) και (4.28) προκύπτει πως η εκτιμώμενη τιμή της πιθανότητας αθέτησης για μια βαθμίδα είναι εκθετικά μικρότερη από την εκτίμηση της αμέσως προηγούμενης χειρότερης βαθμίδας. Επίσης, όπως έχουμε ήδη επισημάνει, όσο καλύτερο είναι το μοντέλο αξιολόγησης τόσο πιο υψηλές τιμές παίρνουν τα AUC και AR . Συνδυάζοντας αυτήν την πληροφορία με τις τελευταίες σχέσεις είναι εμφανές πως όσο μεγαλύτερη διακριτική ισχύ έχει το μοντέλο τόσο μικρότερες εκτιμήσεις θα έχουμε για τις βαθμίδες χαμηλού ρίσκου.

4.5 Η χρήση της γλώσσας R στην εύρεση της παραμέτρου κοιλότητας

Το κρίσιμο σημείο για την εκτίμηση της πιθανότητας αθέτησης με την μεθοδολογία που αναφέραμε σε αυτό το Κεφάλαιο είναι η εκτίμηση της παραμέτρου κοιλότητας k . Όπως ήδη αναφέραμε ο ενδεδειγμένος τρόπος σύμφωνα με τον Burgt(2008) είναι η εύρεση εκείνης της τιμής του k που ελαχιστοποιεί το $RMSE$. Για να εκτιμήσουμε την συγκεκριμένη παράμετρο με τον προτεινόμενο τρόπο χρησιμοποιήσαμε την γλώσσα προγραμματισμού R και το πακέτο *mltools*(Gorman,

2018). Συγκεκριμένα από το εν λόγω πακέτο χρησιμοποιήθηκε η εντολή $rmse(preds,actuals)$. Το όρισμα $preds$ αφορά τις εκτιμήσεις για το σωρευτικό ποσοστό αθετήσεων που λαμβάνουμε από την σχέση (4.5) για μια δεδομένη τιμή του k ενώ το όρισμα $actuals$ είναι το πραγματικό σωρευτικό ποσοστό αθετήσεων που έχουν συμβεί στην εκάστοτε βαθμίδα που λαμβάνουμε κατευθείαν από τα δεδομένα μας. Με την εντολή $rmse$ υπολογίζουμε την τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος. Για να μπορέσουμε λοιπόν να βρούμε την τιμή του k που ελαχιστοποιεί το $RMSE$ για το χαρτοφυλάκιο 7, ορίσαμε ένα διάνυσμα για τις πιθανές τιμές του k από το 0 έως και το 40, με βήμα 0.01 και εκτελέσαμε μία επαναληπτική εντολή `for` ελέγχοντας για κάθε τιμή του k την τιμή που λαμβάνει το $RMSE$ και κρατώντας την τιμή του k που το ελαχιστοποιεί.

4.6 Σχολιασμός της μεθόδου

Η μέθοδος αυτού του Κεφαλαίου βασίζεται στην προσαρμογή μίας κοίλης συνάρτησης της μορφής (4.5) στην καμπύλη CAP. Η κύρια παράμετρος για την εκτίμηση της πιθανότητας αθέτησης είναι η κοιλότητα. Όσο υψηλότερη, θετική η τιμή του k τόσο υψηλότερη διακριτική ισχύ έχει το μοντέλο βαθμονόμησης. Στην Ενότητα 4.4 είδαμε ότι το μοντέλο παρουσιάζει μεγάλες διαφορές στις εκτιμήσεις των πιθανοτήτων αθέτησης ανάλογα με την βαθμίδα στην οποία έχουν συμβεί οι αθετήσεις. Επίσης δείξαμε πως υπάρχει εκθετική μείωση της εκτιμώμενης πιθανότητας αθέτησης προχωρώντας από τις βαθμίδες υψηλού κινδύνου στις βαθμίδες χαμηλού κινδύνου. Μετά τον καθορισμό του k η εκτίμηση των πιθανοτήτων αθέτησης γίνονται από την σχέση (4.20) στην οποία το CT είναι η παρατηρούμενη πιθανότητα αθέτησης όλου του χαρτοφυλάκιου. Με αυτό τον τρόπο μπορεί να εκτιμηθεί η πιθανότητα αθέτησης ακόμα και για μια βαθμίδα στην οποία δεν έχει συμβεί καμία αθέτηση. Να τονίσουμε πως με τη συγκεκριμένη μεθοδολογία δεν μπορεί να εκτιμηθεί η πιθανότητα αθέτησης αν δεν έχουν συμβεί αθετήσεις στο χαρτοφυλάκιο. Προφανώς αυτό είναι ένα μειονέκτημα της προτεινόμενης μεθόδου καθώς στα χαρτοφυλάκια υψηλής φερεγγυότητας είναι αρκετά πιθανό να συναντήσουμε κάτι τέτοιο.

5. Προσομοίωση και αξιολόγηση των μεθόδων

Σε αυτό το Κεφάλαιο θα προβούμε σε προσομοίωση, προκειμένου να διαπιστώσουμε και να αξιολογήσουμε τις μεθοδολογίες που αναφέρθηκαν στα Κεφάλαια 2, 3, 4.

5.1 Προσομοίωση και αξιολόγηση της συντηρητικής αρχής εκτίμησης

Προκειμένου να αξιολογήσουμε την αποτελεσματικότητα της μεθόδου που πρότειναν οι Pluto and Tasche(2005),σε αυτήν την ενότητα θα προκαθορίσουμε τα p_A, p_B, p_G και θα προσομοιώσουμε δεδομένα. Στην συνέχεια θα εκτιμήσουμε τις πιθανότητες αθέτησης σύμφωνα με τις μεθόδους που αναλύθηκαν στις Ενότητες 2.1 έως και 2.5. Συγκεκριμένα για την μεθοδολογία της Ενότητας 2.5 θα προβούμε σε μία περαιτέρω ανάλυση προκειμένου να δείξουμε πόσο επηρεάζουν τις εκτιμώμενες πιθανότητες οι παράμετροι τ, ρ .

5.1.1 Αξιολόγηση της μεθόδου υπό την υπόθεση της ανεξαρτησίας

Πριν προβούμε στην αξιολόγηση της μεθοδολογίας που αναλύθηκε στην Ενότητα 2.2 θα δείξουμε πως μπορούμε να υπολογίσουμε αριθμητικά τις πιθανότητες αθέτησης. Παρακάτω θα αναλύσουμε έναν τρόπο υπολογισμού του p_G , που δίνεται από την σχέση (2.11). Η σχέση (2.11) μπορεί να γραφεί:

$$1 - \sum_{i=0}^c \binom{n_G}{i} p_G^i (1 - p_G)^{n_G - i} \leq \gamma. \quad (5.1)$$

Επίσης γνωρίζουμε ότι αν μια τυχαία μεταβλητή Y ακολουθεί κατανομή Beta με ακέραιους παραμέτρους α, β για την αθροιστική συνάρτηση κατανομής της ισχύει:

$$F(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^y x^{\alpha-1} (1-x)^{\beta-1} dx = \sum_{i=\alpha}^{\alpha+\beta-1} \binom{\alpha+\beta-1}{i} y^i (1-y)^{\alpha+\beta-1-i}.$$

Ισοδύναμα,

$$F(y) = 1 - \sum_{i=0}^{\alpha-1} \binom{\alpha+\beta-1}{i} y^i (1-y)^{\alpha+\beta-1-i}.$$

Από την παραπάνω σχέση προκύπτει ότι η σχέση (5.1) μπορεί να γραφεί

$$F(p_G) \leq \gamma \quad (5.2)$$

όπου η F είναι η αθροιστική συνάρτηση κατανομής της Beta με παραμέτρους $c + 1$ και $n_G - c$. Οπότε η p_G μπορεί να υπολογιστεί ως ποσοστιαίο σημείο της κατανομής Beta. Ακολουθώντας παρόμοια διαδικασία μπορούμε να καταλήξουμε στο συμπέρασμα ότι η p_A μπορεί να υπολογιστεί ως ποσοστιαίο σημείο της κατανομής Beta με παραμέτρους $d + 1$ και $n_A + n_B + n_G - d$ ενώ η p_B μπορεί να υπολογιστεί ως ποσοστιαίο σημείο της κατανομής Beta με παραμέτρους $b + c + 1$ και $n_B + n_G - b - c$.

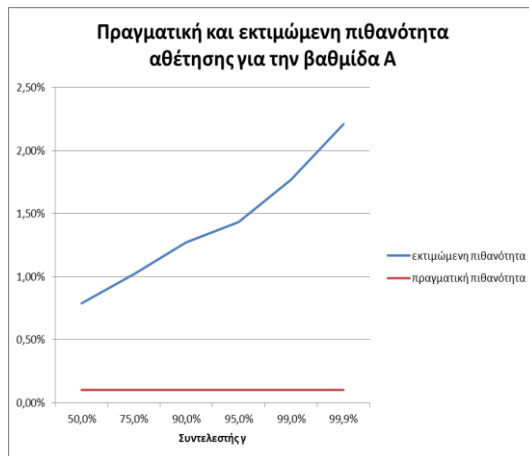
Για την προσομοίωση χρησιμοποιήσαμε το εξής χαρτοφυλάκιο:

| Χαρτοφυλάκιο | Οφειλέτες Α βαθμίδας | Οφειλέτες Β βαθμίδας | Οφειλέτες Γ βαθμίδας | p_A | p_B | p_Γ |
|--------------|-------------------------|-------------------------|-------------------------|-------|-------|------------|
| 8 | 80 | 300 | 370 | 0.1% | 0.5% | 1% |

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.79% | 1.02% | 1.27% | 1.43% | 1.77% | 2.21% |
| p_B | 0.87% | 1.13% | 1.40% | 1.58% | 1.96% | 2.45% |
| p_Γ | 1.16% | 1.56% | 2.00% | 2.29% | 2.92% | 3.73% |

Πίνακας 5.1 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 8

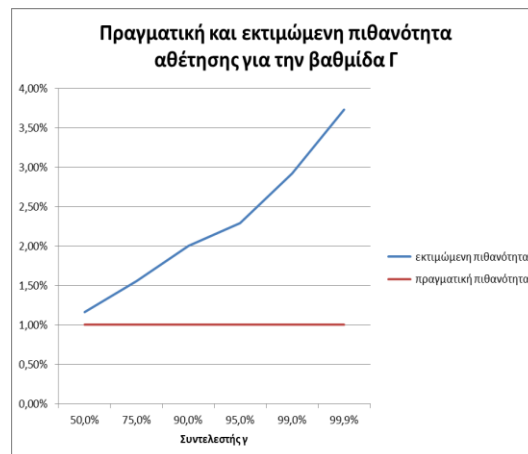
Από τα αποτελέσματα του πίνακα που προέκυψε από την προσομοίωση διαπιστώνουμε πόσο συντηρητική είναι η μέθοδος εκτίμησης που πρότειναν οι Pluto and Tasche(2005). Διαπιστώνουμε ότι οι εκτιμήσεις που πήραμε για όλες τις βαθμίδες, για όλες τις τιμές του συντελεστή γ είναι μεγαλύτερες από τις πραγματικές πιθανότητες αθέτησης. Πιο συγκεκριμένα, παρατηρούμε ότι για την βαθμίδα Α για $\gamma = 99.9\%$ η πιθανότητα που εκτιμήθηκε είναι πάνω από είκοσι φορές μεγαλύτερη από την πραγματική. Η υπερεκτίμηση των πιθανοτήτων αθέτησης είναι εμφανής και στα παρακάτω γραφήματα



Σχήμα 5.1 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για την βαθμίδα Α



Σχήμα 5.2 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για την βαθμίδα Β



Σχήμα 5.3 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για την βαθμίδα Γ

5.1.2 Αξιολόγηση της μεθόδου υπό την υπόθεση της εξάρτησης του δείγματος για μονοετή περίοδο παρατήρησης

Όπως έχουμε ήδη αναφέρει, είναι λογικό να υποθέσουμε ότι υπάρχει εξάρτηση μεταξύ των παρατηρήσεων του δείγματος. Στην παρούσα Ενότητα θα προβούμε σε προσομοίωση προκειμένου να αξιολογήσουμε την μέθοδο που αναλύθηκε στις Ενότητες 2.3 και 2.4.

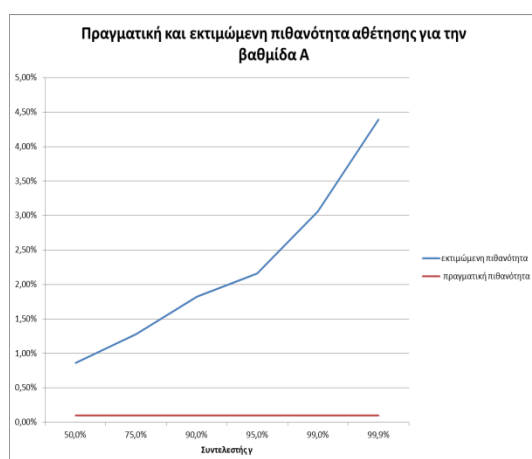
Για την προσομοίωση θα χρησιμοποιήσουμε το παρακάτω χαρτοφυλάκιο:

| Χαρτοφυλάκιο | Οφειλέτες <i>A</i> βαθμίδας | Οφειλέτες <i>B</i> βαθμίδας | Οφειλέτες <i>Γ</i> βαθμίδας | p_A | p_B | p_Γ | ρ |
|--------------|-----------------------------------|-----------------------------------|-----------------------------------|-------|-------|------------|--------|
| 8 | 80 | 300 | 370 | 0.1% | 0.5% | 1% | 3% |

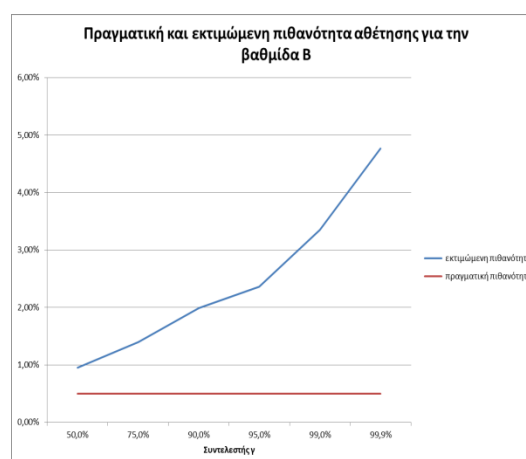
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.86% | 1.28% | 1.82% | 2.16% | 3.06% | 4.39% |
| p_B | 0.95% | 1.40% | 1.99% | 2.36% | 3.35% | 4.77% |
| p_Γ | 1.26% | 1.91% | 2.73% | 3.22% | 4.61% | 6.58% |

Πίνακας 5.2 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 8 για $\rho = 3\%$

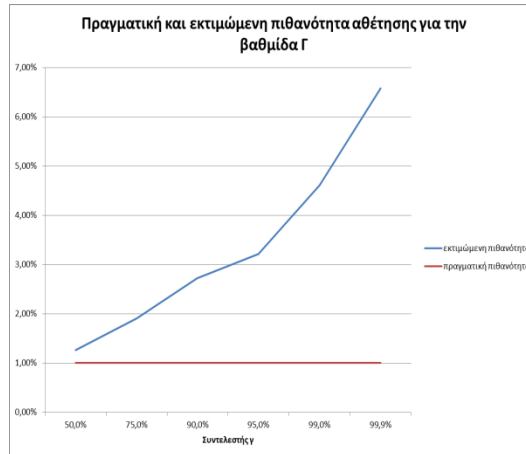
Όπως ήταν αναμενόμενο η εισαγωγή του ρ οδήγησε σε επιπλέον υπερεκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 8. Ενδεικτικά, για $\gamma = 99.9\%$ η εκτίμηση που λάβαμε για την p_A είναι σχεδόν διπλάσια από την αντίστοιχη εκτίμηση του Πίνακα 5.1. Μπορούμε λοιπόν να αποφανθούμε πως η εισαγωγή ενός συντελεστή συσχέτισης όταν η περίοδος παρατήρησης των παρατηρήσεων του δείγματος είναι 1 έτος, οδηγεί σε μία άκρως αναποτελεσματική μέθοδο εκτίμησης.



Σχήμα 5.4 Σύγκριση της πραγματικής και της εκτιμώμενης p_A για το χαρτοφυλάκιο 8



Σχήμα 5.5 Σύγκριση της πραγματικής και της εκτιμώμενης p_B για το χαρτοφυλάκιο 8



Σχήμα 5.6 Σύγκριση της πραγματικής και της εκτιμώμενης p_{Γ} για το χαρτοφυλάκιο 8

5.1.3 Αξιολόγηση της μεθόδου υπό την υπόθεση της εξάρτησης του δείγματος για πολυετή περίοδο παρατήρησης

Όπως είδαμε, η μέθοδος που αναλύθηκε στην Ενότητα 2.3 και 2.4 οδηγεί σε υπερεκτίμηση των πιθανοτήτων αθέτησης ακόμα και για πολύ μικρές τιμές των συντελεστών ρ και γ . Στην παρούσα Ενότητα θα ασχοληθούμε με την μεθοδολογία της Ενότητας 2.5 και θα παραθέσουμε αριθμητικά αποτελέσματα για το παρακάτω χαρτοφυλάκιο:

| Χαρτοφυλάκιο | Οφειλέτες Α βαθμίδας | Οφειλέτες Β βαθμίδας | Οφειλέτες Γ βαθμίδας | p_A | p_B | p_{Γ} |
|--------------|----------------------------|----------------------------|----------------------------|-------|-------|--------------|
| 9 | 300 | 200 | 250 | 0.6% | 1.5% | 3% |

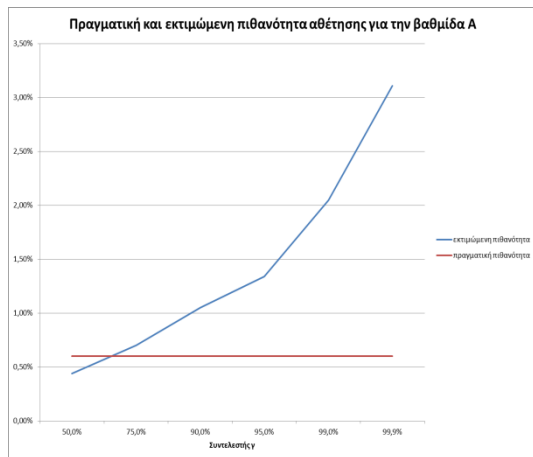
Για το χαρτοφυλάκιο 8 έχουμε επίσης,

$$T = 5, \rho = 15\%, \tau = 30\%, \gamma = 50\%, 75\%, 90\%, 95\%, 99\%, 99.9\%$$

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|--------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.44% | 0.70% | 1.05% | 1.34% | 2.05% | 3.11% |
| p_B | 0.62% | 0.97% | 1.45% | 1.82% | 2.75% | 4.11% |
| p_{Γ} | 0.79% | 1.26% | 1.88% | 2.34% | 3.53% | 5.22% |

Πίνακας 5.3 Εκτίμηση των p_A, p_B, p_{Γ} για το χαρτοφυλάκιο 9 με $T = 5, \rho = 15\%, \tau = 30\%$

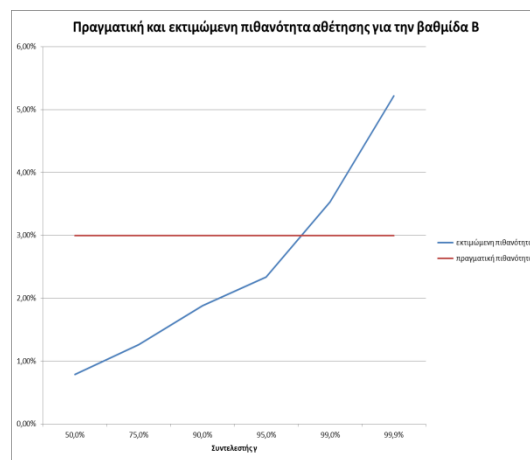
Παρατηρούμε ότι για μικρές τιμές συντελεστή εμπιστοσύνης υπάρχει υποεκτίμηση των πιθανοτήτων ενώ αντιθέτως η επιλογή $\gamma = 99.9\%$ μας οδηγεί σε υπερεκτίμηση του πραγματικού κινδύνου. Είναι εμφανές ωστόσο, πως με τη συγκεκριμένη μεθοδολογία και επιλέγοντας υψηλή τιμή για το συντελεστή εμπιστοσύνης μπορούμε να πάρουμε “καλύτερα” αποτελέσματα από ότι στην περίπτωση που αναλύθηκε στις Ενότητες 2.1 έως και 2.4. Όσα σχολιάστηκαν παραπάνω φαίνονται και στα παρακάτω διαγράμματα.



Σχήμα 5.7 Σύγκριση της πραγματικής και της εκτιμώμενης p_A για το χαρτοφυλάκιο 9



Σχήμα 5.8 Σύγκριση της πραγματικής και της εκτιμώμενης p_B για το χαρτοφυλάκιο 9



Σχήμα 5.9 Σύγκριση της πραγματικής και της εκτιμώμενης p_Γ για το χαρτοφυλάκιο 9

Επιπλέον θα προχωρήσουμε σε μια συγκριτική ανάλυση των αποτελεσμάτων προκειμένου να εξετάσουμε πόσο σημαντικοί παράγοντες για τις εκτιμώμενες πιθανότητες είναι οι τιμές των τ, ρ . Αρχικά θα εκτιμήσουμε τα p_A, p_B, p_Γ για το χαρτοφυλάκιο 9 κρατώντας ίδιες τιμές για τα T, τ αλλάζοντας ωστόσο την τιμή του ρ από 15% σε 12%. Οι εκτιμήσεις που λαμβάνουμε λόγω αυτής της αλλαγής δίνονται στον Πίνακα 5.4.

| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
|------------|-------|-------|-------|-------|-------|-------|
| p_A | 0.41% | 0.64% | 0.93% | 1.16% | 1.70% | 2.53% |
| p_B | 0.59% | 0.90% | 1.29% | 1.60% | 2.31% | 3.40% |
| p_Γ | 0.76% | 1.18% | 1.67% | 2.07% | 2.99% | 4.40% |

Πίνακας 5.4 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 9 με $T=5, \rho=12\%, \tau=30\%$

Όπως ήταν αναμενόμενο οι πιθανότητες που εκτιμήθηκαν για $\rho = 12\%$ είναι λίγο μικρότερες από ότι στην προηγούμενη περίπτωση. Παρατηρούμε όμως πως και για τις 2 επιλεγμένες τιμές του ρ οι εκτιμήσεις παρουσιάζουν παρόμοια εικόνα. Η πραγματική πιθανότητα αθέτησης για την βαθμίδα A αντιστοιχεί σε μία εκτίμηση για γ κοντά στο 0.75, η πιθανότητα αθέτησης για την βαθμίδα B αντιστοιχεί σε τιμή του

συντελεστή εμπιστοσύνης μεταξύ 0.90 και 0.95 ενώ για την βαθμίδα Γ η πιθανότητα αθέτησης υπερεκτιμάται μόνο στην περίπτωση που επιλεγεί το $\gamma = 0.999$.

Τέλος θα εξετάσουμε πόσο σημαντικός παράγοντας για την μέθοδο είναι η συσχέτιση τ . Θα χρησιμοποιήσουμε πάλι το χαρτοφυλάκιο 9 επιλέγοντας τις ακόλουθες τιμές, $T = 5, \rho = 12\%, \tau = 20\%$.

| | | | | | | |
|------------|-------|-------|-------|-------|-------|-------|
| γ | 50% | 75% | 90% | 95% | 99% | 99.9% |
| p_A | 0.40% | 0.62% | 0.88% | 1.08% | 1.54% | 2.26% |
| p_B | 0.57% | 0.87% | 1.22% | 1.49% | 2.13% | 3.05% |
| p_Γ | 0.75% | 1.13% | 1.60% | 1.93% | 2.76% | 3.96% |

Πίνακας 5.5 Εκτίμηση των p_A, p_B, p_Γ για το χαρτοφυλάκιο 9 με $T = 5, \rho = 12\%, \tau = 20\%$

Συγκρίνοντας τα αποτελέσματα αυτού του πίνακα με τα αποτελέσματα του Πίνακα 5.4 βλέπουμε ότι η επιλογή μικρότερης τιμής για το τ οδήγησε σε εκτίμηση μικρότερων τιμών για τις πιθανότητες αθέτησης ωστόσο και σε αυτήν την περίπτωση ισχύει η παρατήρηση που κάναμε προηγουμένως

5.2 Προσομοίωση και αξιολόγηση της προσέγγισης Bayes

Στην παρούσα ενότητα θα προσομοιάσουμε δεδομένα προκειμένου να αξιολογήσουμε την αποτελεσματικότητα της προσέγγισης Bayes, για τις εκ των προτέρων κατανομές που πρότειναν οι Kiefer(2009) και Tasche(2013).

5.2.1 Αξιολόγηση της μεθόδου λαμβάνοντας ως εκ των προτέρων κατανομή την Beta

Θεωρούμε το παρακάτω χαρτοφυλάκιο:

| Χαρτοφυλάκιο | Οφειλέτες χαρτοφυλακίου | Πιθανότητα αθέτησης |
|--------------|-------------------------|---------------------|
| 10 | 1200 | 1% |

Για το χαρτοφυλάκιο 10 θα μελετήσουμε την μέθοδο με τις παρακάτω εκ των προτέρων κατανομές:

- α) Beta με παραμέτρους $\alpha = 2$ και $\beta = 500$
- β) Beta με παραμέτρους $\alpha = 4$ και $\beta = 50$

Συμβολίζοντας με θ_1 την εκτίμηση με χρήση της Beta(2,500), θ_2 την εκτίμηση με χρήση της Beta(4,50) και MLE τον εκτιμητή μέγιστης πιθανοφάνειας βάσει της προσομοίωσης που πραγματοποιήθηκε λαμβάνουμε τα παρακάτω αποτελέσματα

| N | θ_1 | θ_2 | MLE |
|------|------------|------------|-----|
| 1200 | 0.82% | 1.28% | 1% |

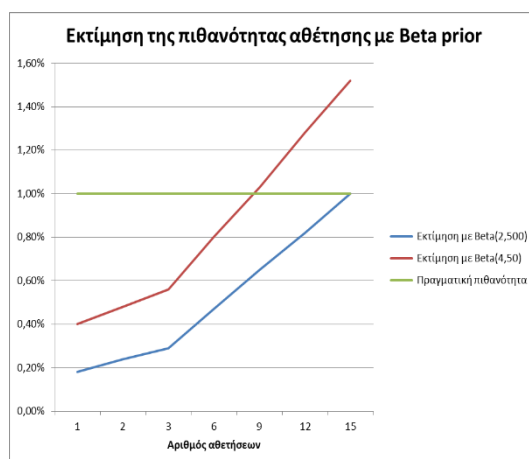
Πίνακας 5.6 Εκτίμηση της πιθανότητας αθέτησης με χρήση Beta εκ των προτέρων κατανομής για το χαρτοφυλάκιο 10

Όπως γίνεται εύκολα αντιληπτό, από την προσομοίωση που κάναμε παρότι και οι 2 εκτιμήσεις είναι σχετικά κοντά στην πραγματική πιθανότητα αθέτησης, αμφότερες

επηρεάζονται σε μεγάλο βαθμό από την επιλογή της εκ των προτέρων κατανομής. Για να γίνει αυτό κατανοητό αρκεί να παρατηρήσουμε το εξής:

Στην περίπτωση της Beta(4,50) βλέπουμε ότι έχουμε λάβει μία εκτίμηση ελαφρώς μεγαλύτερη από την πραγματική πιθανότητα. Αυτό οφείλεται στο γεγονός ότι λάβαμε μια “συντηρητική” εκ των προτέρων κατανομή. Η κατανομή Beta(4,50) έχει μέση τιμή 7.4%, πρακτικά αυτό σημαίνει ότι εκ των προτέρων είχαμε υπερεκτιμήσει τον κίνδυνο σε μεγάλο βαθμό και λόγω αυτής της υπερεκτίμησης λάβαμε μια υψηλότερη εκτίμηση για την πιθανότητα αθέτησης. Αντιθέτως, με την κατανομή Beta(2,500) είχαμε υποεκτιμήσει τον κίνδυνο, καθώς η μέση τιμή σε αυτήν την περίπτωση είναι σχεδόν 0.4%, αυτό είχε ως αποτέλεσμα να οδηγηθούμε σε μία μικρή υποεκτίμηση της πιθανότητας αθέτησης. Στα παρακάτω γραφήματα παρουσιάζεται η εκτίμηση της πιθανότητας αθέτησης για το χαρτοφυλάκιο 10 για τις προαναφερθείσες εκ των προτέρων κατανομές και για τις τιμές

$$r = 1,2,3,6,9,12,15$$



Σχήμα 5.10 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 10 με χρήση Beta prior

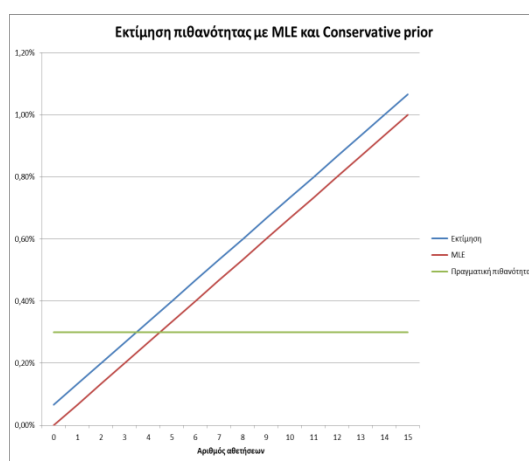
Είναι εμφανές και από το διάγραμμα πως ο αριθμός των αθετήσεων επηρεάζει αρκετά την εκτίμηση των πιθανοτήτων. Επίσης με την Beta(4,50) η πιθανότητα αθέτησης υπερεκτιμάται αν συμβούν πάνω από 9 αθετήσεις ενώ το πρόβλημα της υποεκτίμησης λόγω της Beta(2,500) επιλύεται αν στο χαρτοφυλάκιο συμβούν οι 12 αναμενόμενες αθετήσεις.

5.2.2 Αξιολόγηση της μεθόδου με χρήση της συντηρητικής εκ των προτέρων κατανομής

Σε αυτήν την ενότητα θα εξετάσουμε την αποτελεσματικότητα της μεθόδου που αναλύθηκε στην Ενότητα 3.4.2 για το παρακάτω χαρτοφυλάκιο:

| Χαρτοφυλάκιο | Οφειλέτες χαρτοφυλακίου | Πιθανότητα αθέτησης |
|--------------|-------------------------|---------------------|
| 11 | 1500 | 0.3% |

Κάνοντας προσομοίωση για το συγκεκριμένο χαρτοφυλάκιο προέκυψε ότι η εκτίμηση για την πιθανότητα αθέτησης είναι 0.36% μία τιμή πολύ κοντά στην πραγματική πιθανότητα. Όπως γίνεται εμφανές στο παρακάτω διάγραμμα η εκτίμηση που πρότεινε ο Tasche(2013) παρουσιάζει μεγάλη ευαισθησία στον αριθμό των αθετήσεων. Σε σύγκριση με τον εκτιμητή μέγιστης πιθανοφάνειας υπάρχει το πλεονέκτημα πως μπορεί να υπάρξει εκτίμηση για την πιθανότητα αθέτησης ακόμα και με μηδενικό αριθμό αθετήσεων. Γενικά ο συγκεκριμένος εκτιμητής είναι πολύ κοντά με τον MLE καθώς και οι 2 κάνουν μια “καλή” εκτίμηση μόνο όταν ο αριθμός των αθετήσεων που λαμβάνουμε από την προσομοίωση είναι κοντά στον αναμενόμενο αριθμό αθετήσεων βάσει της πραγματικής πιθανότητας αθέτησης.



Σχήμα 5.11 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 11 με χρήση conservative prior και MLE

5.2.3 Αξιολόγηση της μεθόδου με χρήση της ομοιόμορφης εκ των προτέρων κατανομής

Θα εκτιμήσουμε την πιθανότητα αθέτησης με χρήση της ομοιόμορφης κατανομής για το παρακάτω χαρτοφυλάκιο

| Χαρτοφυλάκιο | Οφειλέτες χαρτοφυλακίου | Πιθανότητα αθέτησης |
|--------------|-------------------------|---------------------|
| 12 | 1000 | 1% |

Στον παρακάτω πίνακα παρουσιάζονται οι εκτιμήσεις της πιθανότητας αθέτησης:

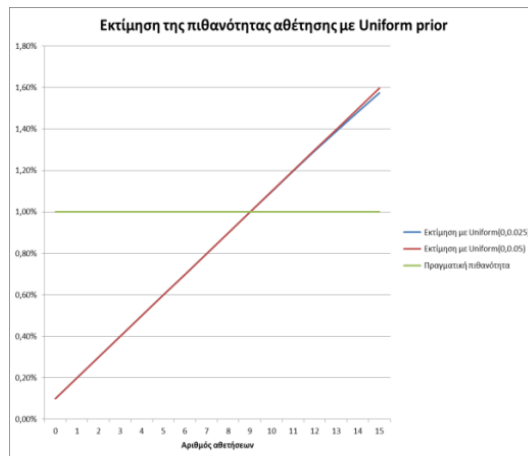
- α) με χρήση της ομοιόμορφης εκ των προτέρων κατανομής για τις τιμές:
 $u = 0.025, 0.05, 0.1, 1$
- β) με χρήση της συντηρητικής εκ των προτέρων κατανομής
- γ) με την μέθοδο των Pluto and Tasche(2005), υπό την υπόθεση της ανεξαρτησίας, για $\gamma = 50\%$ και 75%

| | |
|-----------|--------|
| Εκτιμητής | n=1000 |
|-----------|--------|

| | |
|-----------------|---------|
| (0,0.025) | 1.0885% |
| (0,0.05) | 1.0994% |
| (0,0.1) | 1.1015% |
| (0,1) | 1.1058% |
| $\hat{\theta}$ | 1.0985% |
| $\theta_{0,75}$ | 1.2976% |
| $\theta_{0,5}$ | 1.0667% |

Πίνακας 5.7 Εκτίμηση της πιθανότητας αθέτησης με την ομοιόμορφη κατανομή, την συντηρητική κατανομή και τον συντηρητικό εκτιμητή για $\gamma=50\%$ και 75%

Όπως είχαμε διαπιστώσει και στην Ενότητα 3.4.3 η εκτίμηση με χρήση της ομοιόμορφης ως εκ των προτέρων κατανομή δίνει σχεδόν ίδια αποτελέσματα με την συντηρητική εκ των προτέρων κατανομή ενώ σε σύγκριση με τον συντηρητικό εκτιμητή των Pluto and Tasche(2005) η εκτίμηση είναι κοντά σε αυτές που λαμβάνουμε αν επιλέξουμε συντελεστή εμπιστοσύνης 50% ή 75% . Στο παρακάτω διάγραμμα παρουσιάζεται η εκτίμηση της πιθανότητας αθέτησης για το χαρτοφυλάκιο 11 για διάφορες τιμές του r και για $u = 0.025$ ή 0.05 . Όπως γίνεται άμεσα αντιληπτό ο αριθμός των παρατηρούμενων αθετήσεων είναι ο κρίσιμος παράγοντας για την εκτίμηση της πιθανότητας αθέτησης.



Σχήμα 5.12 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 12 με χρήση Ομοιόμορφης εκ των προτέρων κατανομής

5.3 Προσομοίωση και αξιολόγηση της καμπύλης CAP

Σε αυτή την Ενότητα θα προσομοιάσουμε δεδομένα εκτιμώντας την πιθανότητα αθέτησης για ένα χαρτοφυλάκιο μέσω της μεθόδου που πρότεινε ο Burgt(2007). Θα εκτιμήσουμε την παράμετρο κοιλότητας μέσω της ελαχιστοποίησης του $RMSE$ και των προσεγγιστικών τύπων και θα συγκρίνουμε τα αποτελέσματα για το παρακάτω χαρτοφυλάκιο:

| Χαρτοφυλάκιο | Οφειλέτες A βαθμίδας | Οφειλέτες B βαθμίδας | Οφειλέτες Γ βαθμίδας | Οφειλέτες Δ βαθμίδας | Οφειλέτες E βαθμίδας |
|--------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 13 | 300 | 200 | 150 | 300 | 250 |

Η βαθμίδα A είναι αυτή με την μικρότερη πιστοληπτική ικανότητα και η E αυτή με τη μεγαλύτερη. Οι πιθανότητες αθέτησης για κάθε βαθμίδα για κάθε βαθμίδα του χαρτοφυλακίου 13 παρατίθενται στον επόμενο πίνακα.

| Χαρτοφυλάκιο | p_A | p_B | p_Γ | p_Δ | p_E |
|--------------|-------|-------|------------|------------|--------|
| 13 | 3% | 1% | 0.1% | 0.05% | 0.001% |

Κάνοντας προσομοίωση για το εν λόγω χαρτοφυλάκιο και χρησιμοποιώντας την σχέση (4.17) προέκυψε ότι η $RMSE = 0.029$ και $k = 10.98$. Επίσης τα AUC , AR εκτιμήθηκαν 0.86 και 0.73 αντίστοιχα. Οι πιθανότητες αθέτησης για κάθε βαθμίδα του χαρτοφυλακίου 13 δίνονται παρακάτω.

| | |
|------------|--------|
| p_A | 2.40% |
| p_B | 0.56% |
| p_Γ | 0.22% |
| p_Δ | 0.07% |
| p_E | 0.018% |

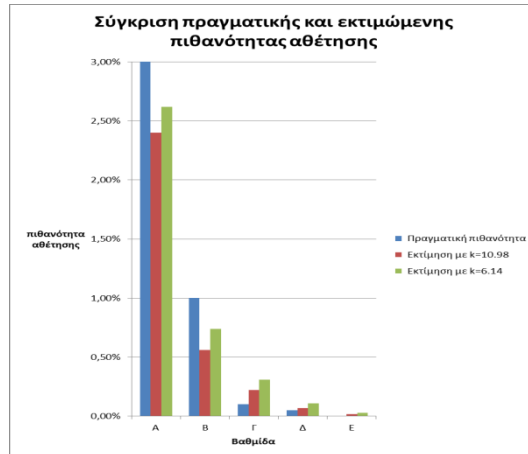
Πίνακας 5.8 Εκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 13 για $k = 10.98$

Χρησιμοποιώντας στην προσομοίωση τον σύνθετο κανόνα του τραπεζίου προκειμένου να εκτιμήσουμε τους δείκτες AUC και AR λαμβάνουμε τις τιμές 0.83 και 0.67 αντίστοιχα. Εκτιμώντας την παράμετρο κοιλότητας από την σχέση (4.13) έχουμε $k = 6.14$. Σε αυτή την περίπτωση οι εκτιμώμενες πιθανότητες αθέτησης είναι:

| | |
|------------|--------|
| p_A | 2.62% |
| p_B | 0.74% |
| p_Γ | 0.31% |
| p_Δ | 0.11% |
| p_E | 0.029% |

Πίνακας 5.9 Εκτίμηση των πιθανοτήτων αθέτησης για το χαρτοφυλάκιο 13 για $k = 6.14$

Διαπιστώνουμε πως η χρήση διαφορετικών μεθόδων για την εκτίμηση του k μας οδήγησε σε αρκετά διαφορετικές τιμές της παραμέτρου. Ελαχιστοποιώντας το $RMSE$ η τιμή που λαμβάνουμε για το k είναι σχεδόν διπλάσια σε σχέση με την τιμή που προκύπτει από την χρήση των προσεγγιστικών τύπων. Παρόλα αυτά και στις 2 περιπτώσεις είναι σαφές πως το μοντέλο βαθμονόμησης έχει υψηλή διακριτική ισχύ. Παρατηρούμε επίσης πως με την εκτίμηση του k μέσω του $RMSE$ οι πιθανότητες αθέτησης προέκυψαν πιο μικρές σε σύγκριση με την περίπτωση που έγινε χρήση των προσεγγιστικών τύπων. Όπως προκύπτει από τα αποτελέσματα των Πινάκων 5.8 και 5.9 η πιθανότητα αθέτησης για τις 2 χαμηλότερες βαθμίδες πιστοληπτικής ικανότητας A και B υποεκτιμήθηκαν, σε μικρό βαθμό. Για τις βαθμίδες Γ και Δ οι εκτιμήσεις υπερβαίνουν ελαφρώς τις πραγματικές πιθανότητες αθέτησης, ενώ για την βαθμίδα E η πιθανότητα αθέτησης έχει υπερεκτιμηθεί σε πολύ μεγάλο βαθμό. Στα παραπάνω συμπεράσματα μπορούμε να καταλήξουμε και με από το Σχήμα (5.13).



Σχήμα 5.13 Σύγκριση της πραγματικής και εκτιμώμενης πιθανότητας αθέτησης για το χαρτοφυλάκιο 13

Εφόσον παρατηρήσαμε πως στις 2 βαθμίδες με τους λιγότερο φερέγγυους οφειλέτες η προτεινόμενη μεθοδολογία οδηγεί σε υποεκτίμηση της πραγματικής πιθανότητας αθέτησης, έχει νόημα να εξετάσουμε αν η χρήση κάποιας από τις προηγούμενες μεθόδους δύναται να λύσει το πρόβλημα. Στον πίνακα που ακολουθεί θα παρουσιάσουμε τις εκτιμήσεις που προκύπτουν με την χρήση της συντηρητικής εκ των προτέρων κατανομής, της ομοιόμορφης στο (0,1) και του πιο συντηρητικού εκτιμητή χωρίς να υπάρχει εξάρτηση του δείγματος με $\gamma=75\%$.

| Βαθμίδα | Συντηρητική εκ των προτέρων | Ομοιόμορφη στο (0,1) | Συντηρητικός εκτιμητής με $\gamma=75\%$ |
|---------|-----------------------------|----------------------|---|
| A | 3.26% | 3.25% | 3.93% |
| B | 1.43% | 1.42% | 2.80% |

Πίνακας 5.10 Εκτίμηση της πιθανότητας αθέτησης για τις βαθμίδες A και B με χρήση της συντηρητικής εκ των προτέρων, της ομοιόμορφης στο (0,1) και του συντηρητικού εκτιμητή

Όπως αναμέναμε και από τα αποτελέσματα της προσομοίωσης των προηγούμενων Ενοτήτων διαπιστώνουμε πως το πρόβλημα της υποεκτίμησης μπορεί να λυθεί με χρήση των προηγούμενων μεθόδων. Η χρήση της ομοιόμορφης ή της συντηρητικής κατανομής ως εκ τω προτέρων μας οδηγεί σε εκτιμήσεις αρκετά κοντά στις πραγματικές. Μέσω του συντηρητικού εκτιμητή έχουμε υπερεκτίμηση των ζητούμενων πιθανοτήτων και ιδιαίτερα για την βαθμίδα B όπου η εκτιμώμενη τιμή είναι σχεδόν 3 φορές μεγαλύτερη της πραγματικής.

6. Σύνοψη

Στο πρώτο κεφάλαιο της εργασίας γίνεται μια εισαγωγή σε βασικές έννοιες του πιστωτικού κινδύνου και τον τρόπο με τον οποίο τα ΠΙ καλούνται να αποφασίσουν τη χορήγηση ή μη ενός δανείου σε κάποιον πελάτη βάσει κάποιων κριτηρίων. Γίνεται επίσης αναφορά σε συνήθεις στατιστικές τεχνικές και στον λόγο που αυτές αποτυγχάνουν να εκτιμήσουν αξιόπιστα την πιθανότητα αθέτησης σε ένα χαρτοφυλάκιο υψηλής φερεγγυότητας.

Οι τρεις μέθοδοι που αναφέρουμε, προκειμένου να επιτύχουμε τη ζητούμενη εκτίμηση είναι η μέθοδος του πιο συντηρητικού εκτιμητή, η προσέγγιση Bayes και η καμπύλη ισχύος. Από τα αριθμητικά παραδείγματα του 2^{ου} Κεφαλαίου διαπιστώνουμε πως εν γένει η πιο συντηρητική αρχή εκτίμησης δίνει πολύ υψηλές εκτιμήσεις για τις πιθανότητες αθέτησης, ειδικά αν θεωρήσουμε πως υπάρχει εξάρτηση των παρατηρήσεων του δείγματος για μονοετή περίοδο παρατήρησης. Το πρόβλημα αυτό μπορεί να λυθεί, αν εφαρμόσουμε τη μέθοδο έχοντας περίοδο παρατήρησης πάνω από ένα έτος. Το γεγονός όμως, πως για να εφαρμοστεί η μέθοδος πρέπει να προσδιορισθούν τρεις παράμετροι, αποτελεί πρόβλημα.

Στο 3^ο Κεφάλαιο παραθέσαμε την προσέγγιση Bayes. Κρίσιμο σημείο της εκτίμησης αποτελεί η επιλογή της εκ των προτέρων κατανομής. Επειδή η εκ των προτέρων κατανομή εκφράζει τις προσδοκίες που υπάρχουν για την πραγματική τιμή της πιθανότητας αθέτησης, ο ειδικός που επιλέγει την κατανομή οφείλει να είναι αμερόληπτος στην επιλογή αυτής, καθώς σε αντίθετη περίπτωση τα αποτελέσματα δεν θα είναι αξιόπιστα. Μέσω της προσέγγισης Bayes και της πιο συντηρητικής αρχής εκτίμησης μπορούμε να εκτιμήσουμε την πιθανότητα αθέτησης ακόμα και στην περίπτωση που στο δείγμα μας δεν έχει συμβεί καμία αθέτηση. Αυτή τους η ιδιότητα τις καθιστά ιδιαίτερες χρήσιμες, διότι σε ένα χαρτοφυλάκιο υψηλής φερεγγυότητας είναι πολύ πιθανό να μην έχουν παρατηρηθεί αθετήσεις.

Η καμπύλη ισχύος ως επί το πλείστον χρησιμοποιείται, προκειμένου να αξιολογηθεί το μοντέλο βαθμονόμησης που χρησιμοποιεί ένα ΠΙ. Στο 4^ο Κεφάλαιο αναλύουμε τον τρόπο με τον οποίο μοντελοποιούμε την καμπύλη ισχύος μέσω μιας κοίλης συνάρτησης, προκειμένου να εκτιμήσουμε την πιθανότητα αθέτησης για ένα χαρτοφυλάκιο. Ένα πολύ σημαντικό μειονέκτημα της μεθόδου είναι πως είναι αδύνατον να εκτιμηθεί η πιθανότητα αθέτησης, αν δεν έχει παρατηρηθεί καμία αθέτηση στο χαρτοφυλάκιο.

Στο 5^ο και τελευταίο Κεφάλαιο της διπλωματικής προσομοιώνουμε σενάρια για κάποια χαρτοφυλάκια έχοντας καθορίσει τις πραγματικές τιμές των πιθανοτήτων, ώστε να αξιολογήσουμε τις μεθόδους. Από τα αποτελέσματα διαπιστώνουμε πως η μέθοδος της πιο συντηρητικής εκτίμησης υπερεκτιμά σε πολύ μεγάλο βαθμό την πιθανότητα

αθέτησης στην περίπτωση που θεωρούμε ότι υπάρχει εξάρτηση μεταξύ των παρατηρήσεων του δείγματος με την περίοδο παρατήρησης να είναι 1 έτος . Στην πολυετή περίοδο παρατήρησης οι εκτιμήσεις είναι πιο κοντά στις πραγματικές, ωστόσο το αν θα υπερεκτιμήσουμε ή θα υποεκτιμήσουμε την πιθανότητα αθέτησης εξαρτάται από την επιλογή του συντελεστή γ . Από την προσομοίωση προέκυψε πως χρησιμοποιώντας την προσέγγιση Bayes και επιλέγοντας *Beta* εκ των προτέρων κατανομή, η εκτίμηση της πιθανότητας εξαρτάται σε πολύ μεγάλο βαθμό από την επιλογή των παραμέτρων α και β . Στην περίπτωση που επιλέξουμε την συντηρητική κατανομή ή την ομοιόμορφη ως εκ των προτέρων, η πιθανότητα υπερεκτιμάται σε μικρό βαθμό. Τέλος, η χρήση της καμπύλης ισχύος φαίνεται να οδηγεί σε υποεκτίμηση της πιθανότητας αθέτησης για τους οφειλέτες των πιο επικίνδυνων βαθμίδων και σε υπερεκτίμηση για τους οφειλέτες των ακίνδυνων.

Παράρτημα

Για τα αριθμητικά αποτελέσματα που παρουσιάστηκαν στα προηγούμενα κεφάλαια χρησιμοποιήθηκε η γλώσσα προγραμματισμού R. Παρακάτω θα παραθέσουμε τους κώδικες βάσει των οποίων υπολογίστηκαν οι πίνακες.

Κεφάλαιο 2

Θα αναφέρουμε τους κώδικες μέσω των οποίων υπολογίστηκαν οι πιθανότητες αθέτησης για το χαρτοφυλάκιο 1 χρησιμοποιώντας ως συντελεστή εμπιστοσύνης το 50%

Ενότητα 2.1

Κώδικας 1

```
na=100
nb=250
ng=150
g=0.50
pa=1-(1-g)^(1/(na+nb+ng))
pb=1-(1-g)^(1/(nb+ng))
pg=1-(1-g)^(1/ng)
```

Ο Κώδικας 1 αποτελεί τον πιο απλό τρόπο υπολογισμού για την περίπτωση που περιγράφεται στην Ενότητα 2.1 . Εναλλακτικά μπορεί να χρησιμοποιηθεί το πακέτο LDPD και ο παρακάτω κώδικας

Κώδικας 2

```
library(LDPD)
na=100
nb=250
ng=150
da=0
db=0
dg=0
confidenceinterval=0.50
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
PTOnePeriodPD(portfolio,defaults,conf.interval=confidenceinterval) .
```

Ενότητα 2.2

Και σε αυτήν την περίπτωση μπορούμε να πάρουμε αριθμητικά αποτελέσματα με 2 τρόπους. Ο 1^{ος} τρόπος περιγράφεται στο Κεφάλαιο 5 και βασίζεται στην χρήση ποσοστιαίων σημείων της κατανομής Beta ενώ ο 2^{ος} είναι μέσω του πακέτου LDPD

Κώδικας 3

```
na=100
nb=250
ng=150
nall=na+nb+ng
da=0
```

```

db=1
dg=1
dall=da+db+dg
a=dall+1
b=nall-dall
qbeta(0.50,a,b)
anew=db+dg+1
bnew=nb+ng-db-dg
qbeta(0.50,anew,bnew)
atonos=dg+1
btonos=ng-dg
qbeta(0.50,atonos,btonos)

```

Κώδικας 4

```

library(LDPD)
na=100
nb=250
ng=150
da=0
db=1
dg=1
confidenceinterval=0.50
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
PTOnePeriodPD(portfolio,defaults,conf.interval=confidenceinterval)

```

Ενότητα 2.3

Για την συγκεκριμένη ενότητα, όπως και για τις 2.4 και 2.5, τα αποτελέσματα προέκυψαν με την χρήση του πακέτου LDPD και της εντολής PTMultiPeriodPD. Ο κώδικας δίνεται παρακάτω

Κώδικας 5

```

library(LDPD)
na=100
nb=250
ng=150
da=0
db=0
dg=0
confidenceinterval=0.50
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
r=0.18
t=0
T=1
simulations=10^3
m=100
PD=0
for(i in 1:m)

```

```

{
X=PTMultiPeriodPD(portfolio,defaults,r,t,T,simulations,conf.interval=confidenceinte
rval)
PD=PD+X
}
PD/m

```

Ενότητα 2.4

Προφανώς χρησιμοποιείται ο ίδιος κώδικας αλλάζοντας απλά τον αριθμό των αθετήσεων

Κώδικας 6

```

library(LDPD)
na=100
nb=250
ng=150
da=0
db=1
dg=1
confidenceinterval=0.50
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
r=0.18
t=0
T=1
simulations=10^3
m=100
PD=0
for(i in 1:m)
{
X=PTMultiPeriodPD(portfolio,defaults,r,t,T,simulations,conf.interval=confidenceinte
rval)
PD=PD+X
}
PD/m

```

Ενότητα 2.5

Χρησιμοποιούμε και πάλι τον ίδιο κώδικα με αυτόν των Ενοτήτων 2.3 και 2.4 αλλάζοντας ωστόσο τον χρόνο T και την συσχέτιση t. Ο Κώδικας 7 αντιστοιχεί στην περίπτωση που δεν έχουν συμβεί αθετήσεις στο χαρτοφυλάκιο ενώ ο Κώδικας 8 είναι για την περίπτωση που έχουν παρατηρηθεί αθετήσεις

Κώδικας 7

```

library(LDPD)
na=100
nb=250
ng=150

```

```

da=0
db=0
dg=0
confidenceinterval=0.50
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
r=0.20
t=0.40
T=5
simulations=10^3
m=100
PD=0
for(i in 1:m)
{
X=PTMultiPeriodPD(portfolio,defaults,r,t,T,simulations,conf.interval=confidenceinte
rval)
PD=PD+X
}
PD/m

```

Κώδικας 8

```

library(LDPD)
na=100
nb=250
ng=150
da=0
db=1
dg=1
confidenceinterval=0.50
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
r=0.20
t=0.40
T=5
simulations=10^3
m=100
PD=0
for(i in 1:m)
{
X=PTMultiPeriodPD(portfolio,defaults,r,t,T,simulations,conf.interval=confidenceinte
rval)
PD=PD+X
}
PD/m

```

Κεφάλαιο 3

Θα αναφέρουμε τους κώδικες μέσω των οποίων υπολογίστηκαν οι πιθανότητες αθέτησης για το χαρτοφυλάκιο 1 σε κάθε μία από τις περιπτώσεις που αναλύθηκαν στις ενότητες του 3^{ου} Κεφαλαίου

Ενότητα 3.4.1&3.4.2

Για αυτές τις 2 ενότητες θα αναφέρουμε τις εκτιμήσεις σε έναν μόνο κώδικα καθώς οι υπολογισμοί είναι αρκετά εύκολοι.

Κώδικας 9

```
a=2.1
b=300
r=2
n=500
pbeta=(a+r)/(a+n+b)
pconservative=(r+1)/(n+1)
```

Ενότητα 3.4.3

Και στην περίπτωση της ομοιόμορφης εκ των προτέρων κατανομής ο κώδικας είναι αρκετά απλός καθώς όπως είδαμε στην αντίστοιχη ενότητα τα αποτελέσματα προκύπτουν άμεσα με χρήση της αθροιστικής συνάρτησης κατανομής της Beta.

Κώδικας 10

```
u=0.025
n=500
r=2
ektimisi=(r+1)*pbeta(u,r+2,n-r+1)/((n+2)*pbeta(u,r+1,n-r+1))
```

Κεφάλαιο 4

Ενότητα 4.4

Με τον επόμενο Κώδικα 11 καθορίζουμε την τιμή της παραμέτρου k για το χαρτοφυλάκιο 7 ελαχιστοποιώντας το $RMSE$ και υπολογίζουμε τα AUC και AR

Κώδικας 11

```
library(mltools)
x=c(100,175,225,150,50)
cumx=cumsum(x)/sum(x)
y=c(3,2,0,0,0)
cumy=cumsum(y)/sum(y)
k=seq(1,40,0.01)
RMS=10^7
concavity=0
for(i in k)
{
predictions=(1-exp(-i*cumx))/(1-exp(-i))
RMSnew=rmse(predictions,cumy)
if(RMSnew<RMS){
RMS=RMSnew
concavity=i
}
}
AUCreal=(1/1-exp(-concavity))-(1/concavity)
```

$$AR_{real}=2*AUC_{real}-1$$

Προκειμένου να υπολογίσουμε τους δείκτες AUC, AR και την παράμετρο k χρησιμοποιώντας τον σύνθετο κανόνα του τραπεζίου και τους προσεγγιστικούς τύπους για το χαρτοφυλάκιο 7, χρησιμοποιούμε τον Κώδικα 12

Κώδικας 12

```
x=c(100,175,225,150,50)
cumx=cumsum(x)/sum(x)
y=c(3,2,0,0,0)
cumy=cumsum(y)/sum(y)
trapezoidal=rep(0,length(x))
for(i in 1:length(x))
{
if(i==1){
trapezoidal[i]=cumx[i]*cumy[i]*0.5
}else{
trapezoidal[i]=(cumx[i]-cumx[i-1])*(cumy[i]+cumy[i-1])*0.5
}
}
AUC=sum(trapezoidal)
AR=2*AUC-1
k1=1/(1-AUC)
```

Έχοντας καθορίσει πλέον την τιμή του k με έναν από τους προηγούμενους Κώδικες μπορούμε να εκτιμήσουμε τις πιθανότητες αθέτησης με τον επόμενο Κώδικα. Η μεταβλητή concavity περιέχει την τιμή του k .

Κώδικας 13

```
CT=sum(y)/sum(x)
Xr=rep(0,length(x))
for(i in 1:length(x))
{
Xr[i]=(sum(x[1:i])-x[i]/2)/sum(x)
}
PD=concavity*CT*exp(-concavity*Xr)/(1-exp(-concavity))
```

Κεφάλαιο 5

Ενότητα 5.1.1

Κώδικας 14

```
na=80
nb=300
ng=370
pa=0.1/100
pb=0.5/100
pg=1/100
g=0.50
EktimisiPa=0
EktimisiPb=0
```

```

EktimisiPg=0
m=1000
for(i in 1:m)
{
Y1=rbinom(na,1,pa)
da=sum(Y1)
Y2=rbinom(nb,1,pb)
db=sum(Y2)
Y3=rbinom(ng,1,pg)
dg=sum(Y3)
n=na+nb+ng
d=da+db+dg
a=d+1
b=n-d
Xa=qbeta(g,a,b)
anew=db+dg+1
bnew=nb+ng-db-dg
Xb=qbeta(g,anew,bnew)
atonos=dg+1
btonos=ng-dg
Xg=qbeta(g,atonos,btonos)
EktimisiPa=EktimisiPa+Xa
EktimisiPb=EktimisiPb+Xb
EktimisiPg=EktimisiPg+Xg
}
EktimisiPa*100/m
EktimisiPb*100/m
EktimisiPg*100/m

```

Στον κώδικα 11 υπολογίζουμε την εκτίμηση των πιθανοτήτων για τον Πίνακα 5.1 βάσει της μεθόδου των ποσοστιαίων σημείων που αναλύθηκε στην Ενότητα 5.1.1. Εναλλακτικά θα μπορούσε να χρησιμοποιηθεί ο παρακάτω κώδικας

Κώδικας 15

```

library(LDPD)
p1=0.1/100
p2=0.5/100
p3=1/100
na=80
nb=300
ng=370
confidenceinterval=0.50
X=c(0,0,0)
m=1000
for(i in 1:m)
{
Y1=rbinom(na,1,p1)
da=sum(Y1)
Y2=rbinom(nb,1,p2)
db=sum(Y2)
Y3=rbinom(ng,1,p3)

```

```

dg=sum(Y3)
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
Y=PTOnePeriodPD(portfolio,defaults,conf.interval=confidenceinterval)
X=X+Y
}
X/m

```

Ενότητα 5.1.2

Κώδικας 16

```

library(LDPD)
p1=1/100
p2=0.5/100
p3=1/100
na=80
nb=300
ng=370
confidenceinterval=0.50
X=c(0,0,0)
m=1000
for(i in 1:m)
{
Y1=rbinom(na,1,p1)
da=sum(Y1
Y2=rbinom(nb,1,p2)
db=sum(Y2
Y3=rbinom(ng,1,p3)
dg=sum(Y3
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
r=0.03
t=0
T=1
simulations=10^3
Y=PTMultiPeriodPD(portfolio,defaults,r,t,T,simulations,conf.interval=confidenceinte
rval)
X=X+Y
}
z=X/m

```

Ενότητα 5.1.3

Ο κώδικας 14 δίνει τις εκτιμήσεις των πιθανοτήτων για τον Πίνακα 5.3, για τους Πίνακες 5.4 και 5.5 αρκεί να τρέξουμε πάλι τον κώδικα 14 αλλάζοντας τα ρ , τ .

Κώδικας 17

```

library(LDPD)
p1=0.6/100
p2=1.5/100

```

```

p3=3/100
na=300
nb=200
ng=250
confidenceinterval=0.50
X=c(0,0,0)
m=1000
for(i in 1:m)
{
Y1=rbinom(na,1,p1)
da=sum(Y1)
Y2=rbinom(nb,1,p2)
db=sum(Y2)
Y3=rbinom(ng,1,p3)
dg=sum(Y3)
portfolio<-c(ng,nb,na)
defaults<-c(dg,db,da)
r=0.15
t=0.30
T=5
simulations=10^3
Y=PTMultiPeriodPD(portfolio,defaults,r,t,T,simulations,conf.interval=confidence
interval)
X=X+Y
}
z=X/m

```

Ενότητα 5.2.1

Κώδικας 18

```

p=1/100
n=1200
a=2
b=500
atonos=4
btonos=50
m=1000
X=0
Z=0
MLE=0
for(i in 1:m)
{
Y=rbinom(n,1,p)
Y
r=sum(Y)
estimation=(a+r)/(a+b+n)
X=X+estimation
estimation2=(atonos+r)/(atonos+btonos+n)
Z=Z+estimation2
mle=r/n

```

```

MLE=MLE+mle
}
X*100/m
Z*100/m
MLE*100/m

```

Ενότητα 5.2.2

Κώδικας 19

```

library(LDPD)
n=1500
p=0.3/100
m=1000
conserv=0
for(i in 1:m)
{
Y=rbinom(n,1,p)
r=sum(Y)
conservative_ektimisi=(r+1)/(n+1)
conserv=conserv+conservative_ektimisi
}
conserv/m

```

Ενότητα 5.2.3

Κώδικας 20

```

library(LDPD)
u=0.025
n=1000
p=1/100
m=1000
unif=0
conserv=0
prudent0.50=0
prudent0.75=0
for(i in 1:m)
{
Y=rbinom(n,1,p)
r=sum(Y)
uniform_ektimisi=(r+1)*pbeta(u,r+2,n-r+1)/((n+2)*pbeta(u,r+1,n-r+1))
unif=unif+uniform_ektimisi
conservative_ektimisi=(r+1)/(n+1)
conserv=conserv+conservative_ektimisi
ektimisi_0.50=PTOnePeriodPD(n,r,conf.interval=0.5)
prudent0.50=prudent0.50+ektimisi_0.50
ektimisi_0.75=PTOnePeriodPD(n,r,conf.interval=0.75)
prudent0.75=prudent0.75+ektimisi_0.75
}
prudent0.75/m

```

```
prudent0.50/m
conserv/m
unif*100/m
```

Ενότητα 5.3

Οι Κώδικες 21 και 22 αφορούν την προσομοίωση που έγινε στην Ενότητα 5.3. Στον Κώδικα 21 υπολογίζουμε τις πιθανότητες αθέτησης εκτιμώντας το k βάσει της ελαχιστοποίησης του RMSE ενώ στον Κώδικα 22 χρησιμοποιούμε τον σύνθετο κανόνα του τραπεζίου και τον προσεγγιστικό τύπο για την τιμή του k .

Κώδικας 21

```
library(mltools)
n1=300
n2=200
n3=150
n4=300
n5=250
x=c(n1,n2,n3,n4,n5)
cumx=cumsum(x)/sum(x)
Xr=rep(0,length(x))
for(i in 1:length(x))
{
Xr[i]=(sum(x[1:i])-x[i]/2)/sum(x)
}
p1=3/100
p2=1/100
p3=0.1/100
p4=0.05/100
p5=0.001/100
k=seq(1,40,0.01)
m=1000
Preal=rep(0,length(x))
CT=0
con=0
RMSE=0
AUC=0
AR=0
for(i in 1:m)
{
y1=rbinom(n1,1,p1)
y2=rbinom(n2,1,p2)
y3=rbinom(n3,1,p3)
y4=rbinom(n4,1,p5)
y5=rbinom(n5,1,p5)
d1=sum(y1)
d2=sum(y2)
d3=sum(y3)
d4=sum(y4)
d5=sum(y5)
y=c(d1,d2,d3,d4,d5)
```

```

cumy=cumsum(y)/sum(y)
ct=sum(y)/sum(x)
CT=CT+ct
RMS=10^7
for(i in k)
{
predictions=(1-exp(-i*cumx))/(1-exp(-i))
RMSnew=rmse(predictions,cumy)
if(RMSnew<=RMS){
RMS=RMSnew
concavity=i
}
}
RMSE=RMSE+RMS
con=con+concavity
AUCreal=(1/1-exp(-concavity))-(1/concavity)
ARreal=2*AUCreal-1
PDrconca=concavity*ct*exp(-concavity*Xr)/(1-exp(-concavity))
Preal=Preal+PDrconca
AUC=AUC+AUCreal
AR=AR+ARreal
}
mesoCT=CT/m
mesoRMSE=RMSE/m
mesoConcavity=con/m
mesoPD=Preal*100/m
mesoAUC=AUC/m
mesoAR=AR/m
Κώδικας 22
library(mltools)
n1=300
n2=200
n3=150
n4=300
n5=250
x=c(n1,n2,n3,n4,n5)
cumx=cumsum(x)/sum(x)
Xr=rep(0,length(x))
for(i in 1:length(x))
{
Xr[i]=(sum(x[1:i])-x[i]/2)/sum(x)
}
p1=3/100
p2=1/100
p3=0.1/100
p4=0.05/100
p5=0.001/100
k=seq(1,40,0.01)
m=1000
Ptrapezoidal=rep(0,length(x))

```



```

CT=0
contrap=0
ARtrapezoidal=0
AUCtrapezoidal=0
for(i in 1:m)
{
y1=rbinom(n1,1,p1)
y2=rbinom(n2,1,p2)
y3=rbinom(n3,1,p3)
y4=rbinom(n4,1,p5)
y5=rbinom(n5,1,p5)
d1=sum(y1)
d2=sum(y2)
d3=sum(y3)
d4=sum(y4)
d5=sum(y5)
y=c(d1,d2,d3,d4,d5)
cumy=cumsum(y)/sum(y)
ct=sum(y)/sum(x)
CT=CT+ct
RMS=10^7
trapezoidal=rep(0,length(x))
for(i in 1:length(x))
{
if(i==1){
trapezoidal[i]=cumx[i]*cumy[i]*0.5
}else{
trapezoidal[i]=(cumx[i]-cumx[i-1])*(cumy[i]+cumy[i-1])*0.5
}
}
auctrapezoidal=sum(trapezoidal)
artrapezoidal=2*auctrapezoidal-1
concatrapezoidal=1/(1-auctrapezoidal)
contrap=contrap+concatrapezoidal
PDtrapezoidal=concatrapezoidal*ct*exp(-concatrapezoidal* $X_r$ )/(1-exp(-
concatrapezoidal))
Ptrapezoidal=Ptrapezoidal+PDtrapezoidal
AUCtrapezoidal=AUCtrapezoidal+auctrapezoidal
ARtrapezoidal=ARtrapezoidal+artrapezoidal
}
mesoCT=CT/m
mesoConcavity=contrap/m
mesoAUC=AUCtrapezoidal/m
mesoAR=ARtrapezoidal/m
Pithanotites=Ptrapezoidal*100/m

```

Αναφορές

1. Altman E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, **23**, 189-209.
2. Basel Committee on Banking Supervision (2005). Basel Committee Newsletter No.6: Validation of Low-Default Portfolios in the Basel II Framework, Discussion Paper, Bank of International Settlements.
3. Burgt M. (2008). Calibrating low-default portfolios, using the Cumulative Accuracy Profile, *Journal of Risk Model Validation*, **1**, 17-33.
4. Casella G and Berger R (2002). *Statistical Inference*. Duxbury Press, Second Edition.
5. Clifford T., Marianski A., Sebestyen K. (2013). Low Default Portfolio modeling. Probability of Default Calibration Conundrum, Deloitte.
6. Dzidzevičiūtė L. (2012). Estimation of Default Probability for Low Default Portfolios. *Ekonomika*, **91**, 132-156.
7. Engelmann B., Hayden E., Tasche D. (2003). Measuring the discriminative power of rating systems. Discussion Paper No.1/2003 Series 2: Banking and Financial Supervision, Deutsche Bundesbank.
8. Engelmann B. and Rauhmeier R. (2011). *The Basel II Risk Parameters: Estimation, Validation, Stress Testing - with Applications to Loan Risk Management*. Springer Berlin Heidelberg, Second edition.
9. Kiefer N.M. (2007). The Probability Approach to Default Estimation, *Risk*, **20**, 146-150.
10. Kiefer N.M. (2009). Default Estimation for Low Default Portfolios, *Journal of Empirical Finance*, **16**, 164-173.
11. Pluto K. and Tasche D. (2005). Thinking positively, *Risk*, **18**, 72-78.
12. Tasche D. (2013). Bayesian estimation of probabilities of default for low default portfolios, *Journal of Risk Management in Financial Institutions*, **1**, 302-326.
13. Venter E. (2016). Probability of Default Calibration for Low Default Portfolios: Revisiting the Bayesian Approach, University of Stellenbosch.

Ιστοσελίδες

<https://CRAN.R-project.org/package=LDPD>

<https://CRAN.R-project.org/package=mltools>