



Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
Τμήμα Μαθηματικών
Τομέας Στατιστικής και Επιχειρησιακής Έρευνας

ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΣΤΟΧΑΣΤΙΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΟΔΥΣΣΕΑ ΚΑΝΑΒΕΤΑ

Αθήνα
2014

Στην γυναίκα μου Νάνσυ
και στον γιό μας Θανάση

Ολοκληρώνοντας τη διδακτορική διατριβή θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντά μου, καθηγητή κ. Α. Μπουρνέτα για την καθοδήγησή του, την υπομονή του και την επιμονή του κατά την διάρκεια όλων αυτών των ετών. Μου πρόσφερε απλόχερα υποστήριξη για την εκπόνηση της διατριβής και μου έδειξε τον δρόμο για να πορευτώ στον επιστημονικό μου βίο. Χάρη σε αυτόν είχα την δυνατότητα να συνεργαστώ κατά την διάρκεια της εκπόνησης της διδακτορικής μου διατριβής με σημαίνουσες επιστημονικές προσωπικότητες εντός και εκτός Ελλάδος.

Θα ήθελα ακόμη να ευχαριστήσω τον καθηγητή κ. Χ. Παπαγεωργίου και τον καθηγητή κ. Γ. Κοολε που συμμετείχαν στην τριμελή συμβουλευτική επιτροπή για τη στήριξή τους και τις γνώσεις που μου μετέδωσαν σε όλη τη διάρκεια των μεταπτυχιακών σπουδών μου. Επιπλέον, ευχαριστώ τους αναπληρωτές καθηγητές κ. Ε. Κυριακίδη, κ. Κ. Μηλολιδάκη, κ. Α. Οικονόμου και κ. Δ. Φακίνο που μου έκαναν την τιμή να συμμετέχουν στην επταμελή εξεταστική επιτροπή της διδακτορικής διατριβής μου.

Τέλος, θα ήθελα να ευχαριστήσω τη σύζυγο μου Νάνσυ η οποία πέρασε μαζί μου όλες τις χαρές και τις δυσκολίες που μπορεί να έχει μια πολυετής διαδρομή για την εκπόνηση μιας διατριβής, όταν συμβαίνει ταυτόχρονα με την οικογενειακή μας πορεία. Ευχαριστώ το γιό μας ο οποίος μαζί με την σύζυγο μου είναι ό,τι πολυτιμότερο έχω και κάνουν τη καθημερινότητά μου πιο ευχάριστη. Επίσης θέλω να ευχαριστήσω τους γονείς μας οι οποίοι μας βοήθησαν να κάνουμε πραγματικότητα αυτό που ονειρευτήκαμε.

Περιεχόμενα

| | | |
|----------|--|-----------|
| 1 | Εισαγωγή | 9 |
| 1.1 | Ακολουθιακή δειγματοληψία | 9 |
| 1.2 | Προβλήματα αποθέματος | 11 |
| 1.3 | Δομή της διατριβής | 12 |
| 2 | Προσαρμοστικές πολιτικές για ακολουθιακή δειγματοληψία υπό ελλιπή πληροφόρηση και ασυμπτωτικό περιορισμό στο κόστος | 15 |
| 2.1 | Εισαγωγή | 15 |
| 2.2 | Περιγραφή μοντέλου | 17 |
| 2.2.1 | Πλαίσιο πλήρους πληροφόρησης | 18 |
| 2.2.2 | Πλαίσιο ελλιπούς πληροφόρησης | 21 |
| 2.3 | Κατασκευή μιας συνεπούς πολιτικής | 23 |
| 2.4 | Ταχύτητα σύγκλισης και προσομοιώσεις | 30 |
| 2.5 | Ανασκόπηση και επεκτάσεις | 34 |
| 3 | Αποδοτική πολιτική για ακολουθιακή δειγματοληψία κάτω από ελλιπή πληροφόρηση και περιορισμούς σε κάθε βήμα | 37 |
| 3.1 | Εισαγωγή | 37 |
| 3.2 | Περιγραφή μοντέλου | 40 |
| 3.3 | Ένα κάτω φράγμα για την συνάρτηση απώλειας | 48 |
| 3.4 | Κατασκευή μιας ασυμπτωτικά αποδοτικής πολιτικής | 55 |
| 4 | Πολιτικές παραγγελιών για δύο προϊόντα με ανταλλαξιμότητα | |

| | |
|--|-----------|
| στη ζήτηση | 65 |
| 4.1 Εισαγωγή | 65 |
| 4.2 Περιγραφή μοντέλου | 68 |
| 4.2.1 Προσέγγιση 1 | 69 |
| 4.2.2 Προσέγγιση 2 | 72 |
| 4.3 Προσέγγιση 1: Μέτρα απόδοσης | 74 |
| 4.3.1 Μεταβατική κατανομή | 74 |
| 4.3.2 Αναμενόμενο όφελος | 80 |
| 4.3.3 Submodularity του αναμενόμενου οφέλους | 81 |
| 4.4 Προσέγγιση 2: Μέτρα απόδοσης | 83 |
| 4.4.1 Στάσιμη κατανομή | 83 |
| 4.4.2 Αναμενόμενο όφελος | 86 |
| 4.4.3 Submodularity του αναμενόμενου οφέλους | 87 |
| 4.5 Αλγόριθμος | 88 |
| 4.6 Αριθμητική ανάλυση | 90 |
| 4.7 Ανασκόπηση και επεκτάσεις | 94 |
| Βιβλιογραφία | 97 |

Κεφάλαιο 1

Εισαγωγή

1.1 Ακολουθιακή δειγματοληψία

Μέχρι την εμφάνιση της ακολουθιακής ανάλυσης, η στατιστική θεωρία είχε περιοριστεί στη μελέτη και την ανάλυση επιλογής πειραμάτων στα οποία το μέγεθος και τα στοιχεία των δειγμάτων ήταν εντελώς προσδιορισμένα πριν το πείραμα ξεκινήσει. Η αλλαγή άρχισε να υπάρχει με την εισαγωγή θεωρίας σχετικά με την ακολουθιακή διάταξη πειραμάτων, στα οποία το μέγεθος και τα στοιχεία των δειγμάτων δεν είναι σταθερά αλλά προκύπτουν ως συναρτήσεις των παρατηρήσεων. Η παρακάτω ιστορική ανασκόπηση των αρχών της εξακολουθητικής ανάλυσης βασίζεται στην εργασία του *Robbins* (1952).

Η πρώτη σημαντική γενίκευση από το σταθερό μέγεθος δείγματος ήρθε στο πλαίσιο του βιομηχανικού ελέγχου ποιότητας, με την μέθοδο επιθεώρησης διπλής δειγματοληψίας των *Dodge* και *Romig* το 1929 και το 1941. Σε αυτή υπάρχει μόνο ένας πληθυσμός για δειγματοληψία, και η ερώτηση είναι αν η αναλογία των ελλατωματικών σε μία παρτίδα είναι σε ένα δεδομένο επίπεδο. Ένα προκαταρκτικό δείγμα από n_1 αντικείμενα έχει επιλεγεί από μια παρτίδα και έστω x από αυτά είναι ελλατωματικά. Αν το x είναι μικρότερο από μια σταθερή τιμή a η παρτίδα γίνεται δεκτή χωρίς περαιτέρω δειγματοληψία, αλλά αν $a \leq x \leq b$ τότε ένα δεύτερο δείγμα μεγέθους n_2 επιλέγεται και η απόφαση του να δεχτούμε ή να απορρίψουμε την παρτίδα μπαίνει στον αριθμό των ελλατωματικών σε ολόκληρο πλέον το δείγμα μεγέθους $n_1 + n_2$ αντικειμένων. Το συνολικό μέγεθος δείγματος n είναι μια τυχαία μεταβλητή με δύο τιμές, n_1 και $n_1 + n_2$, και η τιμή n

είναι στοχαστικά εξαρτημένη από τις παρατηρήσεις. Μία λογική επέκταση της ιδέας της διπλής δειγματοληψίας ήρθε κατά την διάρκεια του Δευτέρου Παγκοσμίου πολέμου με την ανάπτυξη, κυρίως από τον *Wald*, της ακολουθιακής ανάλυσης, στην οποία οι παρατηρήσεις είναι επιλεγμένες μία προς μία και η απόφαση να τερματίσουμε την δειγματοληψία και να δεχτούμε ή να απορρίψουμε την παρτίδα μπορεί να έρθει σε κάθε στάδιο της δειγματοληψίας. Το συνολικό μέγεθος δείγματος n τώρα είναι μια τυχαία μεταβλητή ικανή αρχικά να πάρει απείρως πολλές τιμές, αν και στην πράξη συνήθως τίθεται ένα πεπερασμένο άνω όριο για το n . Το όφελος των σταθερού μεγέθους ακολουθιακών δειγματοληψιών προκύπτει από το γεγονός ότι σε κάποιες περιστάσεις η επιλογή ενός ακολουθιακού σχεδίου επιφέρει μία αξιοσημείωτη μείωση του μέσου μεγέθους δείγματος που απαιτείται για να μειώσει την πιθανότητα λανθασμένης απόφασης σε ένα επιθυμητά χαμηλό επίπεδο.

Η εισαγωγή των ακολουθιακών μεθόδων δειγματοληψίας ελευθέρωσε τους στατιστικούς από τον περιορισμό των δειγμάτων σταθερού μεγέθους. Ωστόσο, δεν είναι μόνο το μέγεθος δείγματος που εμπλέκεται στο σχεδιασμό ενός πειράματος. Σε πολλά στατιστικά προβλήματα συναντούμε περισσότερους του ενός πληθυσμούς να εμπλέκονται και σε συνδυασμό με το κάθε πρόβλημα πρέπει να δούμε από ποιόν πληθυσμό θα πάρουμε δείγμα σε κάθε στάδιο. Ένα παράδειγμα θα βοηθήσει να κατανοήσουμε το προηγούμενο. Υποθέτουμε ότι έχουμε δύο πληθυσμούς κανονικά κατανομημένους με άγνωστες μέσες τιμές μ_1, μ_2 και διασπορές σ_1^2, σ_2^2 , και θέλουμε να εκτιμήσουμε την τιμή της διαφοράς $\mu_1 - \mu_2$. Για να εστιάσουμε στο θέμα του κεφαλαίου υποθέτουμε ότι το συνολικό μέγεθος δείγματος είναι n , σταθερό. Τώρα τίθεται το θέμα πώς πρέπει οι n παρατηρήσεις να επιλεγθούν ανάμεσα στους δύο πληθυσμούς. Αν ορίσουμε \bar{x}_1, \bar{x}_2 τους μέσους των δειγμάτων μεγέθους n_1, n_2 των δύο πληθυσμών τότε $\bar{x}_1 - \bar{x}_2$ είναι μια αμερόληπτη εκτιμήτρια της διαφοράς $\mu_1 - \mu_2$, με διασπορά $\sigma^2 = (\sigma_1^2/n_1) + (\sigma_2^2/n_2)$. Για σταθερό $n = n_1 + n_2$, το σ^2 ελαχιστοποιείται για $n_1/n_2 = \sigma_1/\sigma_2$. Αν ο προηγούμενος λόγος είναι γνωστός εκ των προτέρων τότε όλα είναι εντάξει. Αν ο λόγος είναι άγνωστος, και η δειγματοληψία μπορεί να γίνει σε δύο στάδια, τότε είναι βολικό να πάρουμε προκαταρκτικά δείγματα μεγέθους m από τον κάθε πληθυσμό, να χρησιμοποιήσουμε τις τιμές ώστε να λάβουμε μια εκτίμηση του σ_1/σ_2 , και οι υπόλοιπες $n - 2m$ παρατηρήσεις να κατανομηθούν στους δύο πληθυσμούς σύμφωνα με την εκτιμήτρια του σ_1/σ_2 . Και τώρα το θέμα είναι ποιά είναι η καλύτερη τιμή για το m . Αν το m είναι μικρό δεν

θα προκύψει ακριβής εκτιμήτρια για το σ_1/σ_2 . Αν το m είναι μεγάλο τότε οι $n - 2m$ παρατηρήσεις είναι τόσο λίγες που δεν επιτρέπουν την πλήρη χρησιμοποίηση της προσεγγιστικής γνώσης του σ_1/σ_2 . Γενικά μπορούμε να υποθέσουμε διατάξεις στις οποίες οι παρατηρήσεις επιλέγονται μία προς μία με την απόφαση να είναι από ποιό πληθυσμό πρέπει να είναι η επόμενη παρατήρηση υπολογιζόμενων και όλων των προηγούμενων, το συνολικό μέγεθος δείγματος n πρέπει να είναι σταθερό ή να είναι μια τυχαία μεταβλητή εξαρτημένη από τις παρατηρήσεις.

Παρά την συνολική απουσία θεωρίας, ένα αξιοσημείωτο πρωτοποριακό εγχείρημα στο πνεύμα της ακολουθιακής μελέτης επιτεύχθηκε το 1938 από τον *Mahalanobis* προσδιορίζοντας την επιφάνεια κάτω από το φυτό *jute* στην Βεγγάλη. Προκαταρκτικές εκτιμήσεις έγιναν σε μικρή κλίμακα για να εκτιμήσουν τις τιμές βασικών παραμέτρων, η γνώση των οποίων ήταν απαραίτητη για την επαρκή μελέτη μιας διαδοχικά μεγάλης κλίμακας καταμέτρησης. Σε μια επόμενη έκδοση ο *Mahalanobis* έδωσε προσοχή στην επιθυμία να επανεξετάσει την μελέτη κάθε πειράματος όσο τα δεδομένα συσσωρεύονται. Το θέμα βέβαια είναι πόσο καλά μπορούμε να το κάνουμε αυτό.

Είμαστε υπόχρεοι στον *Wald* (1950) για την πρώτη αξιοσημείωτη συμβολή στην θεωρία της ακολουθιακής μελέτης. Η εργασία του θέτει το πρόβλημα σε πλήρη γενικότητα και δίνει τη σκιαγράφηση μιας γενικής επαγωγικής μεθόδου ως λύση. Η πιθανότητα σύνθετων προβλημάτων ήταν μεγάλη, αφού υπεισέρχονται πολύπλοκοι υπολογισμοί με δεσμευμένες πιθανότητες, και σαφείς τρόποι δεν ήταν ακόμα διαθέσιμοι για χειρισμό προβλημάτων πρακτικού ενδιαφέροντος. Παρ' όλα αυτά τα επόμενα αποτελέσματα στη θεωρία ακολουθιακής μελέτης είναι μεγίστης σημασίας για τη μαθηματική στατιστική.

1.2 Προβλήματα αποθέματος

Η θεωρία ελέγχου αποθεμάτων ασχολείται με συστήματα εισόδου - εξόδου που παριστάνουν το επίπεδο ενός αγαθού σε μια αποθήκη ή παραγωγική μονάδα. Σκοπός της είναι ο βέλτιστος σχεδιασμός της παραγωγής (ή παραγγελίας) και της αποθήκευσης ενός αγαθού ώστε να ελαχιστοποιείται το κόστος της διαχείρισης του συστήματος.

Τα συστήματα ελέγχου αποθεμάτων αφθονούν στις διαδικασίες παραγωγής και παρουσιάζουν εξαιρετική ποικιλομορφία λόγω της διαφορετικής φύσης των διαφόρων αγαθών και των ειδικών συνθηκών που επικρατούν στην αγορά τους.

Μια πρώτη διάκριση των μοντέλων γίνεται με βάση τη φύση της ζήτησης (προσδιοριστική ή στοχαστική) και την επιθεώρηση του αποθέματος (συνεχή ή περιοδική). Συνδυάζοντας αυτά τα κριτήρια έχουμε 4 βασικές κατηγορίες μοντέλων. Φυσικά στη διαμόρφωση του μοντέλου υπεισέρχονται πολλοί άλλοι παράγοντες που επηρεάζουν τη μελέτη του καθοριστικά, όπως η ύπαρξη ή όχι πάγιων τελών παραγγελίας, η ανοχή ή όχι ελλειμάτων, η ύπαρξη εκπτώσεων για μεγάλες ποσότητες παραγγελίας, η δυνατότητα άμεσης παραλαβής παραγγελίας με επιπλέον κόστος κλπ.

Η θεωρία αποθεμάτων καλείται να απαντήσει σε ερωτήματα τύπου: Ποια πρέπει να είναι η ποσότητα αγαθού που θα παραγγείλω ή θα παράγω; Πότε πρέπει να δοθεί η παραγγελία ή να αρχίσει η παραγωγή; Πόσο πρέπει να είναι το μέγεθος της αποθήκης μου; Για την ουσιαστική μελέτη αυτών των συστημάτων τα βασικά εργαλεία είναι η Θεωρία Πιθανοτήτων και Στοχαστικών Ανελιξεων, αφού τα περισσότερα ρεαλιστικά μαθηματικά πρότυπα ελέγχου αποθεμάτων είναι στοχαστικά (τυχαιοκρατικά). Επιπλέον απαιτείται γνώση της κλασικής θεωρίας βελτιστοποίησης (μηδενισμός μερικών παραγώγων, μέθοδος πολλαπλασιαστών Lagrange κλπ.) καθώς και του δυναμικού προγραμματισμού.

1.3 Δομή της διατριβής

Στην παρούσα διατριβή παρουσιάζονται προβλήματα σε δύο βασικές περιοχές: (α) Προσαρμοστική δειγματοληψία υπό ελλιπή πληροφόρηση και επιπλέον περιορισμούς και (β) Βέλτιστες πολιτικές παραγγελιών σε σύστημα αποθεμάτων με πεπερασμένη χωρητικότητα και μερική υποκατάσταση ζήτησης. Η λογική συσχέτιση ανάμεσα στα προβλήματα που μελετήθηκαν στις δύο περιοχές περιγράφεται παρακάτω.

Στην πρώτη κατηγορία προβλημάτων μελετήθηκαν υποδείγματα προσαρμοστικής εκτίμησης και ελέγχου σε στοχαστικούς πληθυσμούς με μερικώς γνωστές κατανομές, στα οποία ο σκοπός είναι να αναπτυχθούν προσαρμοστικές πολιτικές λήψης διαδοχικών δειγμάτων έτσι ώστε να μεγιστοποιείται το αναμενόμενο μέσο αποτέλεσμα ανά βήμα κάτω από ένα εξωγενή περιορισμό ως προς το κόστος δειγματοληψίας ανά περίοδο. Τα προβλήματα που μελετήθηκαν σε αυτή την περιοχή είναι γενικά και περιλαμβάνουν εφαρμογές εκτός των άλλων και σε διαχείριση αποθεμάτων με άγνωστη κατανομή ζήτησης, στα οποία ζητείται μια προσαρμοστική πολιτική παραγγελιών που μεγιστοποιεί το αναμενόμενο κέρδος ανά μονάδα χρόνου. Επικεντρώθηκα σε ένα συγκεκριμένο πρόβλημα

αυτής της περιοχής, στο οποίο γίνονται παραγγελίες για δύο προϊόντα ταυτόχρονα σε περιβάλλον εφημεριδοπώλη (τα προϊόντα διατηρούνται για μια περίοδο μόνο) και μερική αμφίδρομη υποκατάσταση ζήτησης. Για το πρόβλημα αυτό αποδείξαμε ιδιότητες της βέλτιστης πολιτικής για την περίπτωση που οι παράμετροι του προβλήματος (κατανομές ζήτησης και πιθανότητες υποκατάστασης) είναι πλήρως γνωστές. Τα αποτελέσματα αυτά παρέχουν το απαραίτητο υπόβαθρο για την εφαρμογή και των προσαρμοστικών μεθόδων δειγματοληψίας, όταν κάποιες ή όλες από τις παραπάνω παραμέτρους δεν είναι γνωστές.

Στη συνέχεια παρουσιάζονται τα βασικά αποτελέσματα της διατριβής.

Το πρόβλημα προσαρμοστικής βελτιστοποίησης που μελετήθηκε είναι το παρακάτω: Θεωρούνται k στατιστικοί πληθυσμοί τέτοιοι ώστε οι διαδοχικές παρατηρήσεις από τον πληθυσμό a , $a = 1, \dots, k$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με αθροιστική συνάρτηση πιθανότητας $F(x|\theta_a)$ και μέση τιμή $\mu_a(\theta_a)$, όπου θ_a είναι ένα άγνωστο διάνυσμα παραμέτρων που παίρνει τιμές σε ένα γνωστό σύνολο Θ_a . Σε κάθε βήμα ο αποφασίζων πρέπει να επιλέξει ένα πληθυσμό και να πάρει μια παρατήρηση από αυτόν. Η δειγματοληψία από τον πληθυσμό a επιφέρει ένα γνωστό κόστος C_a ανά παρατήρηση. Το πρόβλημα είναι να βρεθεί μια προσαρμοστική πολιτική δειγματοληψίας τέτοια ώστε το μέσο κόστος δειγματοληψίας ανά περίοδο δεν υπερβαίνει ένα δοθέν άνω φράγμα C_0 και ταυτόχρονα η αναμενόμενη τιμή της μέσης αξίας των παρατηρήσεων ανά βήμα σε άπειρο ορίζοντα μεγιστοποιείται σύμφωνα με κάποιο κατάλληλο κριτήριο που λαβαίνει υπόψη την ελλιπή πληροφόρηση.

Σε αυτό το γενικό πρόβλημα μελετήθηκαν δύο περιπτώσεις.

Στο Κεφάλαιο 2 της διατριβής παρουσιάζεται η περίπτωση όπου ο περιορισμός κόστους πρέπει να ικανοποιείται κατά αναμενόμενη τιμή, δηλαδή το αναμενόμενο μέσο κόστος ανά περίοδο δε μπορεί να υπερβαίνει το C_0 . Για το πρόβλημα αυτό προτάθηκε μια κατηγορία πολιτικών δειγματοληψίας η οποία αποδείχθηκε ότι είναι ισχυρά συνεπής, δηλαδή το αναμενόμενο μέσο κέρδος ανά μονάδα χρόνου συγκλίνει σχεδόν βέβαια στην αντίστοιχη ποσότητα κάτω από πλήρη πληροφόρηση.

Στο Κεφάλαιο 3 παρουσιάζεται η σημαντικά δυσκολότερη μορφή του προβλήματος στην οποία αφενός ο περιορισμός του κόστους δειγματοληψίας πρέπει να ικανοποιείται για κάθε υλοποίηση της διαδικασίας και όχι απλώς κατά αναμενόμενη τιμή, και αφετέρου η πολιτική είναι εκτός από ισχυρά συνεπής και αποτελεσματική, δηλαδή ο ρυθμός

σύγκλισης του μέσου κέρδους ανά περίοδο σε αυτό κάτω από πλήρη πληροφόρηση μεγιστοποιείται σύμφωνα με ένα κατάλληλα ορισμένο ασυμπτωτικό κριτήριο απώλειας. Για το πρόβλημα αυτό αποδείχθηκαν δύο βασικά αποτελέσματα. Πρώτον αναπτύχθηκε ένα κριτήριο για την αποδοτικότητα μιας πολιτικής, δηλαδή αποδείχθηκε ότι για κάθε καλή πολιτική η απώλεια λόγω έλλειψης πληροφόρησης δε μπορεί να είναι χαμηλότερη από ένα ασυμπτωτικό κάτω φράγμα. Δεύτερον προτάθηκε μια πολιτική δειγματοληψίας για την οποία αποδείχθηκε ότι η απώλεια πιάνει το κάτω φράγμα, και επομένως η πολιτική αυτή είναι αποδοτική.

Στο Κεφάλαιο 4 της διατριβής παρουσιάζεται ένα πλαίσιο εφαρμογής των παραπάνω αποτελεσμάτων στη διαχείριση αποθεμάτων με δύο προϊόντα. Συγκεκριμένα παρουσιάζεται ένα πρόβλημα αποθεμάτων στο οποίο ο διαχειριστής στην αρχή κάθε περιόδου κάνει ταυτόχρονα παραγγελίες για δύο προϊόντα. Τα προϊόντα έχουν περιορισμένη διάρκεια ζωής και δε μπορούν να χρησιμοποιηθούν για περισσότερους από ένα κύκλους παραγγελιών, ενώ υπάρχει περιορισμός στη συνολική ποσότητα παραγγελίας λόγω πεπερασμένης χωρητικότητας της αποθήκης. Τέλος υπάρχει αμφίδρομη μερική υποκατάσταση ζήτησης, δηλαδή όταν ένας πελάτης δε βρίσκει το προϊόν που επιθυμεί, με δοσμένη πιθανότητα το υποκαθιστά με το άλλο προϊόν, αν αυτό είναι διαθέσιμο. Το πρόβλημα υποκατάστασης ζήτησης έχει μελετηθεί αρκετά στη βιβλιογραφία της διαχείρισης αποθεμάτων, αλλά λόγω της πολυπλοκότητας της ανάλυσης συνήθως γίνονται σημαντικές απλοποιητικές παραδοχές όσον αφορά τα ποσοστά ζήτησης που ικανοποιούνται μέσω υποκατάστασης. Στο πρόβλημα που παρουσιάζεται εδώ έγινε η πιο ρεαλιστική υπόθεση ότι η ζήτηση για τα δύο προϊόντα ακολουθεί δύο ανεξάρτητες διαδικασίες Poisson. Με τον τρόπο αυτό ενσωματώνεται φυσιολογικά στην ανάλυση το σημαντικό χαρακτηριστικό του προβλήματος ότι οι ελλείψεις που δημιουργούνται εξαρτώνται όχι μόνο από τη συνολική ζήτηση των δύο προϊόντων αλλά και από τις συγκεκριμένες χρονικές στιγμές που έρχεται η ζήτηση. Στο πρόβλημα αυτό αναπτύχθηκαν μοντέλα Μαρκοβιανών διαδικασιών συνεχούς χρόνου και μέσω ανάλυσης μεταβατικής ή οριακής κατανομής αναπτύχθηκαν κλειστές εκφράσεις για το κέρδος ανά μονάδα χρόνου ως συνάρτηση των ποσοτήτων παραγγελίας. Επίσης αποδείχθηκε ότι η συνάρτηση κέρδους είναι sub-modular ως προς τις ποσότητες, γεγονός που επιτρέπει σημαντική απλοποίηση στον προσδιορισμό της βέλτιστης πολιτικής.

Κεφάλαιο 2

Προσαρμοστικές πολιτικές για ακολουθιακή δειγματοληψία υπό ελλιπή πληροφόρηση και ασυμπτωτικό περιορισμό στο κόστος

2.1 Εισαγωγή

Σε αυτό το Κεφάλαιο θεωρούμε το πρόβλημα ακολουθιακής δειγματοληψίας από k ανεξάρτητους στατιστικούς πληθυσμούς με άγνωστες κατανομές. Το ζητούμενο είναι να μεγιστοποιήσουμε την αναμενόμενη επιστροφή ανα περίοδο που επιτυγχάνεται σε άπειρο ορίζοντα, κάτω από τον περιορισμό το αναμενόμενο κόστος δειγματοληψίας ανα περίοδο να μην ξεπερνά ένα άνω όριο. Η εισαγωγή του κόστους δειγματοληψίας εισάγει μια νέα διάσταση στα καθιερωμένα προβλήματα ανάμεσα στον πειραματισμό και την μεγιστοποίηση που αντιμετωπίζονταν στα προβλήματα ελέγχου κάτω από ελλιπή πληροφόρηση. Το κόστος δειγματοληψίας μπορεί να απαγορεύει την χρησιμοποίηση πληθυσμών με μεγάλες μέσες επιστροφές επειδή το δειγματοληπτικό τους κόστος μπορεί να είναι πολύ

υψηλό. Για αυτό, ο αποφασίζων πρέπει να ταυτοποιεί το υποσύνολο των πληθυσμών με τον καλύτερο συνδυασμό επιστροφών σε σχέση με το κόστος και να κατανέμει την δειγματοληπτική του προσπάθεια με βέλτιστο τρόπο.

Από μαθηματικής πλευράς, αυτή η κλάση προβλημάτων ενσωματώνει στατιστικές μεθοδολογίες σε προβλήματα μαθηματικού προγραμματισμού. Για αυτό, κάτω από πλήρη πληροφόρηση, το πρόβλημα κατανομής της προσπάθειας κάτω από τους περιορισμούς στο κόστος μοντελοποιείται σαν ένα πρόβλημα γραμμικού ή μη γραμμικού προγραμματισμού. Ωστόσο όταν κάποιες από τις παραμέτρους του προβλήματος είναι άγνωστες εκ των προτέρων και πρέπει να εκτιμηθούν μέσω πειραματισμού, ο αποφασίζων πρέπει να σχεδιάσει προσαρμοστική γνώση και πολιτικές ελέγχου οι οποίες εξασφαλίζουν γνώση για τις παραμέτρους και την ίδια στιγμή εξασφαλίζουν ότι το όφελος που θυσιάζεται για γνώση είναι όσο το δυνατό μικρότερο.

Το μοντέλο σε αυτό το Κεφάλαιο ανήκει στην γενική περιοχή των multi-armed bandit προβλημάτων, τα οποία ορίστηκαν από τον [8], ο οποίος πρότεινε μια απλή προσαρμοστική πολιτική για ακολουθιακή δειγματοληψία από δύο άγνωστους πληθυσμούς με σκοπό να μεγιστοποιήσει την αναμενόμενη επιστροφή ανα μονάδα χρόνου σε άπειρο ορίζοντα. Στο [4] γενίκευσαν τα αποτελέσματα κατασκευάζοντας ασυμπτωτικά αποδοτικές προσαρμοστικές πολιτικές με τον βέλτιστο ρυθμό σύγκλισης της μέσης επιστροφής στην βέλτιστη τιμή κάτω από πλήρη πληροφόρηση και δείχνουν ότι η πεπερασμένου ορίζοντα απώλεια λόγω της ελλιπούς πληροφόρησης αυξάνει με λογαριθμικό ρυθμό. Στο [2] απέδειξαν ότι υπάρχουν απλούστερες αποδοτικές πολιτικές που βασίζονται σε δείκτες για την περίπτωση της κανονικής κατανομής με άγνωστες μέσες τιμές, και στο [14] επέκτειναν τα αποτελέσματα των αποδοτικών πολιτικών στην απαραμετρική περίπτωση διακριτών κατανομών με γνωστό στήριγμα.

Σε μια πεπερασμένη εκδοχή στο [3] ανέπτυξαν μια minimax εκδοχή των αποτελεσμάτων του [4] για δύο πληθυσμούς, ενώ στο [1] κατασκευάζουν πολιτικές οι οποίες επίσης επιτυγχάνουν την λογαριθμική απώλεια ομοιόμορφα στο χρόνο, αντί μόνο ασυμπτωτικά.

Σε όλες τις εργασίες που αναφέρθηκαν παραπάνω δεν υπάρχουν περιορισμοί στην δειγματοληψία. Προβλήματα με προσαρμοστικές πολιτικές και περιορισμούς δεν αναφέρονται πολλά στην βιβλιογραφία λόγω της δυσκολίας που παρουσιάζουν. Στο [9] θεωρείται ένα multi-armed bandit μοντέλο με περιορισμούς αλλά υιοθετείται Μπεϋζιανό υπόβαθρο και το πρόβλημα προσεγγίζεται μέσω των δεικτών του Gittins. Το άρθρο

προτείνει διάφορες ευριστικές πολιτικές (heuristic). Στο [6] επίσης θεωρούν ένα πρόβλημα εκτίμησης της κατανομής ενός μόνο πληθυσμού με δειγματοληπτικό κόστος κάτω από την υπόθεση ότι ο αριθμός των χρηστών που θα οφελούν εξαρτάται από την εκτίμηση της επιστροφής. Τέλος, στο [5] αναλύεται μια εκδοχή του multi-armed bandit προβλήματος με Μπευζιανή προσέγγιση, Bernoulli επιστροφές και Βήτα εκ των προτέρων κατανομή των άγνωστων παραμέτρων, όπου υπάρχει ένα συνολικό διαθέσιμο ποσό για πειραματισμό, το οποίο πρέπει να κατανεμηθεί για να κάνουν δειγματοληψία από τους διαφορετικούς πληθυσμούς.

Μια άλλη προσέγγιση η οποία είναι πιο κοντά σε αυτή που υιοθετούμε εδώ είναι να θεωρήσουμε την οικογένεια των στοχαστικών προσεγγίσεων και των αλγορίθμων ενισχυτικής γνώσης. Η γενική ιδέα είναι ότι η επιλογή σε κάθε βήμα δίνεται ακολουθώντας μια πολιτική τυχαιοποίησης με πιθανότητες τυχαιοποίησης οι οποίες προσαρμόζονται μετά την παρατήρηση της επιστροφής σε κάθε περίοδο. Το προσαρμοστικό σχέδιο βασίζεται σε ένα προσεγγιστικό αλγόριθμο. Αλγόριθμοι τέτοιου τύπου αναλύονται στο [7] για την πιο γενική περίπτωση όπου οι επιστροφές των πληθυσμών έχουν Μαρκοβιανή δυναμική ατί να είναι ανεξάρτητες και ισόνομες.

Σε αυτό το Κεφάλαιο θα δούμε την κατασκευή μιας οικογένειας πολιτικών για τις οποίες η μέση επιστροφή ανα περίοδο συγκλίνει στην βέλτιστη τιμή κάτω από πλήρη πληροφόρηση για όλες τις κατανομές των ανεξάρτητων πληθυσμών με πεπερασμένες μέσες τιμές. Με αυτή την έννοια γενικεύει τα αποτελέσματα του [8] εισάγοντας το κόστος δειγματοληψίας. Το Κεφάλαιο έχει την εξής δομή. Στην δεύτερη ενότητα, περιγράφουμε το μοντέλο στο πλαίσιο της πλήρους και ελλιπούς πληροφόρησης. Στην τρίτη Ενότητα, κατασκευάζουμε μια κλάση προσαρμοστικών πολιτικών δειγματοληψίας και αποδεικνύουμε ότι είναι συνεπής. Στην τέταρτη Ενότητα, μελετάμε την ταχύτητα σύγκλισης των πολιτικών που προτείνονται μέσω προσομοίωσης. Τέλος στην πέμπτη Ενότητα συνοψίζουμε.

2.2 Περιγραφή μοντέλου

Θεωρούμε το ακόλουθο πρόβλημα προσαρμοστικής δειγματοληψίας. Υπάρχουν k ανεξάρτητοι στατιστικοί πληθυσμοί, $i = 1, \dots, k$. Οι διαδοχικές παρατηρήσεις από έναν πληθυσμό i αποτελούν μια ακολουθία από ανεξάρτητες και ισόνομες τυχαίες μεταβλητές

X_{i1}, X_{i2}, \dots που ακολουθούν μια κατανομή μιας μεταβλητής με πυκνότητα $f_i(\cdot)$ ως προς ένα μη εκφυλισμένο μέτρο ν . Έτσι το στοχαστικό μοντέλο καθορίζεται μοναδικά από το διάνυσμα $f = (f_1, \dots, f_k)$ των ανεξάρτητων συναρτήσεων πυκνότητας πιθανότητας. Δεδομένης της f έστω $\underline{\mu}(f)$ το διάνυσμα των μέσων τιμών, π.χ. $\mu_i(f) = E^{f_i}(X_i)$. Ο τύπος της f δεν είναι γνωστός. Σε κάθε περίοδο ο πειραματιστής πρέπει να επιλέξει έναν πληθυσμό ώστε να πάρει ένα και μοναδικό δείγμα. Το να κάνουμε δειγματοληψία από τον πληθυσμό i επιφέρει ένα κόστος c_i ανά δειγματοληψία και χωρίς βλάβη της γενικότητας υποθέτουμε ότι $c_1 \leq c_2 \leq \dots \leq c_k$, αλλά όχι όλα ίσα. Το ζητούμενο είναι να μεγιστοποιήσουμε την μέση αναμενόμενη επιστροφή ανά περίοδο υπό τον περιορισμό το μέσο αναμενόμενο δειγματοληπτικό κόστος ανά περίοδο σε άπειρο ορίζοντα να μην ξεπερνά ένα δεδομένο άνω όριο C_0 . Χωρίς βλάβη της γενικότητας υποθέτουμε ότι $c_1 \leq C_0 < c_k$. Πράγματι αν $C_0 < c_1$ τότε το πρόβλημα είναι μη εφικτό. Από την άλλη πλευρά αν $C_0 \geq c_k$ ο περιορισμός στο κόστος είναι περιττός. Έστω $d = \max\{j : c_j \leq C_0\}$. Τότε $1 \leq d < k$ και $c_d \leq C_0 < c_{d+1}$.

2.2.1 Πλαίσιο πλήρους πληροφόρησης

Αρχικά θα αναλύσουμε το πρόβλημα πλήρους πληροφόρησης. Αν όλες οι $f_i(\cdot)$ είναι γνωστές, τότε το πρόβλημα μπορεί να μοντελοποιηθεί μέσω γραμμικού προγραμματισμού. Θεωρούμε μια τυχαιοποιημένη πολιτική δειγματοληψίας η οποία σε κάθε περίοδο επιλέγει τον πληθυσμό j με πιθανότητα x_j , για $j = 1, \dots, k$. Για να βρούμε μια πολιτική η οποία μεγιστοποιεί την αναμενόμενη επιστροφή, μπορούμε να διατυπώσουμε το ακόλουθο γραμμικό πρόγραμμα στην τυπική του μορφή:

$$\begin{aligned}
 z^* &= \max \sum_{j=1}^k \mu_j x_j \\
 &\sum_{j=1}^k c_j x_j + y = C_0 \\
 &\sum_{j=1}^k x_j = 1 \\
 &x_j \geq 0, \forall j.
 \end{aligned} \tag{2.1}$$

Σημειώνουμε ότι το z^* εξαρτάται από την f μόνο μέσω του διανύσματος $\underline{\mu}(f)$, π.χ. το z^* είναι το ίδιο για όλες τις περιπτώσεις συναρτήσεων πυκνότητας πιθανότητας με την ίδια μέση τιμή μ . Οπότε από εδώ και στο εξής θα υποδηλώνουμε το z^* σαν μια συνάρτηση του άγνωστου διανύσματος των μέσων τιμών $\underline{\mu}$.

Στην ανάλυση του προβλήματος χρειαζόμαστε επίσης το δυικό πρόβλημα γραμμικού προγραμματισμού ($\Delta\Pi$) του (2.1),

$$\begin{aligned} z_D^* &= \min g + C_0\lambda \\ g + c_1\lambda &\geq \mu_1 \\ &\vdots \\ g + c_k\lambda &\geq \mu_k \\ g \in \mathbb{R}, \lambda &\geq 0, \end{aligned}$$

με δύο μεταβλητές τις λ και g οι οποίες αντιστοιχούν στον πρώτο και στον δεύτερο περιορισμό του (2.1), αντιστοίχως.

Ο Βασικός πίνακας B αντιστοιχεί σε μια Βασική Εφικτή Λύση (ΒΕΛ) του προβλήματος (2.1) και μπορεί να πάρει μια από τις δύο μορφές:

Στην πρώτη περίπτωση, οι βασικές μεταβλητές είναι οι x_i, x_j , για δύο πληθυσμούς i, j , με $c_i \leq C_0 \leq c_j$, $c_i < c_j$, και ο βασικός πίνακας είναι

$$\mathbf{B} = \begin{pmatrix} c_i & c_j \\ 1 & 1 \end{pmatrix}.$$

Η ΒΕΛ είναι τότε

$$x_i = \frac{c_j - C_0}{c_i - c_j}, \quad x_j = \frac{C_0 - c_i}{c_i - c_j}, \quad \text{και } x_m = 0 \text{ για } m \neq i, j, \quad y = 0,$$

με

$$z(\underline{x}) = \mu_i x_i + \mu_j x_j.$$

Η λύση είναι μη εκφυλισμένη όταν $c_i < C_0 < c_j$ και εκφυλισμένη όταν $C_0 = c_i$ ή $C_0 = c_j$. Στην τελευταία περίπτωση, αυτό αντιστοιχεί στο να κάνουμε δειγματοληψία από έναν και μόνο πληθυσμό τον $l = i$ ή $l = j$, αντίστοιχα:

$$x_l = 1, \quad x_m = 0 \quad \forall m \neq l, \quad y = 0,$$

με

$$z(\underline{x}) = \mu_i.$$

Η δεύτερη περίπτωση μιας ΒΕΛ αντιστοιχεί στις βασικές μεταβλητές x_i, y για έναν πληθυσμό i με $c_i \leq C_0$. Ο βασικός πίνακας είναι

$$\mathbf{B} = \begin{pmatrix} c_i & 1 \\ 1 & 0 \end{pmatrix}.$$

Σε αυτή την περίπτωση η ΒΕΛ αντιστοιχεί σε δειγματοληψία από τον πληθυσμό i και μόνο

$$x_i = 1, \quad x_m = 0 \quad \forall m \neq i, \quad y = C_0 - c_i,$$

με

$$z(\underline{x}) = \mu_i.$$

Η λύση είναι μη εκφυλισμένη αν $c_i < C_0$, διαφορετικά είναι εκφυλισμένη.

Από τα παραπάνω ακολουθεί ότι μια ΒΕΛ είναι εκφυλισμένη αν $x_l = 1$ για κάποιο l με $c_l = C_0$. Οποιοσδήποτε βασικός πίνακας B περιέχει την x_l σαν βασική μεταβλητή αντιστοιχεί σε αυτή την ΒΕΛ.

Για μια ΒΕΛ x έστω

$$b = \{i : x_i > 0\}.$$

Τότε, είτε $b = \{i, j\}$ για κάποια i, j με $i \leq d \leq j$, ή $b = \{i\}$ για κάποιο $i \leq d$. Υπάρχει ένα προς ένα αντιστοιχία ανάμεσα στις βασικές εφικτές λύσεις και τα σύνολα των b αυτής της μορφής. Χρησιμοποιούμε το K για να δηλώσουμε το σύνολο των ΒΕΛ, ή ισοδύναμα

$$K = \{b : b = \{i, j\}, i \leq d \leq j \text{ ή } b = \{i\}, i \leq d\}.$$

Αφού η εφικτή περιοχή του (2.1) είναι φραγμένη, τότε το K είναι πεπερασμένο.

Για έναν βασικό πίνακα B , έστω $v^B = (\lambda^B, g^B)$ το δυικό διάνυσμα που αντιστοιχεί στον B , π.χ., $v^B = \mu_B B^{-1}$, όπου $\mu_B = (\mu_i, \mu_j)$, ή $\mu_B = (\mu_i, 0)$, ανάλογα με την μορφή του B .

Όσον αφορά την βελτιστότητα, μια ΒΕΛ είναι βέλτιστη αν και μόνο αν για τουλάχιστον έναν βασικό πίνακα B τα ελαττωμένα κόστη (reduced costs/dual slacks) είναι όλα μη αρνητικά:

$$\phi_\alpha^B \equiv c_\alpha \lambda^B + g^B - \mu_\alpha \geq 0, \quad \alpha = 1, \dots, k.$$

Ένας βασικός πίνακας B που ικανοποιεί αυτή τη συνθήκη είναι βέλτιστος. Να σημειώσουμε ότι αν μια βέλτιστη ΒΕΛ είναι εκφυλισμένη, τότε δεν είναι απαραίτητο όλοι οι βασικοί πίνακες που αντιστοιχούν σε αυτή να είναι βέλτιστοι.

Είναι εύκολο να δείξουμε ότι τα ελαττωμένα κόστη (reduced costs) μπορούν να εκφραστούν σαν γραμμικοί συνδυασμοί $\phi_\alpha^B = \underline{w}_\alpha^B \underline{\mu}$, όπου \underline{w}_α^B είναι ένα κατάλληλα ορισμένο διάνυσμα το οποίο είναι ανεξάρτητο του $\underline{\mu}$.

Τέλος ορίζουμε το σύνολο των βέλτιστων λύσεων του (2.1) για ένα $\underline{\mu}$,

$$s(\underline{\mu}) = \{b \in K : b \text{ αντιστοιχεί σε μια ΒΕΛ}\}.$$

Μια βέλτιστη λύση του (2.1) προσδιορίζει τις πιθανότητες τυχαιοποίησης οι οποίες διασφαλίζουν την μεγιστοποίηση της μέσης επιστροφής που υπόκεινται στον περιορισμό στο κόστος. Παρατηρούμε ότι ένας εναλλακτικός τρόπος εφαρμογής της βέλτιστης λύσης, χωρίς τυχαιοποίηση, είναι να επιλέγουμε περιοδικά από όλους τους πληθυσμούς με τέτοιο τρόπο ώστε το ποσοστό των δειγματοληψιών από κάθε πληθυσμό j να είναι ίσο με x_j . Επειδή τα ποσοστά δειγματοληψιών είναι ρητοί αριθμοί, η παραπάνω ισοδυναμία ισχύει όταν και οι πιθανότητες τυχαιοποίησης που προκύπτουν από τη λύση του π.γ.π. είναι επίσης ρητοί. Αυτή η απαίτηση δεν είναι περιοριστική γιατί αρκεί να υποθέσουμε ό,τι τα κόστη και οι μέσες τιμές είναι ρητοί αριθμοί (γνωρίζουμε ότι η βέλτιστη λύση ενός π.γ.π. με ρητούς συντελεστές είναι ρητή).

2.2.2 Πλαίσιο ελλιπούς πληροφόρησης

Σε αυτή την Ενότητα υποθέτουμε ότι οι κατανομές των πληθυσμών είναι άγνωστες. Ειδικότερα κάνουμε την παρακάτω υπόθεση.

Υπόθεση 2.2.1. *Οι κατανομές των επιστροφών είναι ανεξάρτητες, και οι αναμενόμενες τιμές $\mu_\alpha = E(X_\alpha) < \infty$, $\alpha = 1, \dots, k$.*

Έστω F το σύνολο όλων των $f = (f_1, \dots, f_k)$ που ικανοποιούν την Υπόθεση 2.2.1. Η κλάση F είναι το σύνολο των αποτελεσματικών παραμέτρων στο πλαίσιο ελλιπούς πληροφόρησης. Κάτω από ελλιπή πληροφόρηση, μια πολιτική όπως αυτή της Ενότητας 2.2.1, η οποία εξαρτάται από τις πραγματικές τιμές των $\underline{\mu}$, δεν είναι πλέον αποδεκτή. Για αυτό περιορίζουμε την προσοχή μας στην κλάση των προσαρμοστικών πολιτικών,

οι οποίες εξαρτώνται μόνο από τις προηγούμενες παρατηρήσεις των επιλογών και των επιστροφών.

Ειδικότερα, έστω ότι με $A_t, X_t, t = 1, 2, \dots$ δηλώνουμε τον επιλεγμένο πληθυσμό και την παρατηρούμενη επιστροφή στην περίοδο t . Έστω τώρα $h_t = (\alpha_1, x_1, \dots, \alpha_{t-1}, x_{t-1})$ η ιστορία των δράσεων και των παρατηρήσεων που είναι διαθέσιμες την περίοδο t .

Μια προσαρμοστική πολιτική είναι ορισμένη σαν μια κολουθία $\pi = (\pi_1, \pi_2, \dots)$ από ιστορικά ανεξάρτητες κατανομές πιθανοτήτων στο $\{1, \dots, k\}$, έτσι ώστε

$$\pi_t(j, h_t) = P(A_t = j | h_t).$$

Δεδομένης της ιστορίας h_n , έστω ότι $T_n(\alpha)$ δηλώνει τον αριθμό των φορών που ο πληθυσμός α έχει επιλεγεί κατά την διάρκεια των πρώτων n περιόδων

$$T_n(\alpha) = \sum_{t=1}^n 1\{A_t = \alpha\}.$$

Έστω S_n^π να είναι η επιστροφή μέχρι την περίοδο n :

$$S_n^\pi = \sum_{t=1}^n X_t,$$

και C_n^π το συνολικό κόστος μέχρι την περίοδο n :

$$C_n^\pi = \sum_{t=1}^n c_{A_t}.$$

Αυτές οι ποσότητες μπορούν να χρησιμοποιηθούν για να ορίσουν τις επιθυμητές ιδιότητες μιας προσαρμοστικής πολιτικής, δηλαδή της εφικτότητας και της συνέπειας.

Ορισμός 2.2.1. Μια πολιτική π καλείται εφικτή (*feasible*) αν

$$\limsup_{n \rightarrow \infty} \frac{E^\pi(C_n^\pi)}{n} \leq C_0, \quad \forall f \in F. \quad (2.2)$$

Ορισμός 2.2.2. Μια πολιτική π καλείται συνεπής (*consistent*) αν είναι εφικτή και

$$\lim_{n \rightarrow \infty} \frac{S_n^\pi}{n} = z^*(\underline{\mu}), \quad \text{σχ.β.} \quad \forall f \in F.$$

Έστω Π^F και Π^C να δηλώνουν τις κλάσεις των εφικτών και συνεπών πολιτικών, αντίστοιχα. Οι παραπάνω ιδιότητες είναι εύλογες απαιτήσεις για μια προσαρμοστική πολιτική. Η πρώτη εξασφαλίζει ότι το μακροπρόθεσμο μέσο κόστος δειγματοληψίας δεν ξεπερνά το διαθέσιμο ποσό. Ο δεύτερος ορισμός σημαίνει ότι η μακροπρόθεσμη μέση αμοιβή ανά περίοδο η οποία επιτυγχάνεται από την π συγκλίνει με πιθανότητα ένα στην βέλτιστη αναμενόμενη τιμή η οποία θα μπορούσε να επιτευχθεί κάτω από πλήρη πληροφόρηση, για όλες τις πιθανές κατανομές πληθυσμών οι οποίες ικανοποιούν την Υπόθεση 2.2.1.

Παρατηρούμε εδώ ότι η συνέπεια όπως ορίζεται στον Ορισμό 2.2.2 είναι ισοδύναμη με την έννοια της ισχυρής συνέπειας ενός εκτιμητή.

2.3 Κατασκευή μιας συνεπούς πολιτικής

Η ερώτηση κλειδί στο πλαίσιο ελλιπούς πληροφόρησης είναι αν υπάρχουν εφικτές και, ακόμα πιο σημαντικό, συνεπείς πολιτικές και πως μπορούν να κατασκευασθούν.

Είναι πολύ εύκολο να δείξουμε ότι υπάρχουν εφικτές πολιτικές, αφού τα κόστη δειγματοληψίας είναι γνωστά. Για αυτό κάθε τυχαίοποιημένη πολιτική, όπως ορίστηκε στην Ενότητα 2.2.1, με πιθανότητες τυχαίοποίησης οι οποίες ικανοποιούν το Π.Γ.Π. (2.1) είναι εφικτή για κάθε κατανομή f . Έτσι, $\Pi^F \neq \emptyset$.

Από την άλλη πλευρά, η κατασκευή συνεπών πολιτικών δεν είναι τόσο απλή. Μια συνεπής πολιτική πρέπει να επιτυγχάνει τρεις στόχους. Πρώτον να είναι εφικτή. Δεύτερον να είναι ικανή να εκτιμήσει τις μέσες επιστροφές από όλους τους πληθυσμούς. Για να επιτευχθεί ο δεύτερος στόχος χρειάζεται να πάρουμε άπειρες φορές κάθε πληθυσμό, πράγμα που μας οδηγεί στο να πάρουμε άπειρες φορές και μη βέλτιστους πληθυσμούς που σημαίνει ό,τι θα επηρεαστεί το μέσο όφελος άρα ο τρίτος στόχος είναι, μακροπρόθεσμα, η δειγματοληψία που πραγματοποιεί από μη βέλτιστους πληθυσμούς να είναι αρκετά σπάνια δηλαδή το ποσοστό επιλογής τους να τείνει στο μηδέν ώστε να μην επηρεάζει το μέσο όφελος.

Σε αυτή την Ενότητα αποδεικνύουμε την ύπαρξη μιας κλάσης συνεπών πολιτικών. Η κατασκευή ακολουθεί την βασική ιδέα του [8], και βασίζεται σε αραιές ακολουθίες από θετικούς ακεραίους, οι όροι των οποίων αυξάνουν την μεταξύ τους απόσταση όσο αυξάνεται ο χρονικός ορίζοντας ώστε το ποσοστό των όρων τους να τείνει στο μηδέν,

οι οποίες είναι προσαρμοσμένες ώστε να εξασφαλίζουν και την εφικτότητα.

Αρχίζουμε με μερικούς ορισμούς. Για κάθε πληθυσμό j , έστω $\hat{\mu}_{j,t}$, $t = 1, 2, \dots$ να είναι ισχυρός συνεπής εκτιμητής της μ_j , π.χ. $\lim_{t \rightarrow \infty} \hat{\mu}_{j,t} = \mu_j$ σχ.β.- f_j . Τέτοιες εκτιμήτριες υπάρχουν. Για παράδειγμα από την Υπόθεση 2.2.1, ο δειγματικός μέσος $\bar{X}_{j,t} = \frac{1}{t} \sum_{k=1}^t X_{j,k}$ είναι ισχυρά συνεπής.

Για κάθε n , έστω $\hat{\underline{\mu}}_n = (\hat{\mu}_{j,T_j(n)}, j = 1, \dots, k)$ να είναι το διάνυσμα των εκτιμητών του $\underline{\mu}$ το οποίο βασίζεται στην ιστορία μέχρι την περίοδο n . Επίσης έστω ότι το $\hat{z}_n = z(\hat{\underline{\mu}}_n)$ δηλώνει την βέλτιστη τιμή του Π.Γ.Π. (2.1) όπου χρησιμοποιήθηκαν οι εκτιμητές αντί των άγνωστων μέσων τιμών στην αντικειμενική συνάρτηση. Το \hat{z}_n θα αναφέρεται σαν το Ισοδύναμο Π.Γ.Π. Παρατηρούμε ότι το $s(\hat{\underline{\mu}}_n)$ είναι το σύνολο των βέλτιστων ΒΕΛ του \hat{z}_n .

Η λύση του \hat{z}_n αντιστοιχεί σε μια πολιτική δειγματοληψίας η οποία καθορίζεται από ένα βέλτιστο διάνυσμα \hat{x}_n , έτσι ώστε $\hat{z}_n = \hat{\underline{\mu}}_n' \hat{x}_n$.

Τώρα θα ορίσουμε μια κλάση πολιτικών δειγματοληψίας, η οποία θα δείξουμε ότι είναι συνεπής. Θεωρούμε k μη επικαλυπτόμενες αραιές ακολουθίες από θετικούς ακεραίους,

$$\tau_j = \{\tau_{j,m}, m = 1, 2, \dots\}, j = 1, \dots, k,$$

έτσι ώστε

$$\lim_{m \rightarrow \infty} \frac{\tau_{j,m}}{m} = \infty, j = 1, \dots, k. \quad (2.3)$$

Τώρα ορίζουμε την πολιτική π^0 η οποία στην περίοδο n επιλέγει έναν πληθυσμό j με πιθανότητα ίση με

$$\pi^0(j|h_n) = \begin{cases} 1 & , \text{αν } \tau_{j,m} = n \text{ για κάποιο } m \geq 1 \\ \hat{x}_{n,j} & , \text{διαφορετικά} \end{cases}$$

όπου \hat{x}_n είναι μια βέλτιστη ΒΕΛ του Ισοδύναμου Π.Γ.Π. \hat{z}_n .

Η κύρια ιδέα στην π^0 είναι ότι στις περιόδους οι οποίες συμπίπτουν με όρους της ακολουθίας τ_j , ο πληθυσμός j επιλέγεται ανεξάρτητα με την ιστορία. Αυτές οι περιπτώσεις αναφέρονται ως εξαναγκαστικές επιλογές του πληθυσμού j . Ο σκοπός των εξαναγκαστικών επιλογών είναι να εξασφαλίσουν ότι θα κάνουμε δειγματοληψία από όλους τους πληθυσμούς αρκετά συχνά, έτσι ώστε το διάνυσμα των εκτιμητών $\hat{\underline{\mu}}_n$ να συγκλίνει στην πραγματική μέση τιμή $\underline{\mu}$ όταν $n \rightarrow \infty$.

Από την άλλη πλευρά, επειδή οι ακολουθίες τ_j είναι αραιές, το ποσοστό των περιόδων που συμβαίνουν εξαναγκαστικές επιλογές συγκλίνει στο μηδέν για όλα τα j , έτσι ώστε η δειγματοληψία από τους μη βέλτιστους πληθυσμούς να μην επηρεάζει την μέση επιστροφή μακροπρόθεσμα.

Στις υπόλοιπες χρονικές περιόδους, οι οποίες δεν συμπίπτουν με αυτές μιας αραιής ακολουθίας, η πολιτική δειγματοληψίας είναι αυτή που προτείνεται από το Ισοδύναμο Π.Γ.Π., π.χ., ο πειραματιστής γενικά επιλέγει με πιθανότητες τυχαιοποίησης από τους πληθυσμούς, οι οποίοι, βασιζόμενοι στην ιστορία, φαίνεται να είναι βέλτιστοι.

Στο επόμενο θεώρημα αποδεικνύουμε το κυρίως αποτέλεσμα της Ενότητας, δηλαδή ότι $\pi^0 \in \Pi^C$. Η απόδειξη προσαρμόζει την κύρια ιδέα του [8] στο πρόβλημα με περιορισμό στο κόστος.

Θεώρημα 2.3.1. *Η πολιτική π^0 είναι συνεπής.*

Πρίν δείξουμε το Θεώρημα 2.3.1, αποδεικνύουμε ένα ενδιάμεσο αποτέλεσμα το οποίο δείχνει ότι αν σε κάποια περίοδο το Ισοδύναμο Π.Γ.Π. δίδει μια βέλτιστη λύση η οποία είναι μη βέλτιστη κάτω από την πραγματική κατανομή f , τότε η εκτίμηση της μέσης τιμής ενός τουλάχιστον πληθυσμού πρέπει να είναι επαρκώς διαφορετική από την πραγματική της τιμή. Χρησιμοποιούμε την supremum νόρμα $\|x\| = \max_j |x_j|$.

Λήμμα 2.3.1. *Για κάθε $\underline{\mu}$ υπάρχει $\epsilon > 0$ έτσι ώστε για κάθε $n = 1, 2, \dots$ αν $b \in s(\hat{\underline{\mu}}_n)$ και $b \notin s(\underline{\mu})$ για κάποιο $b \in K$, τότε $\|\underline{\mu} - \hat{\underline{\mu}}_n\| \geq \epsilon$.*

Απόδειξη. Αφού $b \notin s(\underline{\mu})$, έχουμε ότι για κάθε βασικό πίνακα B' που αντιστοιχεί σε μια ΒΕΛ b υπάρχει τουλάχιστον ένας $m \in \{1, \dots, k\}$ έτσι ώστε $\phi_m^{B'}(\underline{\mu}) < 0$. Οπότε,

$$-\underline{w}_m^{B'} \underline{\mu} = -\phi_m^{B'}(\underline{\mu}) > 0. \quad (2.4)$$

Επιπρόσθετα, αφού $b \in s(\hat{\underline{\mu}}_n)$, υπάρχει ένας βασικός πίνακας B που αντιστοιχεί στο b , έτσι ώστε για κάθε $m \in \{1, \dots, k\}$ έχουμε ότι $\phi_m^B(\hat{\underline{\mu}}_n) \geq 0$, έτσι,

$$\underline{w}_m^B \hat{\underline{\mu}}_n = \phi_m^B(\hat{\underline{\mu}}_n) \geq 0. \quad (2.5)$$

Για αυτόν τον βασικό πίνακα B , ακολουθεί από το (2.4) και (2.5) ότι

$$\underline{w}_m^B \hat{\underline{\mu}}_n - \underline{w}_m^B \underline{\mu} \geq -\phi_m^B(\underline{\mu}) = |\phi_m^B(\underline{\mu})| > 0$$

$$\Rightarrow \underline{w}_m^B(\hat{\underline{\mu}}_n - \underline{\mu}) \geq |\phi_m^B(\underline{\mu})|$$

$$\Rightarrow k \|\underline{w}_m^B\| \|\hat{\underline{\mu}}_n - \underline{\mu}\| \geq |\phi_m^B(\underline{\mu})|$$

$$\Rightarrow \|\hat{\underline{\mu}}_n - \underline{\mu}\| \geq \frac{|\phi_m^B(\underline{\mu})|}{k \|\underline{w}_m^B\|},$$

επειδή από την ιδιότητα $\underline{w}_m^B \underline{\mu} < 0$ ακολουθεί ότι $\|\underline{w}_m^B\| > 0$.

Τώρα έστω

$$\epsilon = \min_{b \in K, b \notin s(\underline{\mu})} \min_{B \in b} \min_{m \in \{1, \dots, k\}} \left\{ \frac{|\phi_m^B(\underline{\mu})|}{k \|\underline{w}_m^B\|} : \phi_m^B(\underline{\mu}) < 0 \right\} > 0.$$

όπου η ελαχιστοποίηση ως προς $B \in b$ είναι πάνω σε όλους τους βασικούς πίνακες που αντιστοιχούν στη ΒΕΛ b .

Τότε $\|\hat{\underline{\mu}}_n - \underline{\mu}\| \geq \epsilon$.

□

Απόδειξη Θεωρήματος 2.3.1.

Για $i = 1, \dots, k$ έστω ότι το

$$SS_i(n) = \sum_{t=1}^n 1\{\tau_{i,m} = t, \text{ για κάποιο } m\},$$

δηλώνει τον αριθμό των περιόδων $\{1, \dots, n\}$ όπου εμφανίζεται μια εξαναγκαστική επιλογή από τον πληθυσμό i .

Επίσης έστω,

$$Y_j^b(n) = \sum_{t=1}^n 1\{b \in s(\hat{\underline{\mu}}_t), b \text{ χρησιμοποιήθηκε στην περίοδο } t, \text{ και επιλέξαμε τον } j, \text{ σύμφωνα με την τυχαιοποίηση } b\}.$$

$$Y^b(n) = \sum_{j \in b} Y_j^b(n),$$

$$Y(n) = \sum_{b \in s(\underline{\mu})} Y^b(n).$$

Αφού αυτά εμπεριέχουν όλες τις πιθανότητες επιλογής σε μια περίοδο, έχουμε ότι

$$n = \sum_{i=1}^k SS_i(n) + \sum_{b \notin s(\underline{\mu})} Y^b(n) + \sum_{b \in s(\underline{\mu})} Y^b(n).$$

Τώρα έστω ότι το W_n δηλώνει το άθροισμα των επιστροφών στις περιόδους όπου έχει χρησιμοποιηθεί η πραγματική βέλτιστη ΒΕΛ:

$$W_n = \sum_{b \in s(\underline{\mu})} \sum_{t=1}^n X_t \cdot 1\{b \text{ χρησιμοποιείται την περίοδο } t\}.$$

Για να δείξουμε το θεώρημα θα αποδείξουμε ότι

$$\lim_{n \rightarrow \infty} \frac{SS_i(n)}{n} = 0, \quad i = 1, \dots, k \quad (2.6)$$

$$\lim_{n \rightarrow \infty} \sum_{b \notin s(\underline{\mu})} \frac{Y^b(n)}{n} = 0, \quad \text{σχ.β.}, \quad (2.7)$$

$$\lim_{n \rightarrow \infty} \frac{W_n}{n} = z^*(\underline{\mu}), \quad \text{σχ.β.}. \quad (2.8)$$

Πρώτον, το (2.6) ισχύει αφού $\tau_{i,m}$ είναι αραιές για όλα τα i . Για να δείξουμε το (2.7), σε περιόδους με μη εξαναγκαστικές επιλογές, έτσι ώστε να κάνουμε δειγματοληψία από μια ΒΕΛ b είναι υποχρεωτικό αλλά όχι αρκετό ότι $b \in s(\hat{\underline{\mu}}_n)$, έτσι

$$Y^b(n) \leq \sum_{t=1}^n 1\{b \in s(\hat{\underline{\mu}}_t)\}.$$

Για κάθε $b \in s(\hat{\underline{\mu}}_t)$ και $b \notin s(\underline{\mu})$, ακολουθεί από το Λήμμα 2.3.1 ότι

$$\|\hat{\underline{\mu}}_t - \underline{\mu}\| \geq \epsilon.$$

Οπότε, για $b \notin s(\underline{\mu})$

$$\begin{aligned} Y^b(n) &\leq \sum_{t=1}^n 1\{b \in s(\hat{\underline{\mu}}_t)\} \\ &\leq \sum_{t=1}^n 1\{\|\hat{\underline{\mu}}_t - \underline{\mu}\| \geq \epsilon\} \end{aligned}$$

έτσι,

$$\frac{Y^b(n)}{n} \leq \frac{1}{n} \sum_{t=1}^n 1_{\{\|\hat{\mu}_t - \underline{\mu}\| \geq \epsilon\}} \rightarrow 0, \quad n \rightarrow \infty, \quad \text{σχ.β.},$$

επειδή $\hat{\mu}_t \rightarrow \underline{\mu}$, σχ.β., αφού το $\hat{\mu}_t$ είναι ισχυρή συνεπής εκτιμήτρια, έτσι το (2.7) ισχύει.

Τώρα για να δείξουμε το (2.8) ξαναγράφουμε το W_n ως

$$\begin{aligned} \frac{W_n}{n} &= \frac{1}{n} \sum_{b \in s(\underline{\mu})} \sum_{t=1}^n X_t \cdot 1_{\{b \text{ χρησιμοποιείται την } t\}} \\ &= \frac{1}{n} \sum_{b \in s(\underline{\mu})} \sum_{j \in b} \sum_{t=1}^n X_t \cdot 1_{\{b \text{ χρησιμοποιείται την } t \text{ και έχουμε επιλέξει τον } j\}} \\ &= \frac{1}{n} \sum_{b \in s(\underline{\mu})} \sum_{j \in b} Y_j^b(n) \cdot \bar{X}_{j, Y_j^b(n)} \\ &= \sum_{b \in s(\underline{\mu})} \frac{Y^b(n)}{n} \cdot \sum_{j \in b} \frac{Y_j^b(n)}{Y^b(n)} \cdot \bar{X}_{j, Y_j^b(n)}. \end{aligned}$$

Από αυτή την έκφραση ακολουθεί ότι

$$\begin{aligned} \frac{W_n}{n} - z^* &= \sum_{b \in s(\underline{\mu})} \frac{Y^b(n)}{n} \cdot \sum_{j \in b} \frac{Y_j^b(n)}{Y^b(n)} \cdot \bar{X}_{j, Y_j^b(n)} - z^* \\ &= \sum_{b \in s(\underline{\mu})} \frac{Y^b(n)}{n} \cdot z_n^b - z^*, \end{aligned}$$

όπου $z_n^b = \sum_{j \in b} \frac{Y_j^b(n)}{Y^b(n)} \cdot \bar{X}_{j, Y_j^b(n)}$.

Αφού $Y(n) = \sum_{b \in s(\underline{\mu})} Y^b(n)$, έχουμε ότι

$$\begin{aligned} \frac{W_n}{n} - z^* &= \sum_{b \in s(\underline{\mu})} \frac{Y^b(n)}{n} \cdot z_n^b - z^* + \frac{Y(n)}{n} z^* - \frac{Y(n)}{n} z^* \\ &= \sum_{b \in s(\underline{\mu})} \frac{Y^b(n)}{n} \cdot (z_n^b - z^*) - (1 - \frac{Y(n)}{n}) z^*. \end{aligned}$$

Για να δείξουμε την (2.8) θα αποδείξουμε ότι

$$\frac{Y^b(n)}{n} \cdot (z_n^b - z^*) \rightarrow 0 \text{ σχ.}\beta. \quad \forall b \in s(\underline{\mu}), \text{ και } \frac{Y(n)}{n} \rightarrow 1, \text{ σχ.}\beta..$$

Η τυχαία μεταβλητή $Y^b(n)$ είναι αύξουσα ως προς n και $0 \leq Y^b(n) \leq n$, έτσι είτε $Y^b(n) \rightarrow \infty$ ή $Y^b(n) \rightarrow M$ για κάποιο $M < \infty$. Ορίζουμε τα παρακάτω ενδεχόμενα:

$$D = \{Y^b(n) \rightarrow \infty\} \text{ και } D^c = \{Y^b(n) \rightarrow M\}.$$

Τώρα έστω $P(D) = p$ και $P(D^c) = 1 - p$. Επίσης έστω

$$A = \left\{ \lim_{n \rightarrow \infty} \frac{Y^b(n)}{n} \cdot (z_n^b - z^*) = 0 \right\}.$$

Τότε $P(A) = P(A|D) \cdot p + P(A|D^c) \cdot (1 - p)$.

Τώρα,

$$\begin{aligned} P(A|D) &= P\left(\lim_{n \rightarrow \infty} \frac{Y^b(n)}{n} \cdot (z_n^b - z^*) = 0 \mid \lim_{n \rightarrow \infty} Y^b(n) = \infty\right) \\ &\geq P\left(\lim_{n \rightarrow \infty} z_n^b - z^* = 0 \mid \lim_{n \rightarrow \infty} Y^b(n) = \infty\right) \\ &= 1, \end{aligned}$$

από τον ισχυρό νόμο των μεγάλων αριθμών, αφού $\frac{Y^b(n)}{n} \leq 1 \quad \forall n$, και

$$P(A|D^c) = P\left(\lim_{n \rightarrow \infty} \frac{Y^b(n)}{n} \cdot (z_n^b - z^*) = 0 \mid \lim_{n \rightarrow \infty} Y^b(n) = M < \infty\right) = 1,$$

αφού σε αυτή την περίπτωση το $z_n^b - z^*$ είναι φραγμένο για κάθε πεπερασμένο n .

Οπότε, $P(A) = 1$, έτσι

$$\frac{Y^b(n)}{n} \cdot (z_n^b - z^*) \rightarrow 0, \quad n \rightarrow \infty, \text{ σχ.}\beta., \quad \forall b \in s(\underline{\mu}).$$

Τελικά,

$$\frac{Y(n)}{n} = \sum_{b \in s(\underline{\mu})} \frac{Y^b(n)}{n} = 1 - \sum_{t=1}^n \frac{SS_t(n)}{n} - \sum_{b \notin s(\underline{\mu})} \frac{Y^b(n)}{n} \rightarrow 1, \text{ σχ.}\beta., \quad n \rightarrow \infty.$$

Έτσι η απόδειξη του Θεωρήματος είναι ολοκληρωμένη.

□

2.4 Ταχύτητα σύγκλισης και προσομοιώσεις

Από τα αποτελέσματα της προηγούμενης ενότητας συμπεραίνουμε ότι υπάρχει σημαντική ευελιξία στην κατασκευή μιας συνεπούς πολιτικής δειγματοληψίας. Πράγματι, κάθε συλλογή από αραιές ακολουθίες η οποία ικανοποιεί την (2.3) εγγυάται ότι το Θεώρημα 2.3.1 ισχύει.

Σε αυτή την ενότητα προτείνουμε μια εκλέπτυνση της έννοιας της συνέπειας και εξετάζουμε πώς η ταχύτητα σύγκλισης της μέσης επιστροφής στην βέλτιστη τιμή επηρεάζεται από διαφορετικούς τύπους αραιών ακολουθιών. Επιπλέον, αφού η ανάλυση ευαισθησίας θα παρουσιαστεί μέσω προσομοίωσης, είναι καταλληλότερο να χρησιμοποιήσουμε την μέση τιμή της απόκλισης ως κριτήριο σύγκλισης. Έτσι θεωρούμε την αναμενόμενη διαφορά της μέσης επιστροφής κάτω από μια συνεπή πολιτική π από την βέλτιστη τιμή:

$$d_n^\pi(\underline{\mu}) = E^\pi \left(\frac{W_n}{n} \right) - z^*(\underline{\mu}).$$

Παρατηρούμε ότι η σχεδόν βέβαιη σύγκλιση του $\frac{W_n}{n}$ στο $z^*(\underline{\mu})$ που αποδεικνύεται στο Θεώρημα 2.3.1 δεν συνεπάγεται σύγκλιση κατα μέση τιμή, αν δεν γίνουν επιπλέον τεχνικές υποθέσεις για τις άγνωστες κατανομές. Για να μπορέσουμε να προσομοιώσουμε, θα υποθέσουμε επιπλέον ότι οι επιστροφές κάθε πληθυσμού είναι απολύτως φραγμένες με πιθανότητα ένα, π.χ., $P(|X_j| \leq u) = 1$, για κάποιο $u > 0$. Κάτω από αυτή την υπόθεση είναι εύκολο να δείξουμε ότι το Θεώρημα 2.3.1 συνεπάγεται

$$\lim_{n \rightarrow \infty} d_n^\pi(\underline{\mu}) = 0, \quad (2.9)$$

για οποιαδήποτε συνεπή πολιτική π και κάθε διάνυσμα $\underline{\mu}$.

Για να εντοπίσουμε την ταχύτητα σύγκλισης στην (2.9), παρουσιάζουμε μια προσομοίωση, για ένα πρόβλημα με $k = 4$ πληθυσμούς. Οι επιστροφές του πληθυσμού i ακολουθούν διωνυμική κατανομή με παραμέτρους (N, p_i) , όπου $p_1 = 0.3, p_2 = 0.5, p_3 = 0.9, p_4 = 0.8$. Το διάνυσμα των αναμενόμενων τιμών είναι $\underline{\mu} = (1.5, 2.5, 4.5, 4)$. Το διάνυσμα με τα $c = (3, 4, 8, 10)$ και το $C_0 = 5$. Κάτω από αυτές τις τιμές η βέλτιστη

πολιτική του προβλήματος πλήρους πληροφόρησης είναι $x = (0, 3/4, 1/4, 0)$, $y = 0$ και $z^*(\underline{\mu}) = 3$, π.χ., είναι βέλτιστο να γίνει τυχαιοποίηση ανάμεσα στους πληθυσμούς 2 και 3, το αναμενόμενο κόστος δειγματοληψίας ανα περίοδο είναι ίσο με 5 και η αναμενόμενη μέση επιστροφή ανα περίοδο ίση με 3.

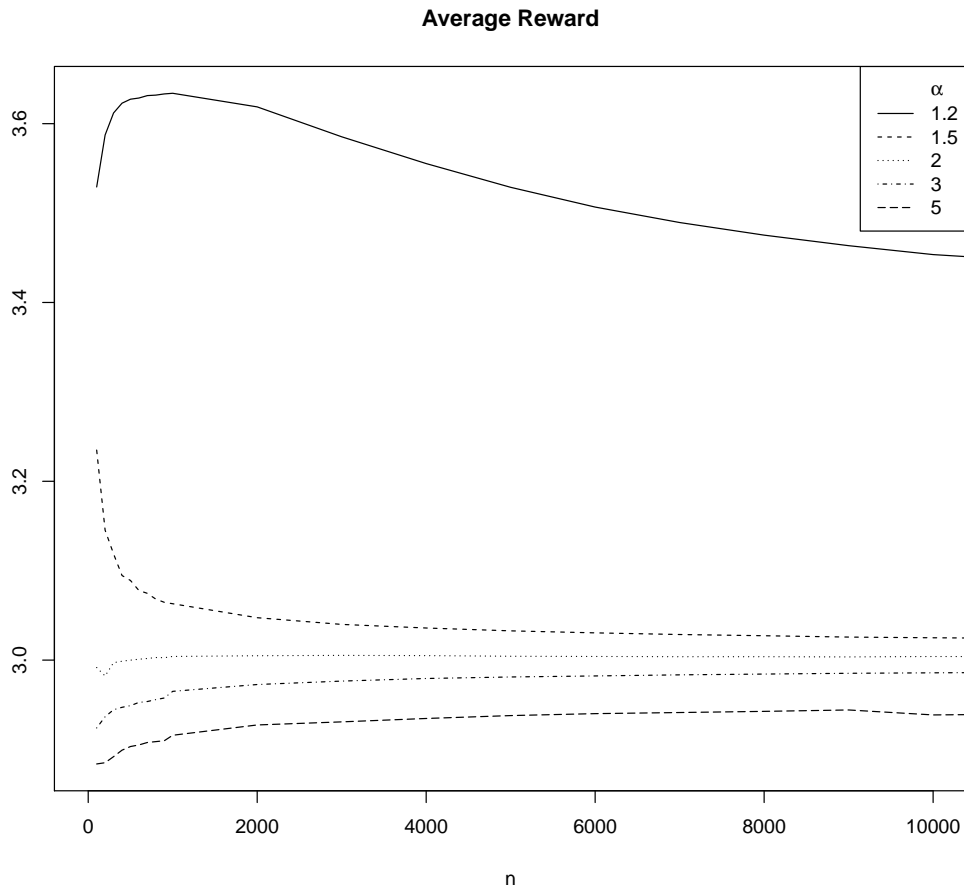
Για το παραπάνω πρόβλημα προσομοιώνουμε την απόδοση μιας συνεπούς πολιτικής με αραιές ακολουθίες του τύπου:

$$\{\tau_{j,m} = \ell_j + m^b, m = 1, 2, \dots, \}, j = 1, \dots, k,$$

όπου ℓ_j είναι κατάλληλα ορισμένες σταθερές οι οποίες διασφαλίζουν ότι οι ακολουθίες δεν επικαλύπτονται, και η εκθετική παράμετρος b είναι κοινή για όλους τους πληθυσμούς. Συγκρίνουμε την ταχύτητα σύγκλισης της (2.9) για πέντε τιμές του b : (1.2, 1.5, 2, 3, 5). Για κάθε τιμή του b η αντίστοιχη πολιτική προσομοιώθηκε με 1000 σενάρια μήκους $n = 10^4$ περιόδων το καθένα, ώστε να δείξουμε μια εκτίμηση της αναμενόμενης μέσης επιστροφής ανα περίοδο $d_n^{\pi}(\underline{\mu})$. Τα αποτελέσματα των προσομοιώσεων παρουσιάζονται στο γράφημα 2.1.

Παρατηρούμε στο γράφημα 2.1 ότι η σύγκλιση είναι πιο αργή για μικρές και μεγάλες τιμές του b και ταχύτερη για ενδιάμεσες τιμές. Ειδικότερα για $b = 1.2$ η διαφορά είναι σχετικά μεγάλη ακόμα και μετά από 10000 περιόδους. Αυτό εξηγείται ως εξής. Για μικρές τιμές του b οι εξαναγκαστικές επιλογές είναι πιο συχνές. Αφενός αυτό έχει το επιθυμητό αποτέλεσμα οι μέσες εκτιμήσεις για όλους τους πληθυσμούς να γίνονται ακριβείς πολύ σύντομα, αφετέρου όμως σημαίνει ότι οι μη βέλτιστοι πληθυσμοί είναι επίσης επιλεγμένοι συχνά λόγω των εξαναγκαστικών επιλογών. Σαν αποτέλεσμα η μέση επιστροφή μπορεί να αποκλίνει από την πραγματική βέλτιστη τιμή για περισσότερες χρονικές περιόδους. Από την άλλη μεριά, για μεγάλες τιμές του b όλες οι ακολουθίες τ_j γίνονται πολύ αραιές και έτσι οι εξαναγκαστικές επιλογές είναι σπάνιες. Σε αυτή την περίπτωση παίρνει περισσότερο χρόνο να εκτιμήσουμε την σύγκλιση, και τα Π.Γ.Π. μπορεί να εισάγουν μη βέλτιστες λύσεις για μεγάλα διαστήματα.

Από την παραπάνω συζήτηση συνεπάγεται ότι οι ενδιάμεσες τιμές του b είναι γενικά περισσότερο προτιμώμενες, αφού προσφέρουν μια καλύτερη ισοροπία ανάμεσα στα δύο αποτελέσματα, γρήγορη εκτίμηση όλων των μέσων τιμών και αποφυγή των μη βέλτιστων πληθυσμών. αυτό είναι επίσης προφανές και στο γράφημα, όπου η τιμή $b = 2$ φαίνεται να είναι η καλύτερη από πλευράς ταχύτητας σύγκλισης.



Σχήμα 2.1: Σύγκριση της Ταχύτητας Σύγκλισης Αραιών Ακολουθιών

Για να αντιμετωπιστεί το ζήτημα της ακρίβειας των ταχυτήτων σύγκλισης που βασίζονται στην προσομίωση, το Γράφημα 2.2 παρουσιάζει μια περιοχή εμπιστοσύνης 95% για την καμπύλη της μέσης επιστροφής που αντιστοιχεί στο $b = 2$, και είναι βασισμένη σε 1000 προσομοιωμένα σενάρια. Η περιοχή εμπιστοσύνης είναι γενικά πολύ στενή (παρατηρείστε ότι οι άξονες στά δύο γραφήματα έχουν διαφορετικές κλίμακες), έτσι η εκτίμηση της αναμενόμενης μέσης επιστροφής είναι αρκετά ακριβής. Αυτή είναι επίσης ίδια περίπτωση και για τις άλλες καμπύλες, οπότε η σύγκριση των ταχυτήτων σύγκλισης είναι έγκυρη. Επιπλέον, το μήκος των διαστημάτων εμπιστοσύνης γίνεται μικρότερο για μεγάλες τιμές περιόδων αφού, όπως αναμέναμε, η σύγκλιση στην πραγματική τιμή είναι καλύτερη για μεγαλύτερης διάρκειας σενάρια.

Ένα άλλο θέμα που πηγάζει από το Γράφημα 2.1 είναι το ακόλουθο. Για $b = 1.2$ η μέση επιστροφή συγκλίνει πολύ αργά στο z^* , αλλά παραμένει πάνω από αυτή για όλη την διάρκεια του σεναρίου. Έτσι θα μπορούσε να υποστηριχτεί ότι, παρ' όλο που η σύγκλιση δεν είναι καλή, αυτή την πολιτική την προτιμούμε, επειδή αποδίδει μεγαλύτερες μέσες επιστροφές σε σχέση με τις άλλες πολιτικές. Αυτό επίσης φαίνεται να διαψεύδει ότι το z^* είναι η μέγιστη αναμενόμενη επιστροφή κάτω από πλήρη πληροφόρηση, αφού υπάρχει μια πολιτική δειγματοληψίας η οποία κάτω από ελλιπή πληροφόρηση αποδίδει καλύτερα.

Ο λόγος για αυτή την φαινομενική αντίφαση σχετίζεται με την μορφή που έχει ο περιορισμός στο κόστος (2.2). Ο περιορισμός απαιτεί το άπειρο ορίζοντα αναμενόμενο κόστος να μην ξεπερνά το C_0 . Αυτό δεν αποκλείει την πιθανότητα ένας ή παραπάνω πληθυσμοί με μεγάλα κόστη δειγματοληψίας και μεγάλες αναμενόμενες επιστροφές να χρησιμοποιούνται για αυθαίρετα μεγάλα διαστήματα πριν πάμε σε μια περιορισμένη-βέλτιστη πολιτική για τον υπόλοιπο άπειρο ορίζοντα. Τέτοιες πολιτικές μπορεί να επιτυγχάνουν μεγαλύτερη μέση επιστροφή από το z^* για μεγάλα διαστήματα, ωστόσο αυτό επιτυγχάνεται επειδή 'δανειζόμαστε', π.χ., παραβιάζοντας τον περιορισμό στο κόστος, επίσης για μεγάλες χρονικές περιόδους. Αφού η (2.2) απαιτείται να ισχύει μόνο στο όριο, τέτοιες πολιτικές επιτρέπονται.

Οι συνεπείς πολιτικές της Ενότητας 2.3 δεν είναι σχεδιασμένες ειδικά να παίρνουν αυτό το όφελος, αλλά ούτε και να το αποφεύγουν. Οπότε, είναι πιθανό, όπως συμβαίνει και για το $b = 2$, η συνεπής πολιτική να επιτυγχάνει μεγαλύτερη επιστροφή από την βέλτιστη μέση επιστροφή για αρκετές χρονικές περιόδους πριν τελικά συγκλίνει στο z^* .

Η παραπάνω συζήτηση δείχνει ότι ο περιορισμός όπως εκφράζεται στην (2.2), μπορεί

να μην είναι κατάλληλος, αν για παράδειγμα το κόστος δειγματοληψίας είναι ένα από ποσό το οποίο πρέπει να πληρώνεται κάθε φορά που παίρνουμε μια παρατήρηση σε μια περίοδο, και υπάρχει ένα ποσό C_0 ανα περίοδο για δειγματοληψία. Σε αυτή την κατάσταση μια πολιτική μπορεί να προτείνει να ξεπεράσουμε το ποσό που διαθέτουμε για δειγματοληψία για μεγάλες χρονικές περιόδους και ακόμα να παραμένει εφικτή, κάτι τέτοιο δεν μπορεί να είναι βιώσιμο στην πραγματικότητα. Σε τέτοιες περιπτώσεις θα είναι πιο ρεαλιστικό να εισάγουμε έναν αυστηρότερο περιορισμό στο μέσο κόστος, για παράδειγμα να απαιτείται ότι η (2.2) ισχύει για όλα τα n και όχι μόνο στο όριο.

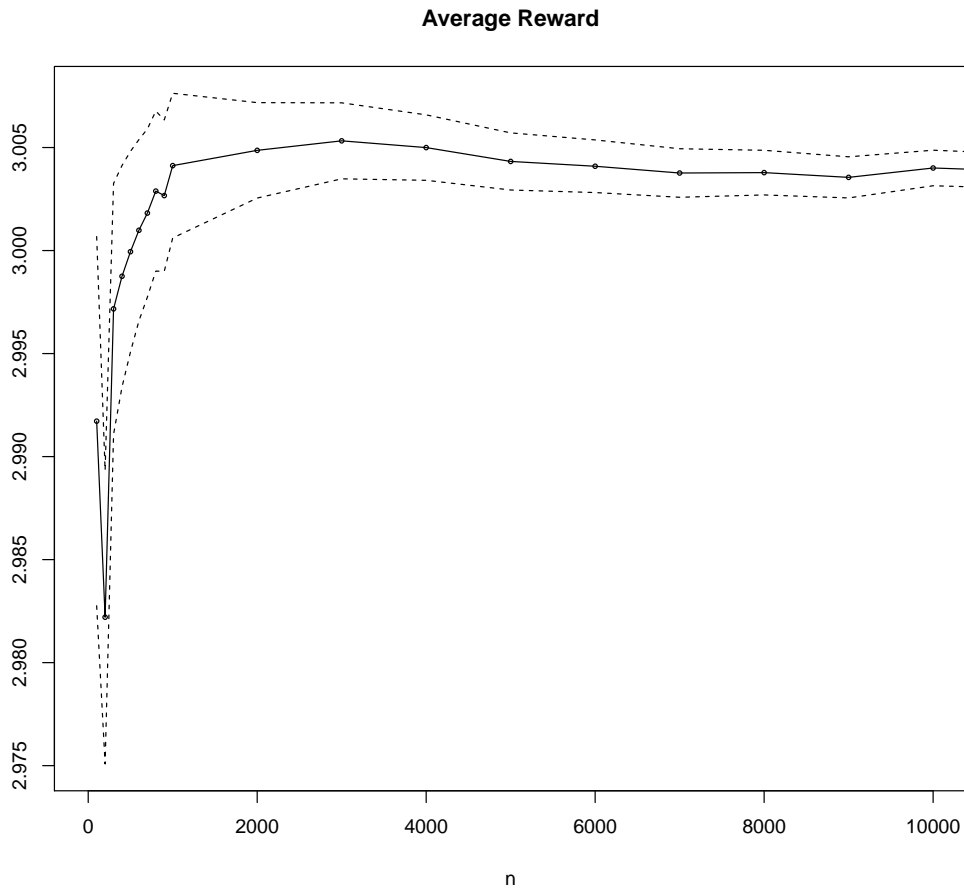
2.5 Ανασκόπηση και επεκτάσεις

Σε αυτό το Κεφάλαιο αναπτύξαμε μια οικογένεια συνεπών προσαρμοστικών πολιτικών για ακολουθιακή δειγματοληψία από k ανεξάρτητους πληθυσμούς με άγνωστες κατανομές κάτω από έναν ασυμπτωτικό περιορισμό στο μέσο κόστος. Η κύρια ιδέα στην ανάπτυξη αυτής της κλάσης πολιτικών είναι να χρησιμοποιήσουμε αραιές ακολουθίες από εξαναγκαστικές επιλογές σε διάφορες περιόδους για κάθε πληθυσμό, να εισάγουμε συνεπείς εκτιμήσεις όλων των άγνωστων μέσων τιμών και στις υπόλοιπες χρονικές περιόδους να χρησιμοποιήσουμε την λύση που εμφανίζεται από το Π.Γ.Π. που χρησιμοποιεί τις εκτιμήσεις αντί τις πραγματικές μέσες τιμές. Επίσης παρουσιάσαμε μια προσομοίωση η οποία συγκρίνει την ταχύτητα σύγκλισης για διάφορες πολιτικές της κλάσης.

Όλα αυτά μπορούν να επεκταθούν σε διάφορες κατευθύνσεις. Πρώτον, όπως δείξαμε στην Ενότητα 2.4, η ασυμπτωτική μορφή του περιορισμού στο κόστος είναι κατα μια έννοια χαλαρή, αφού επιτρέπει το μέσο κόστος δειγματοληψίας να ξεπερνά το άνω φράγμα για αυθαίρετα μεγάλες χρονικές περιόδους και ακόμα να ικανοποιείται στο όριο. Ένα πιο κατάλληλο, μολονότι πιο περίπλοκο, μοντέλο θα μπορούσε να απαιτεί ο περιορισμός στο κόστος να ικανοποιείται σε κάθε περίοδο. Η κατασκευή συνεπούς και ακόμα σημαντικότερα, αποδοτικών πολιτικών κάτω από τον αυστηρότερο περιορισμό στο κόστος θα παρουσιαστεί στο επόμενο Κεφάλαιο.

Μια άλλη επέκταση είναι στην κατεύθυνση του ελέγχου Μαρκοβιανών διαδικασιών. Αντί να υποθέσουμε διακριτούς ανεξάρτητους πληθυσμούς με ανεξάρτητες και ισόνομες παρατηρήσεις, να θεωρήσουμε μια μέσης επιστροφής Μαρκοβιανή διαδικασία αποφάσεων με άγνωστο νόμο μεταπηδήσεων και/ή κατανομές επιστροφών, και ένα ή περισσότερους

μη ασυμπτωτικούς περιορισμούς στο μέσο κόστος. Σε αυτή την περίπτωση το πρόβλημα είναι στο να κατασκευάσεις συνεπείς και, ακόμα πιο σημαντικό, αποδοτικές πολιτικές ελέγχου, επεκτείνοντας τα αποτελέσματα του [14] στην περίπτωση με περιορισμό.



Σχήμα 2.2: Περιοχή Εμπιστοσύνης της Μέσης Αμοιβής για $b = 2$.

Κεφάλαιο 3

Αποδοτική πολιτική για ακολουθιακή δειγματοληψία κάτω από ελλιπή πληροφόρηση και περιορισμούς σε κάθε βήμα

3.1 Εισαγωγή

Σε αυτό το Κεφάλαιο θεωρούμε ένα πρόβλημα προσαρμοστικού ελέγχου το οποίο εκφράζεται σαν ακολουθιακή δειγματοληψία κάτω από ελλιπή πληροφόρηση και περιορισμούς βήματος.

Το πλαίσιο εφαρμογής περιλαμβάνει περιπτώσεις όπου ο αποφασίζων πρέπει να επιλέξει ακολουθιακά ανάμεσα σε κερδοφόρες ευκαιρίες, η επιστροφή των οποίων είναι τυχαία και μερικώς γνωστής κατανομής. Αυτό αυξάνει την ανάγκη για προσαρμοστικές πολιτικές επιλογών, το οποίο ενσωματώνει τον πειραματισμό σε προβλήματα ελέγχου. Σε πολλές εφαρμογές το πρόβλημα είναι πολύ περίπλοκο από το γεγονός ότι οι περιορισμοί επιβάλλονται στην πολιτική επιλογών. Οι περιορισμοί μπορούν να έρθουν από διάφορες πηγές όπως οικονομικό κόστος στο να προσπαθείς σε διαφορετικές επιλογές, νόμιμες ή άλλες απαιτήσεις οι οποίες απαγορεύουν να προσπαθήσουμε σε διάφορες επιλογές για

συγκεκριμένα ποσοστά χρόνου, κτλ.

Η ύπαρξη του κόστους δειγματοληψίας εισάγει μια νέα διάσταση στα καθιερωμένα προβλήματα ανάμεσα στον πειραματισμό και την μεγιστοποίηση. Το κόστος δειγματοληψίας μπορεί να απαγορεύει την χρήση επιλογών με μεγάλες μέσες επιστροφές επειδή το δειγματοληπτικό τους κόστος μπορεί να είναι πολύ υψηλό. Για αυτό, ο αποφασίζων πρέπει να ταυτοποιεί το υποσύνολο των πληθυσμών με τον καλύτερο συνδυασμό επιστροφών σε σχέση με το κόστος και να κατανέμει τους διαθέσιμους πόρους για δειγματοληψία με βέλτιστο τρόπο. Περιλαμβάνοντας περιορισμούς με τέτοιο τρόπο, εμπλουτίζουμε το πλαίσιο εφαρμογής, σε καταστάσεις με περιορισμένους πόρους (όπως π.χ. περιορισμός στην παραγωγή ή την αποθήκευση, το κεφάλαιο κτλ.) που πρέπει να κατανέμονται αποδοτικά κάτω από μερική πληροφόρηση.

Όταν έχουμε πλήρη πληροφόρηση η λύση του προβλήματος κατανομής της προσπάθειας κάτω από τους περιορισμούς στο κόστος μπορεί να μοντελοποιηθεί σαν ένα πρόβλημα γραμμικού ή μη γραμμικού προγραμματισμού. Από την άλλη μεριά όταν έχουμε άγνωστες παραμέτρους του προβλήματος και πρέπει να εκτιμηθούν μέσω πειραματισμού, ο αποφασίζων πρέπει να χρησιμοποιήσει προσαρμοστικές πολιτικές οι οποίες εξασφαλίζουν γνώση για τις παραμέτρους και την ίδια στιγμή εξασφαλίζουν ότι το όφελος που θυσιάζεται για γνώση είναι όσο το δυνατό μικρότερο.

Το θέμα ανάμεσα στον πειραματισμό και την μεγιστοποίηση του οφέλους έχει εξεχούσα θέση στην κλάση των προβλημάτων προσαρμοστικού ελέγχου που εισήχθησαν στις εργασίες [12] και [8] με Μπεϋζιανή και κλασσική προσέγγιση, αντίστοιχα.

Η Μπεϋζιανή μοντελοποίηση σχετίζεται με τα multi-armed bandit προβλήματα, και αναπτύχθηκε κυρίως κάτω από ένα κριτήριο προεξοφλημένου κέρδους πεπερασμένου ορίζοντα ([17], [21] και [18]).

Σε αυτό το Κεφάλαιο ακολουθούμε μη Μπεϋζιανή προσέγγιση, όπου το πρόβλημα μοντελοποιείται μέσω μιας συνάρτησης απώλειας η οποία ελαχιστοποιείται με διάφορες έννοιες ([4], [2] και [14]).

Η υπάρχουσα βιβλιογραφία δεν είναι τόσο εκτεταμένη όταν εμφανίζονται περιορισμοί. Στο προεξοφλημένο πλαίσιο του multi-armed bandit προβλήματος, στο [11], θεωρήθηκαν στοχαστικά bandits με άγνωστο σταθερό περιβάλλον. Πρότειναν έναν αλγόριθμο, βασισμένο στην προσέγγιση steering, και ανέλυσαν την απώλεια συγκρίνοντας την βέλτιστη στάσιμη πολιτική η οποία γνωρίζει τις στατιστικές παραμέτρους

των διαφορετικών πληθυσμών. Ένα multi-armed bandit με περιορισμούς θεωρήθηκε στο [16], όπου για την περίπτωση της γραμμικής ωφέλειας, εισήγαγαν περιορισμούς οι οποίοι συνδέουν τα bandits και έλυσαν το περιορισμένο πρόβλημα βελτιστοποίησης με γενίκευση στηλών. Στο [15] βρήκαν πολιτικές δεικτών για τα προεξοφλημένα bandit προβλήματα με διαθέσιμους περιορισμούς το οποίο σημαίνει ότι τα arms δεν είναι πάντα διαθέσιμα. Επίσης, στο [20] απέδειξαν ότι υπάρχουν πολιτικές δείκτη για το πρόβλημα με περιορισμό χρησιμοποιώντας γραμμικό προγραμματισμό για να αναλύσουν μια πολιτική η οποία προσδιορίζει τις κορυφές με την μέγιστη επικάλυψη των δεικτών του Whittle. Τέλος, στο [19] θεώρησαν το multi-armed bandit πρόβλημα με μη σταθερά μετρικά κόστη.

Στο πλαίσιο της κλασικής προσέγγισης υπάρχουν ακόμα λιγότερες αναφορές. Μια σχετική προσέγγιση είναι αυτή που βασίζεται στις στοχαστικές προσεγγίσεις [27].

Σε αυτή την περιοχή έχουμε θεωρήσει την περίπτωση όπου ο περιορισμός στο κόστος είναι επίσης ασυμπτωτικός, με την έννοια ότι το κόστος, άπειρου ορίζοντα ανα περίοδο δεν επιτρέπεται να ξεπεράσει ένα δεδομένο άνω φράγμα. Για αυτό το πρόβλημα έχουμε κατασκευάσει στο προηγούμενο Κεφάλαιο μια προσαρμοστική πολιτική η οποία είναι συνεπής και επιτυγχάνει την βέλτιστη τιμή του αντίστοιχου προβλήματος κάτω από πλήρη πληροφόρηση. Αυτή η έρευνα είναι δημοσιευμένη στο [13].

Σε αυτό το κεφάλαιο θεωρούμε έναν αυστηρότερο περιορισμό στο κόστος κάθε βήματος, το οποίο απαιτεί το μέσο κόστος των επιλογών να μην ξεπερνά το άνω φράγμα σε κάθε περίοδο και όχι μόνο οριακά. Η βέλτιστη πολιτική του προηγούμενου μοντέλου είναι μια πολιτική τυχαιοποίησης η οποία δεν είναι εφικτή στο αυστηρότερο μοντέλο επειδή η τυχαιοποίηση μπορεί να απαιτεί να επιλέξουμε με τέτοιο τρόπο όπου το μέσο κόστος να ξεπερνά το άνω φράγμα, για πεπερασμένες περιόδους. Έτσι η σειρά των επιλογών μπορεί να γίνει με τέτοιο τρόπο ώστε ο περιορισμός στο κόστος να μην παραβιάζεται ποτέ, στο πλαίσιο ελλιπούς πληροφόρησης. Για αυτό το πρόβλημα έχουμε κατασκευάσει μια κατ' ουσίαν διαφορετική πολιτική επιλογών η οποία δεν είναι μόνο συνεπής, αλλά επίσης ασυμπτωτικά αποδοτική, με την έννοια ότι επιτυγχάνει ένα ασυμπτωτικό κάτω φράγμα με τον ρυθμό αύξησης της συνάρτησης απώλειας. Η πολιτική είναι βασισμένη στην ιδέα των επιλογών σε μπλόκ πεπερασμένου μεγέθους με τέτοιο τρόπο ώστε μέσα σε κάθε μπλόκ ο περιορισμός στο κόστος να ικανοποιείται.

Το Κεφάλαιο έχει την ακόλουθη δομή. Στην δεύτερη Ενότητα περιγράφουμε το μο-

ντέλο σε πλήρη και ελλιπή πληροφόρηση και ορίζουμε την συνάρτηση απώλειας. Στην τρίτη Ενότητα βρίσκουμε ένα κάτω φράγμα για την συνάρτηση απώλειας. Στην τέταρτη Ενότητα κατασκευάζουμε μια προσαρμοστική πολιτική και αποδεικνύουμε ότι είναι ασυμπτωτικά αποδοτική.

3.2 Περιγραφή μοντέλου

Θεωρούμε το πρόβλημα ακολουθιακής δειγματοληψίας από k ανεξάρτητους στατιστικούς πληθυσμούς, $i = 1, \dots, k$. Διαδοχικές δειγματοληψίες από έναν πληθυσμό i αποτελούν μια ακολουθία από ανεξάρτητες και ισόνομες τυχαίες μεταβλητές X_{i1}, X_{i2}, \dots ακολουθώντας μια κατανομή μιας παραμέτρου με πυκνότητα $f_i(\cdot|\theta_i)$ ως προς ένα μη εκφυλισμένο μέτρο ν . Η πυκνότητα $f_i(\cdot|\cdot)$ είναι γνωστή και το θ_i είναι μια παράμετρος η οποία ανήκει σε κάποιο σύνολο Θ_i . Έστω ότι το $\underline{\theta} = (\theta_1, \dots, \theta_k)$ δηλώνει το σύνολο των παραμέτρων, $\underline{\theta} \in \Theta$ όπου $\Theta \equiv \Theta_1 \times \dots \times \Theta_k$. Έτσι το στοχαστικό μοντέλο είναι προσδιορίζεται μοναδικά από το διάνυσμα $\underline{\theta}$. Δοθέντος του $\underline{\theta}$ έστω ότι $\underline{\mu}(\underline{\theta})$ είναι το διάνυσμα των μέσων τιμών, π.χ. $\mu_i(\theta_i) = E_{\underline{\theta}}(X_i)$. Η πραγματική τιμή του $\underline{\theta}$ δεν είναι γνωστή. Σε αυτό το Κεφάλαιο θεωρούμε την περίπτωση όπου οι πυκνότητες $f_i(y|\theta_i)$, $i = 1, \dots, k$, ανήκουν στην εκθετική οικογένεια με φυσική παράμετρο το θ ,

$$f_i(y; \theta) = e^{\theta y - \psi(\theta)},$$

με παραμετρικό χώρο Θ_i . Για την εκθετική οικογένεια έχουμε ότι η $\mu_i(\theta) = \psi'_i(\theta) < \infty$ είναι αύξουσα ως προς θ αφού $\psi''_i(\theta) = \text{Var}_{i,\theta} Y$.

Σε κάθε περίοδο πρέπει να επιλέξουμε έναν πληθυσμό για δειγματοληψία. Κάνοντας δειγματοληψία από τον πληθυσμό i έχουμε ένα κόστος c_i ανα δειγματοληψία και χωρίς βλάβη της γενικότητας υποθέτουμε $c_1 \leq c_2 \leq \dots \leq c_k$, αλλά όχι όλα ίσα και $c_i, C_0 \in \mathbb{Q}$. Το ζητούμενο είναι να μεγιστοποιήσουμε την αναμενόμενη μέση επιστροφή υπό τον περιορισμό ότι το μακροπρόθεσμο μέσο κόστος δειγματοληψίας δεν ξεπερνά ένα δεδομένο άνω φράγμα C_0 για κάθε περίοδο. Χωρίς βλάβη της γενικότητας υποθέτουμε ότι $c_1 \leq C_0 < c_k$. Στην περίπτωση όπου $C_0 < c_1$ το πρόβλημα δεν είναι εφικτό και από την άλλη μεριά όπου $C_0 \geq c_k$ ο περιορισμός στο κόστος είναι περιττός. Έστω $d = \max\{j : c_j \leq C_0\}$. Τότε $1 \leq d < k$ και $c_d \leq C_0 < c_{d+1}$.

Το μοντέλο είναι μια αυστηρότερη εκδοχή του μοντέλου μέσου κόστους του προηγούμενου Κεφαλαίου. Στο παρόν μοντέλο η μορφή του περιορισμού απαιτεί το μέσο κόστος δειγματοληψίας να μην ξεπερνά το C_0 σε κάθε χρονική περίοδο και όχι μόνο στο όριο.

Το πρόβλημα πλήρους πληροφόρησης όπου το $\underline{\theta}$ είναι γνωστό μπορεί να μοντελοποιηθεί σαν ένα Π.Γ.Π όπως στο προηγούμενο Κεφάλαιο. Στο προηγούμενο Κεφάλαιο αποδείξαμε ότι μια πολιτική τυχαιοποίησης όπου σε κάθε περίοδο επιλέγει τον πληθυσμό j με πιθανότητα x_j , για $j = 1, \dots, k$, μεγιστοποιεί την αναμενόμενη επιστροφή όπου οι πιθανότητες τυχαιοποίησης x_j είναι μια βέλτιστη λύση του ακόλουθου Π.Γ.Π.

$$\begin{aligned}
 z^*(\underline{\theta}) &= \max \sum_{j=1}^k \mu_j(\theta_j) x_j \\
 &\sum_{j=1}^k c_j x_j + y = C_0 \\
 &\sum_{j=1}^k x_j = 1 \\
 &x_j \geq 0, \forall j, y \geq 0.
 \end{aligned} \tag{3.1}$$

Η πολιτική αυτή δεν είναι εφικτή στο αυστηρότερο πλαίσιο αυτού του Κεφαλαίου επειδή η τυχαιοποίηση μπορεί να απαιτήσει να κάνουμε δειγματοληψία με τέτοιο τρόπο ώστε να ξεπερνά το μέσο κόστος το C_0 , για κάποιες πεπερασμένες περιόδους. Ωστόσο, κάτω από την υπόθεση των ρητών συντελεστών c_j , κάθε βέλτιστη λύση του (3.1) η οποία είναι ακραίο σημείο είναι επίσης ρητή, έτσι η πολιτική τυχαιοποίησης μπορεί να εφαρμοστεί σαν μια περιοδική πολιτική δειγματοληψίας με συχνότητες c_j . Η σειρά των επιλογών μπορεί να γίνει έτσι ώστε ο περιορισμός στο κόστος δειγματοληψίας να μην παραβιάζεται ποτέ στο πρόβλημα πλήρους πληροφόρησης. Χρησιμοποιούμε γενικεύσεις αυτής της ιδέας στο πλαίσιο ελλιπούς πληροφόρησης στην επόμενη Ενότητα.

Εισάγουμε τώρα κάποιον αναγκαίο συμβολισμό αναφερόμενοι στο Π.Γ.Π. (3.1). Το δυικό του προβλήματος (3.1) είναι

$$\begin{aligned}
z_D^*(\theta) &= \min g + C_0\lambda \\
&g + c_1\lambda \geq \mu_1(\theta_1) \\
&\quad \vdots \\
&g + c_k\lambda \geq \mu_k(\theta_k) \\
&g \in \mathbb{R}, \lambda \geq 0.
\end{aligned}$$

Ένας βασικός πίνακας B είναι της μορφής $\begin{pmatrix} c_i & c_j \\ 1 & 1 \end{pmatrix}$, για κάποιο $i \leq d < j$ ή $\begin{pmatrix} c_i & 1 \\ 1 & 0 \end{pmatrix}$ για κάποιο $i \leq d$, όπου αντίστοιχα επιλέγουμε από το ζεύγος (i, j) ή τον πληθυσμό i . Ορίζουμε την Βασική Εφικτή Λύση (ΒΕΛ) που αντιστοιχεί στον πίνακα B ως $b = \{i, j\}$ ή $b = \{i\}$, αντίστοιχα. Παρατηρούμε ότι στην περίπτωση εκφυλισμένης ΒΕΛ b , περισσότεροι από ένας πίνακες B αντιστοιχούν στο ίδιο b .

Χρησιμοποιούμε το K για δηλώσουμε το σύνολο των ΒΕΛ:

$$K = \{b : b = \{i, j\}, i \leq d \leq j \text{ ή } b = \{i\}, i \leq d\}.$$

Αφού η εφικτή περιοχή του (3.1) είναι φραγμένη, το K είναι πεπερασμένο.

Για ένα βασικό πίνακα B , έστω ότι το $v^B = (\lambda^B, g^B)$ δηλώνει το δυικό διάνυσμα που αντιστοιχεί στον B , π.χ., $v^B = \mu_B(\theta)B^{-1}$, όπου $\mu_B(\theta) = (\mu_i(\theta_i), \mu_j(\theta_j))$, ή $\mu_B(\theta) = (\mu_i(\theta_i), 0)$, εξαρτάται από την μορφή του B .

Όσο αφορά την βελτιστότητα, μια ΒΕΛ είναι βέλτιστη αν και μόνο αν για τουλάχιστον έναν αντίστοιχο βασικό πίνακα B το reduced costs (dual slacks) είναι όλα μη αρνητικά:

$$\phi_\alpha^B(\theta) \equiv c_\alpha\lambda^B + g^B - \mu_\alpha(\theta_\alpha) \geq 0, \alpha = 1, \dots, k.$$

Ένας βασικός πίνακας B που ικανοποιεί αυτή τη συνθήκη είναι βέλτιστος. Παρατηρούμε ότι αν μια βέλτιστη ΒΕΛ είναι εκφυλισμένη, τότε δεν είναι υποχρεωτικά βέλτιστοι όλοι οι βασικοί πίνακες που αντιστοιχούν σε αυτή.

Είναι εύκολο να δείξουμε ότι το ελαττωμένο κόστος (reduced cost) μπορεί να εκφραστεί σαν γραμμικός συνδυασμός των άγνωστων μέσων τιμών των πληθυσμών, π.χ.,

$\phi_\alpha^B(\theta) = \underline{w}_\alpha^B \underline{\mu}(\theta)$, όπου \underline{w}_α^B είναι ένα κατάλληλα ορισμένο διάνυσμα όπου δεν εξαρτάται από το $\underline{\mu}(\theta)$.

Τέλος ορίζουμε το σύνολο των βέλτιστων λύσεων του (3.1) για ένα διάνυσμα $\underline{\mu}(\theta)$,

$$s(\theta) = \{b \in K : b \text{ αντιστοιχεί σε μια βέλτιστη ΒΕΛ}\}.$$

Στο πλαίσιο ελλιπούς πληροφόρησης του προβλήματος η παράμετρος $\underline{\theta}$ είναι άγνωστη. Ειδικότερα κάνουμε τις ακόλουθες υποθέσεις.

Υπόθεση 3.2.1. *Οι κατανομές των επιστροφών είναι ανεξάρτητες.*

Στην περίπτωση ελλιπούς πληροφόρησης είμαστε περιορισμένοι στις προσαρμοστικές πολιτικές οι οποίες εξαρτώνται μόνο από τις παρελθούσες παρατηρήσεις και επιστροφές.

Ειδικότερα, έστω ότι τα $A_t, X_t, t = 1, 2, \dots$ δηλώνουν τον επιλεγμένο πληθυσμό και την παρατηρούμενη επιστροφή στην περίοδο t . Έστω ότι το $h_t = (\alpha_1, x_1, \dots, \alpha_{t-1}, x_{t-1})$ δηλώνει την ιστορία των επιλογών και των παρατηρήσεων που είναι διαθέσιμες μέχρι την περίοδο t .

Μια προσαρμοστική πολιτική ορίζεται σαν μια ακολουθία $\pi = (\pi_1, \pi_2, \dots)$ από ιστορικά εξαρτημένες κατανομές πιθανοτήτων στους $\{1, \dots, k\}$, έτσι ώστε

$$\pi_t(j, h_t) = P(A_t = j | h_t).$$

Δοθέντων μιας πολιτικής π και μιας παραμέτρου $\underline{\theta}$ έστω η $P_{\underline{\theta}}^\pi$ να είναι το μέτρο πιθανότητας το οποίο είναι ορισμένο στις ιστορίες και $E_{\underline{\theta}}^\pi$ είναι η μέση τιμή ως προς το $P_{\underline{\theta}}^\pi$. Τώρα, δοθείσης της ιστορίας h_n , έστω ότι το $T_n(\alpha)$ δηλώνει τον αριθμό των φορών που ο πληθυσμός α έχει επιλεγεί κατά τη διάρκεια των πρώτων n περιόδων

$$T_n(\alpha) = \sum_{t=1}^n 1\{A_t = \alpha\}.$$

Έστω ότι το S_n^π είναι η επιστροφή μέχρι την περίοδο n :

$$S_n^\pi = \sum_{t=1}^n X_t,$$

και το C_n^π είναι το συνολικό κόστος μέχρι την περίοδο n :

$$C_n^\pi = \sum_{t=1}^n c_{A_t}.$$

Αυτές οι ποσότητες μπορούν να χρησιμοποιηθούν για να ορίσουμε τις ιδιότητες μιας προσαρμοστικής πολιτικής. Ειδικότερα:

Ορισμός 3.2.1. *Μια πολιτική π καλείται εφικτή αν*

$$\frac{C_n^\pi}{n} \leq C_0, \quad \forall n = 1, 2, \dots$$

Ορισμός 3.2.2. *Μια πολιτική π καλείται συνεπής αν είναι εφικτή και*

$$\lim_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^\pi(S_n^\pi)}{n} = z^*(\underline{\theta}), \quad \text{σχ.β. } \forall \underline{\theta} \in \Theta \text{ και } P_{\underline{\theta}}^\pi.$$

Ο πρώτος περιορισμός συνεπάγεται ότι το κόστος δειγματοληψίας δεν πρέπει να ξεπερνά το διαθέσιμο ποσό για δειγματοληψία. Αυτό πρέπει να ισχύει για κάθε n και για κάθε υλοποίηση της διαδικασίας και όχι μόνο κατά μέση τιμή ασυμπτωτικά όπως στο προηγούμενο κεφάλαιο. Ο δεύτερος ορισμός είναι η σύγκλιση της μακροπρόθεσμης μέσης επιστροφής ανα περίοδο κάτω από την πολιτική π στη βέλτιστη αναμενόμενη τιμή κάτω από πλήρη πληροφόρηση, για όλες τις πιθανές παραμέτρους των πληθυσμών στο Θ . Τώρα έστω ότι το Π^F και το Π^C δηλώνουν την κλάση των εφικτών και συνεπών πολιτικών, αντίστοιχα.

Στο πλαίσιο ελλιπούς πληροφόρησης είναι λογικό να απαιτήσουμε ότι μια προσαρμοστική πολιτική είναι συνεπής. Σε αυτό το Κεφάλαιο ψάχνουμε πολιτικές οι οποίες έχουν μια πιο ισχυρή ιδιότητα, όπως το ότι η σύγκλιση της αναμενόμενης μέσης επιστροφής στην βέλτιστη τιμή κάτω από πλήρη πληροφόρηση είναι η γρηγορότερη δυνατή. Για αυτό το λόγο ορίζουμε την συνάρτηση απώλειας της πολιτικής π ως την πεπερασμένου ορίζοντα απώλεια της αναμενόμενης επιστροφής σε σχέση με την βέλτιστη πολιτική κάτω από πλήρη πληροφόρηση.

Ορισμός 3.2.3.

$$R_n^\pi(\underline{\theta}) = nz^*(\underline{\theta}) - E_{\underline{\theta}}^\pi(S_n^\pi), \quad \underline{\theta} \in \Theta. \quad (3.2)$$

Παρατηρούμε ότι $\pi \in \Pi^C$ αν και μόνο αν $R_n^\pi(\underline{\theta}) = o(n)$ για όλα τα $\underline{\theta} \in \Theta$. Η συνάρτηση απώλειας ωστόσο είναι χρήσιμη επειδή εκτός από το να χαρακτηρίζει την

συνέπεια μπορεί να χρησιμοποιηθεί για να συγκρίνει συνεπείς πολιτικές ως προς την ταχύτητα σύγκλισης της $E_{\underline{\theta}}^{\pi}(S_n^{\pi})/n$ στο $z^*(\underline{\theta})$, αφού $R_n^{\pi}(\underline{\theta})/n = z^*(\underline{\theta}) - E_{\underline{\theta}}^{\pi}(S_n^{\pi})/n$.

Παρακάτω δίνουμε μια ισοδύναμη έκφραση η οποία συσχετίζει την απώλεια με την λύση του Π.Γ.Π. πλήρους πληροφόρησης. Ανφέρουμε πάλι ότι για κάθε βασικό πίνακα B ο οποίος αντιστοιχεί στην βέλτιστη λύση του Π.Γ.Π. (3.1), για το δυικό Π.Γ.Π. έχουμε ότι

$$z^*(\underline{\theta}) = C_0\lambda^B + g^B \text{ και } \mu_j(\theta_j) = c_j\lambda^B + g^B - \phi_j^B(\underline{\theta}), \quad j = 1, \dots, k.$$

Αντικαθιστώντας το $z^*(\underline{\theta})$ και το $\mu_j(\theta_j)$ στην (3.2)

$$\begin{aligned} R_n^{\pi}(\underline{\theta}) &= nz^*(\underline{\theta}) - \sum_{j=1}^k \mu_j(\theta_j) E_{\underline{\theta}}^{\pi}(T_n(j)) \\ &= \sum_{j=1}^k \phi_j^B(\underline{\theta}) E_{\underline{\theta}}^{\pi}(T_n(j)) + \lambda^B [nC_0 - E_{\underline{\theta}}^{\pi}(C_n^{\pi})] \\ &= \sum_{j=1}^k \phi_j^B(\underline{\theta}) E_{\underline{\theta}}^{\pi}(T_n(j)) + \lambda^B \sum_{j=1}^k (C_0 - c_j) E_{\underline{\theta}}^{\pi}(T_n(j)), \end{aligned} \quad (3.3)$$

για κάθε $\underline{\theta} \in \Theta$ και B βέλτιστο στο $\underline{\theta}$.

Παρατηρούμε ότι αν ο περιορισμός στο κόστος απαιτηθεί να κρατείται κατα μέση τιμή και ασυμπτωτικά ως προς n όπως στο προηγούμενο Κεφάλαιο τότε μπορούμε να κατασκευάσουμε συνεπείς πολιτικές έτσι ώστε το μέρος $nC_0 - E_{\underline{\theta}}^{\pi}(C_n^{\pi})$ και το $R_n^{\pi}(\underline{\theta})$, να παίρνουν αυθαίρετα μικρές αρνητικές τιμές για ένα πεπερασμένο αριθμό περιόδων. Ωστόσο στο παρόν πλαίσιο η συνάρτηση απώλειας μιας εφικτής πολιτικής είναι μη αρνητική για κάθε n , ιδιότητα που είναι λογική για συνάρτηση απώλειας.

Για να ισχυροποιήσουμε τις ιδιότητες σύγκλισης του π , ακολουθούμε την προσέγγιση που παρουσιάζεται στο [4] η οποία έχει χρησιμοποιηθεί εκτενώς στο πλαίσιο χωρίς περιορισμό (π.χ. [10], [14]).

Ορισμός 3.2.4. Μια πολιτική $\pi \in \Pi^F$ είναι ομοιόμορφα καλή αν

$$R_n^{\pi}(\underline{\theta}) = o(n^a), \quad n \rightarrow \infty, \quad \forall a > 0, \quad \forall \underline{\theta} \in \Theta.$$

Το πρώτο ερώτημα εδώ είναι αν υπάρχουν ομοιόμορφα καλές πολιτικές και το δεύτερο αν ανάμεσα στις ομοιόμορφα καλές πολιτικές υπάρχει κάποια βέλτιστη με κάποια έννοια. Σύμφωνα με την προσέγγιση του [4], μια τέτοια πολιτική καλείται αυμπτωτικά αποδοτική ή απλώς αποδοτική (efficient). Για να ταυτοποιήσουμε μια αποδοτική πολιτική, η κύρια ιδέα είναι να δείξουμε ότι η υποχρεωτική συνθήκη για να είναι μια πολιτική ομοιόμορφα καλή είναι η ύπαρξη μιας συνάρτησης $M(\underline{\theta})$ έτσι ώστε $\liminf R_n^\pi(\underline{\theta})/\log n \geq M(\underline{\theta})$ για κάθε $\underline{\theta}$, π.χ. ότι η απώλεια οποιασδήποτε ομοιόμορφα καλής πολιτικής αυξάνει με ρυθμό τουλάχιστον λογαριθμικό. Τότε αν κάποια πολιτική π^0 ικανοποιεί ότι $\limsup R_n^\pi(\underline{\theta})/\log n \leq M(\underline{\theta})$ για όλα τα $\underline{\theta}$ τότε αυτή η πολιτική είναι υποχρεωτικά αποδοτική.

Χρειαζόμαστε κάποιους επιπλέον ορισμούς. Αναφέρουμε πάλι ότι για $b \in K$, η b είναι βέλτιστη λύση του Π.Γ.Π. (3.1) για κάποια $\underline{\theta} \in \Theta$ αν και μόνο αν για τουλάχιστον έναν αντίστοιχο βασικό πίνακα B , $\phi_\alpha^B(\underline{\theta}) \geq 0$, $\alpha = 1, \dots, k$.

Για $b = \{i, j\} \in s(\underline{\theta})$, και $\alpha \neq i, j$ έστω

$$\Delta\Theta_\alpha(\underline{\theta}) = \{\theta'_\alpha \in \Theta_\alpha : s(\underline{\theta}') = \{\{i, \alpha\} \text{ ή } \{\alpha, j\} \text{ ή } \{\alpha\}\}\},$$

όπου $\underline{\theta}' = (\theta_1, \dots, \theta'_\alpha, \dots, \theta_k)$, είναι ένα νέο διάνυσμα τέτοιο ώστε μόνο η παράμετρος θ'_α έχει αλλάξει από θ_α . Το $\Delta\Theta$ είναι το σύνολο των τιμών του Θ_α κάτω από το οποίο το πρόβλημα πλήρους πληροφόρησης υπό την διαταραχή $\underline{\theta}'$ έχει μια μοναδική βέλτιστη ΒΕΛ η οποία περιέχει τον πληθυσμό α .

Τώρα, έστω

$$D(\underline{\theta}) = \{\alpha : \alpha \notin b \text{ για κάθε } b \in s(\underline{\theta}) \text{ και } \Delta\Theta_\alpha(\underline{\theta}) \neq \emptyset\},$$

να είναι οι πληθυσμοί οι οποίοι δεν περιέχονται σε καμμία βέλτιστη λύση κάτω από το σύνολο παραμέτρων $\underline{\theta}$ αλλά, αλλάζοντας μόνο την παράμετρο θ_α , μπορεί να βρεθεί μια μοναδικά βέλτιστη ΒΕΛ που τους περιέχει.

Λήμμα 3.2.1. Για κάθε $\underline{\theta}$, για κάθε βέλτιστο πίνακα B υπό το $\underline{\theta}$, $\exists \rho = \rho(\underline{\theta}, \alpha, B)$ έτσι ώστε για κάθε $\theta'_\alpha \in \Delta\Theta_\alpha(\underline{\theta})$

$$(i) \phi_j^B(\underline{\theta}') = \phi_j^B(\underline{\theta}) \geq 0, \forall j \neq \alpha \text{ και } \phi_\alpha^B(\underline{\theta}') = \phi_\alpha^B(\underline{\theta}) + \mu_\alpha(\theta_\alpha) - \mu_\alpha(\theta'_\alpha) < 0,$$

$$(ii) \mu_{\alpha, \inf}(\underline{\theta}) < \mu_\alpha(\theta'_\alpha) < \mu_{\alpha, \inf}(\underline{\theta}) + \rho, \text{ όπου } \rho > 0 \text{ και } \mu_{\alpha, \inf}(\underline{\theta}) = \phi_\alpha^B(\underline{\theta}) + \mu_\alpha(\theta_\alpha).$$

Απόδειξη. **(ι)** Είναι φανερό ότι το $\phi_j^B(\underline{\theta}') = \phi_j^B(\underline{\theta}) \geq 0$, $\forall j \neq \alpha$ επειδή αλλάξαμε μόνο την παράμετρο του πληθυσμού α και $\phi_j^B(\underline{\theta}') = \phi_j^B(\underline{\theta}) \equiv c_j \lambda^B + g^B - \mu_j(\theta_j)$.

Για έναν πληθυσμό $\alpha \in B(\underline{\theta})$ έχουμε ότι $\alpha \notin b$, για κάθε $b \in s(\underline{\theta})$. Οπότε $\phi_\alpha^B(\underline{\theta}) \equiv c_\alpha \lambda^B + g^B - \mu_\alpha(\theta_\alpha) > 0$, για κάθε B που αντιστοιχεί στο b .

Τώρα, κάθε βέλτιστη $b \in s(\underline{\theta})$ δεν είναι βέλτιστη υπό το $\underline{\theta}' = (\theta_1, \dots, \theta'_\alpha, \dots, \theta_k)$, για κάθε $\theta'_\alpha \in \Delta\Theta_\alpha(\underline{\theta})$, έτσι $s(\underline{\theta}') = \{b'\}$ όπου $b' \notin s(\underline{\theta})$.

Οπότε, για κάθε βέλτιστο πίνακα B υπό το $\underline{\theta}$ έχουμε ότι $\phi_\alpha^B(\underline{\theta}') \equiv c_\alpha \lambda^B + g^B - \mu_\alpha(\theta'_\alpha) < 0$ επειδή ο B δεν είναι βέλτιστος υπό το $\underline{\theta}'$.

Τώρα από $\phi_\alpha^B(\underline{\theta}) = c_\alpha \lambda^B + g^B - \mu_\alpha(\theta_\alpha)$ έχουμε ότι $\phi_\alpha^B(\underline{\theta}') = \phi_\alpha^B(\underline{\theta}) + \mu_\alpha(\theta_\alpha) - \mu_\alpha(\theta'_\alpha) < 0$.

(ιι) Υποθέτουμε χ.β.γ. ότι το $b = \{i, j\}$ είναι μια βέλτιστη λύση υπό το $\underline{\theta}$ με αντίστοιχο βέλτιστο πίνακα B και $b' = \{i, \alpha\}$ είναι μια βέλτιστη λύση υπό το $\underline{\theta}'$ με αντίστοιχο βέλτιστο πίνακα B' . Από το *i*) έχουμε ότι $z^*(\underline{\theta}') > z^*(\underline{\theta})$ αν και μόνο αν $\mu_\alpha(\theta'_\alpha) > \mu_{\alpha, \inf}(\theta)$.

Αφού η b' είναι μοναδικά βέλτιστη υπό το $\underline{\theta}'$ έχουμε ότι $\phi_s^{B'}(\underline{\theta}') > 0$, για κάθε $s \neq i, \alpha$. Τώρα για να ισχύει για αυτή τη κατάσταση χρησιμοποιούμε ότι $\phi_s^B(\underline{\theta}) > 0$ για κάθε $s \neq i, j$ και έχουμε ότι για $s > i$ αυτή ικανοποιεί ότι $\mu_{\alpha, \inf}(\theta) < \mu_\alpha(\theta'_\alpha)$, αλλά για $s < i$ πρέπει να έχουμε $\mu_{\alpha, \inf}(\theta) < \mu_\alpha(\theta'_\alpha) < \mu_{\alpha, \inf}(\theta) + \rho$, όπου ρ είναι μια θετική σταθερά. Έτσι, αν $\mu_{\alpha, \inf}(\theta) < \mu_\alpha(\theta'_\alpha) < \mu_{\alpha, \inf}(\theta) + \rho$ τότε $\phi_s^{B'}(\underline{\theta}') > 0$ για κάθε s . \square

Τώρα μπορούμε να ορίσουμε την ελάχιστη απόσταση, με την έννοια της πληροφορίας Kullback-Leibler, της παραμέτρου θ'_α από την θ_α έτσι ώστε να επιτυγχάνεται ότι ο πληθυσμός α γίνεται βέλτιστος υπό το θ'_α .

$$\begin{aligned} K_\alpha(\underline{\theta}) &= \inf\{I(\theta_\alpha, \theta'_\alpha) : \theta'_\alpha \in \Delta\Theta_\alpha(\underline{\theta})\} \\ &= \inf\{I(\theta_\alpha, \theta'_\alpha) : \theta'_\alpha \in \Theta_\alpha, \mu_{\alpha, \inf}(\theta) < \mu_\alpha(\theta'_\alpha) < \mu_{\alpha, \inf}(\theta) + \rho\}, \end{aligned}$$

όπου $\rho > 0$ και $I(\theta_\alpha, \theta'_\alpha)$ είναι η πληροφορία Kullback-Leibler, και ορίζεται ως

$$I(\theta_\alpha, \theta'_\alpha) = \int_{-\infty}^{+\infty} \log \frac{f(x; \theta_\alpha)}{f(x; \theta'_\alpha)} f(x; \theta_\alpha) dv(x).$$

Παρατηρούμε ότι στην εκθετική οικογένεια ο αριθμός Kullback-Leibler δίνεται από

$$I(\theta, \lambda) = (\theta - \lambda)\psi'(\theta) - (\psi(\theta) - \psi(\lambda)) = \int_{\theta}^{\lambda} (\lambda - t)\psi''(t)dt,$$

έτσι για ένα σταθερό θ , το I είναι αύξουσα ως προς λ , για $\lambda > \theta$ και φθίνουσα για $\lambda < \theta$.

3.3 Ένα κάτω φράγμα για την συνάρτηση απώλειας

Σε αυτή την ενότητα εδραιώνουμε ένα ασυμπτωτικό κάτω φράγμα για την απώλεια μιας ομοιόμορφα καλής πολιτικής.

Από (3.3) και τα δύο μέλη είναι μη αρνητικά, το πρώτο λόγω βελτιστότητας και το δεύτερο λόγω εφικτότητας. Οπότε ακολουθεί ότι μια αναγκαία και επαρκής συνθήκη για μια ομοιόμορφα καλή πολιτική π είναι ότι για $\underline{\theta} \in \Theta$ και κάθε βέλτιστη ΒΕΛ b υπό το $\underline{\theta}$,

$$\phi_j^B(\underline{\theta}) \lim_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^{\pi}(T_n(j))}{n^a} = 0, \text{ για όλα τα } a > 0, j \neq b, \quad (3.4)$$

και επίσης,

$$\lambda^B \lim_{n \rightarrow \infty} \frac{\sum_{j \in b}(C_0 - c_j)E_{\underline{\theta}}^{\pi}(T_n(j))}{n^a} = 0, \text{ για όλους τους } B \text{ που αντιστοιχούν στο } b. \quad (3.5)$$

Χρειαζόμαστε το ακόλουθο Λήμμα και την Πρόταση για να ταυτοποιήσουμε το κάτω φράγμα για την απώλεια.

Λήμμα 3.3.1. Έστω b μοναδικά βέλτιστη ΒΕΛ και B οποιοσδήποτε βέλτιστος πίνακας υπό το $\underline{\theta}$. Τότε

$$(ι) \text{ αν } B = \begin{pmatrix} c_i & c_j \\ 1 & 1 \end{pmatrix}, \text{ για κάποιο } i \leq d < j \Rightarrow \lambda^B > 0,$$

$$(ιι) \text{ αν } B = \begin{pmatrix} c_i & 1 \\ 1 & 0 \end{pmatrix}, \text{ για κάποιο } i \leq d \Rightarrow \lambda^B = 0.$$

Απόδειξη. **(i)** Έστω $\underline{\theta} : s(\underline{\theta}) = \{b\}$, $b = (i, j)$ για $i \leq d < j$, τότε $\lambda^B > 0$ επειδή αν $\lambda^B = 0$ πρέπει να έχουμε περισσότερες από μια λύσεις στο πρωτεύον, το οποίο δεν μπορεί να συμβαίνει επειδή η b είναι μοναδικά βέλτιστη.

(ii) Έστω $\underline{\theta} : s(\underline{\theta}) = \{b\}$, $b = (i)$ για $i \leq d$, τότε $\lambda^B = 0$ από τη δυική λύση και $\phi_j^B(\underline{\theta}) > 0$ για όλα τα $j \neq i$. \square

Πρόταση 3.3.1. Για κάθε ομοιόμορφα καλή πολιτική π και για κάθε $\underline{\theta} \in \Theta$ έχουμε ότι για $\alpha \in D(\underline{\theta})$, κάθε $\underline{\theta}' \in \Delta(\underline{\theta})$ και για κάθε $\beta_n = o(n)$ ισχύει

$$P_{\underline{\theta}'}^\pi[T_n(\alpha) < \beta_n] = o(n^{a-1}), \text{ για κάθε } a > 0.$$

Απόδειξη. Έστω $\alpha \in D(\underline{\theta})$, $\underline{\theta}'_\alpha \in \Delta\Theta_\alpha(\underline{\theta})$, b' μοναδικά βέλτιστη υπό το $\underline{\theta}'$ ($s(\underline{\theta}') = \{b'\}$) και $\alpha \in b'$. Τότε έχουμε δύο περιπτώσεις για την μοναδικά βέλτιστη λύση b' .

Στην πρώτη περίπτωση όπου $b' = \{\alpha\}$ αν b' είναι μη εκφυλισμένη τότε ο βασικός πίνακας είναι $B' = \begin{pmatrix} c_\alpha & 1 \\ 1 & 0 \end{pmatrix}$ και από το Λήμμα 3.3.1 για μια ομοιόμορφα καλή πολιτική $\lambda^B = 0$ έτσι,

$$E_{\underline{\theta}'}^\pi(T_n(j)) = o(n^a), \text{ για κάθε } a > 0, \text{ για κάθε } j \neq b'.$$

Αν b' είναι εκφυλισμένη τότε πρέπει να αληθεύει ότι $c_\alpha = C_0$. Αν θεωρήσουμε κάθε πίνακα $B' = \begin{pmatrix} c_\alpha & c_j \\ 1 & 1 \end{pmatrix}$ τότε $\lambda_{B'} > 0$, έτσι $(C_0 - c_j)E_{\underline{\theta}'}^\pi(T_n(j)) + (C_0 - c_\alpha)E_{\underline{\theta}'}^\pi(T_n(\alpha)) = o(n^a)$ και αφού $C_0 = c_\alpha$ έχουμε ότι $E_{\underline{\theta}'}^\pi(T_n(j)) = o(n^a)$. Επίσης από (3.4) $E_{\underline{\theta}'}^\pi(T_n(i)) = o(n^a)$, για κάθε $i \neq j, \alpha$ έτσι $E_{\underline{\theta}'}^\pi(T_n(j)) = o(n^a)$, για κάθε $j \neq \alpha$.

Επομένως,

$$n - E_{\underline{\theta}'}^\pi(T_n(\alpha)) = o(n^a), \text{ για κάθε } a > 0. \quad (3.6)$$

Αληθεύει επίσης ότι

$$\begin{aligned} E_{\underline{\theta}'}^\pi(T_n(\alpha)) &= \sum_{k=1}^n k \cdot P_{\underline{\theta}'}^\pi[T_n(\alpha) = k] \\ &= \sum_{k=1}^{\lfloor \beta_n \rfloor} k \cdot P_{\underline{\theta}'}^\pi[T_n(\alpha) = k] + \sum_{k=\lfloor \beta_n \rfloor + 1}^n k \cdot P_{\underline{\theta}'}^\pi[T_n(\alpha) = k] \\ &\leq \beta_n P_{\underline{\theta}'}^\pi[T_n(\alpha) \leq \beta_n] + n P_{\underline{\theta}'}^\pi[T_n(\alpha) > \beta_n] \\ &= n - (n - \beta_n) P_{\underline{\theta}'}^\pi[T_n(\alpha) \leq \beta_n]. \end{aligned}$$

Συνεπώς

$$n - E_{\underline{\theta}'}^{\pi}(T_n(\alpha)) \geq (n - \beta_n)P_{\underline{\theta}'}^{\pi}[T_n(\alpha) \leq \beta_n]. \quad (3.7)$$

Από (3.6) και (3.7) δείχνουμε ότι

$$(n - \beta_n)P_{\underline{\theta}'}^{\pi}[T_n(\alpha) \leq \beta_n] = o(n^a), \text{ για κάθε } a > 0,$$

έτσι

$$P_{\underline{\theta}'}^{\pi}[T_n(\alpha) \leq \beta_n] = o(n^{a-1}), \text{ για κάθε } a > 0.$$

Στην δεύτερη περίπτωση χωρίς βλάβη της γενικότητας $c_{\alpha} > C_0$ και $b' = \{j_0, \alpha\}$ έχουμε από το Λήμμα 3.3.1 ότι για μια ομοιόμορφα καλή πολιτική $\lambda^B > 0$, έτσι

$$E_{\underline{\theta}'}^{\pi}(T_n(j)) = o(n^a), \forall a > 0, \forall j \notin b' = \{j_0, \alpha\} \quad (3.8)$$

και

$$(C_0 - c_{j_0})E_{\underline{\theta}'}^{\pi}(T_n(j_0)) + (C_0 - c_{\alpha})E_{\underline{\theta}'}^{\pi}(T_n(\alpha)) = o(n^a), \forall a > 0. \quad (3.9)$$

Αν προσθέσουμε την (3.8) για όλα τα $j \neq \alpha, j_0$ ακολουθεί ότι

$$n - E_{\underline{\theta}'}^{\pi}(T_n(j_0)) - E_{\underline{\theta}'}^{\pi}(T_n(\alpha)) = \varepsilon_n, \text{ όπου } \varepsilon_n = o(n^a), \forall a > 0. \quad (3.10)$$

Διαιρώντας την (3.9) με $c_{\alpha} - c_{j_0}$ και χρησιμοποιώντας την (3.10), δείχνουμε μετά από λίγη άλγεβρα τις ακόλουθες δύο ισότητες

$$\begin{aligned} nx'_{j_0} - E_{\underline{\theta}'}^{\pi}(T_n(j_0)) &= o(n^a), \\ nx'_{\alpha} - E_{\underline{\theta}'}^{\pi}(T_n(\alpha)) &= o(n^a), \forall a > 0. \end{aligned} \quad (3.11)$$

όπου $x'_{j_0} = \frac{c_{\alpha} - C_0}{c_{\alpha} - c_{j_0}}$ και $x'_{\alpha} = \frac{C_0 - c_{j_0}}{c_{\alpha} - c_{j_0}}$ είναι οι πιθανότητες που αντιστοιχούν στη βέλτιστη λύση b' Π.Γ.Π. (3.1) υπό το $\underline{\theta}'$.

Για κάθε n έστω

$$\Gamma_n = \sum_{j \neq \alpha, j_0} T_n(j), \text{ και } F_n = \sum_{j \neq \alpha, j_0} (C_0 - c_j)T_n(j).$$

Έτσι, είναι φανερό ότι

$$F_n \leq \Gamma_n(C_0 - c_1).$$

Επιπλέον, από την (3.10)

$$E_{\theta'}^{\pi}(\Gamma_n) = o(n^a), \quad \forall a > 0. \quad (3.12)$$

τώρα, ξέρουμε ότι

$$nC_0 - C_n = F_n + (C_0 - c_{\alpha})T_n(\alpha) + (C_0 - c_{j_0})T_n(j_0),$$

και από $nC_0 - C_n \geq 0$, $\forall n$, έχουμε ότι

$$(c_{\alpha} - C_0)T_n(\alpha) \leq F_n + (C_0 - c_{j_0})T_n(j_0),$$

οπότε

$$\begin{aligned} \frac{c_{\alpha} - C_0}{c_{\alpha} - c_{j_0}} T_n(\alpha) &\leq \frac{F_n}{c_{\alpha} - c_{j_0}} + \frac{C_0 - c_{j_0}}{c_{\alpha} - c_{j_0}} T_n(j_0) \\ x'_{j_0} T_n(\alpha) &\leq \frac{F_n}{c_{\alpha} - c_{j_0}} + x'_{\alpha} T_n(j_0) \\ (1 - x'_{\alpha}) T_n(\alpha) &\leq \frac{F_n}{c_{\alpha} - c_{j_0}} + x'_{\alpha} T_n(j_0) \\ T_n(\alpha) &\leq \frac{F_n}{c_{\alpha} - c_{j_0}} + x'_{\alpha} (T_n(\alpha) + T_n(j_0)) \\ T_n(\alpha) &\leq \frac{F_n}{c_{\alpha} - c_{j_0}} + x'_{\alpha} (n - \Gamma_n) \\ T_n(\alpha) &\leq nx'_{\alpha} + \frac{F_n}{c_{\alpha} - c_{j_0}} - x'_{\alpha} \Gamma_n, \end{aligned}$$

και από την $F_n \leq \Gamma_n(C_0 - c_1)$

$$\begin{aligned} T_n(\alpha) &\leq nx'_{\alpha} + \frac{\Gamma_n(C_0 - c_1)}{c_{\alpha} - c_{j_0}} - x'_{\alpha} \Gamma_n \\ T_n(\alpha) &\leq nx'_{\alpha} + \Gamma_n \rho(j_0, \alpha) \end{aligned}$$

όπου $\rho(j_0, \alpha) = \frac{c_{j_0} - c_1}{c_\alpha - c_{j_0}} \geq 0$.

Τελικά,

$$nx'_\alpha - T_n(\alpha) + \Gamma_n \rho(j_0, \alpha) \geq 0. \quad (3.13)$$

Έτσι, από την ανισότητα *Markov*, για κάθε $\beta_n = o(n)$

$$\begin{aligned} P_{\underline{\theta}}^\pi(nx'_\alpha - T_n(\alpha) + \Gamma_n \rho(j_0, \alpha) \geq nx'_\alpha - \beta_n) \\ &\leq \frac{E_{\underline{\theta}}^\pi(nx'_\alpha - T_n(\alpha) + \Gamma_n \rho(j_0, \alpha))}{nx'_\alpha - \beta_n} \\ &= \frac{o(n^a)}{nx'_\alpha - \beta_n} = o(n^{a-1}), \quad \forall a > 0. \end{aligned}$$

Επομένως

$$P_{\underline{\theta}}^\pi(T_n(\alpha) \leq \beta_n) \leq P_{\underline{\theta}}^\pi(T_n(\alpha) \leq \beta_n + \Gamma_n \rho(j_0, \alpha)) = o(n^{a-1}), \quad \forall a > 0.$$

Αντικαθιστώντας το $T_n(\alpha) = n - \Gamma_n - T_n(j_0)$ στην (3.13) έχουμε

$$T_n(j_0) - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n \geq 0,$$

τότε

$$\begin{aligned} P_{\underline{\theta}}^\pi(T_n(j_0) \leq \beta_n) &= P_{\underline{\theta}}^\pi(T_n(j_0) - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n \\ &\leq \beta_n - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n) \\ &= P_{\underline{\theta}}^\pi(Z_n \leq \beta_n - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n), \end{aligned}$$

όπου

$$Z_n = T_n(j_0) - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n \geq 0,$$

και

$$E_{\underline{\theta}}^\pi(Z_n) = o(n^a), \quad \forall a > 0 \text{ από (3.11) και (3.12).}$$

Έστω,

$$U_n = \{Z_n \leq \beta_n - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n\}, \text{ τότε}$$

$$\begin{aligned} P_{\underline{\theta}'}^\pi(U_n) &= P_{\underline{\theta}'}^\pi(U_n \cap \{\Gamma_n \leq n\delta\}) + P_{\underline{\theta}'}^\pi(U_n \cap \{\Gamma_n > n\delta\}) \\ &\leq P_{\underline{\theta}'}^\pi(U_n \cap \{\Gamma_n \leq n\delta\}) + P_{\underline{\theta}'}^\pi(\Gamma_n > n\delta) \end{aligned} \quad (3.14)$$

όπου $0 < \delta < \frac{x'_{j_0}}{1 + \rho(j_0, \alpha)}$ έτσι,

$$\begin{aligned} P_{\underline{\theta}'}^\pi(\Gamma_n > n\delta) &\leq \frac{E_{\underline{\theta}'}^\pi(\Gamma_n)}{n\delta} \\ \text{από (3.12)} &= \frac{o(n^a)}{n\delta} = o(n^{a-1}), \forall a > 0. \end{aligned} \quad (3.15)$$

Έστω,

$$\begin{aligned} E_n &= \{U_n \cap \{\Gamma_n \leq n\delta\}\} \\ &= \{Z_n \leq \beta_n - nx'_{j_0} + (1 + \rho(j_0, \alpha))\Gamma_n \text{ και } \Gamma_n \leq n\delta\} \\ &\subseteq \{Z_n \leq \beta_n + [(1 + \rho(j_0, \alpha))\delta - x'_{j_0}]n\}, \\ &= \{Z_n \leq \beta_n - \phi n\}, \text{ όπου,} \end{aligned}$$

$$\phi = x'_{j_0} - (1 + \rho(j_0, \alpha))\delta > x'_{j_0} - (1 + \rho(j_0, \alpha))\frac{x'_{j_0}}{1 + \rho(j_0, \alpha)} = 0.$$

Τώρα για κάθε $\beta_n = o(n)$,

$$\exists n_0 : \beta_n - n\phi < 0 \forall n > n_0$$

και έχουμε ότι

$$P_{\underline{\theta}'}^\pi(E_n) = 0, \forall n > n_0(\phi),$$

έτσι από (3.14),(3.15)

$$P_{\underline{\theta}}^{\pi}(U_n) \leq o(n^{a-1}), \forall a > 0.$$

Τελικά,

$$P_{\underline{\theta}}^{\pi}(T_n(j_0) \leq \beta_n) = o(n^{a-1}), \forall a > 0, \text{ για κάθε } \beta_n = o(n).$$

Μέχρι στιγμής έχουμε δείξει ότι μια αναγκαία συνθήκη για μια ομοιόμορφα καλή πολιτική είναι ότι $\forall \underline{\theta} \in \underline{\Theta}$, και $\forall \alpha \in D(\underline{\theta})$ πρέπει να ισχύει ότι οι ποσότητες $T_n(j_0)$, $T_n(\alpha)$ δεν μπορεί να είναι πολύ μικρές, με την έννοια ότι $P_{\underline{\theta}}^{\pi}(T_n(j_0) \leq \beta_n) = o(n^{a-1})$, $P_{\underline{\theta}}^{\pi}(T_n(\alpha) \leq \beta_n) = o(n^{a-1})$ για κάθε θετική ακολουθία από σταθερές $\beta_n = o(n)$. \square

Λήμμα 3.3.2. Αν $P_{\underline{\theta}}^{\pi}[T_n(\alpha) < \beta_n] = o(n^{a-1})$, για κάθε $a > 0$ και $\beta_n = o(n)$ τότε

$$\lim_{n \rightarrow \infty} P_{\underline{\theta}}^{\pi}[T_n(\alpha) < \frac{\log n}{K_{\alpha}(\underline{\theta})}],$$

για κάθε $\underline{\theta} \in \Theta$ και $\alpha \in \Delta(\underline{\theta})$.

Απόδειξη. Αν πάρουμε $\beta_n = \frac{\log n}{K_{\alpha}(\underline{\theta})}$ τότε $P_{\underline{\theta}}^{\pi}[T_n(\alpha) < \frac{\log n}{K_{\alpha}(\underline{\theta})}] = o(n^{a-1})$ και χρησιμοποιούμε μια αλλαγή στο μέτρο από $\underline{\theta}'$ σε $\underline{\theta}$ και ακολουθούμε τα επιχειρήματα των [4] και έχουμε ότι

$$\lim_{n \rightarrow \infty} P_{\underline{\theta}}^{\pi}[T_n(\alpha) < \frac{\log n}{K_{\alpha}(\underline{\theta})}].$$

\square

Τέλος, μπορούμε να δείξουμε από το Λήμμα 3.3.2 το ακόλουθο Θεώρημα για την συνάρτηση απώλειας.

Θεώρημα 3.3.1. Αν η π είναι ομοιόμορφα καλή πολιτική τότε

$$\liminf_{n \rightarrow \infty} \frac{R_n^{\pi}(\underline{\theta})}{\log n} \geq \sum_{j \in D(\underline{\theta})} \frac{\phi_j^B(\underline{\theta})}{K_j(\underline{\theta})}, \text{ για κάθε } \underline{\theta} \in \Theta.$$

Απόδειξη. Υπενθυμίζουμε την έκφραση της απώλειας,

$$R_n^\pi(\underline{\theta}) = \sum_{j=1}^k \phi_j^B(\underline{\theta}) E_{\underline{\theta}}^\pi(T_n(j)) + \lambda^B [nC_0 - E_{\underline{\theta}}^\pi(C_n^\pi)],$$

και από την προηγούμενη Πρόταση αποδείξαμε ότι

$$\liminf_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^\pi(T_n(j))}{\log n} \geq \frac{1}{K_j(\underline{\theta})}, \text{ για κάθε } j \in D(\underline{\theta}), \text{ } \pi \text{ ομοιόμορφα καλή, και } \underline{\theta} \in \Theta.$$

Επίσης, έχουμε από το δεικτικό ότι $\lambda^B \geq 0$ και από την εφικτότητα ότι $nC_0 - E_{\underline{\theta}}^\pi(C_n^\pi) \geq 0$, για κάθε n . Τελικά, έχουμε ότι οι βέλτιστοι πληθυσμοί υπό το $\underline{\theta}$ έχουν $\phi_j^B(\underline{\theta}) = 0$, έτσι

$$\liminf_{n \rightarrow \infty} \frac{R_n^\pi(\underline{\theta})}{\log n} \geq \sum_{j \in D(\underline{\theta})} \frac{\phi_j^B(\underline{\theta})}{K_j(\underline{\theta})}, \text{ για κάθε } \underline{\theta} \in \Theta.$$

□

3.4 Κατασκευή μιας ασυμπτωτικά αποδοτικής πολιτικής

Σε αυτή την Ενότητα θέλουμε να κατασκευάσουμε μια αποδοτική πολιτική η οποία να επιτυγχάνει το κάτω φράγμα της απώλειας. Για να επιτύχουμε τη σχέση

$$\limsup_{n \rightarrow \infty} \frac{R_n^\pi(\underline{\theta})}{\log n} \leq \sum_{j \in D(\underline{\theta})} \frac{\phi_j^B(\underline{\theta})}{K_j(\underline{\theta})}, \text{ για κάθε } \underline{\theta} \in \Theta,$$

είναι αρκετό να κατασκευάσουμε μια πολιτική π , τέτοια ώστε

$$1. \quad \limsup_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^\pi(T_n(j))}{\log n} \leq \frac{1}{K_j(\underline{\theta})}, \text{ για κάθε } j \in D(\underline{\theta}) \quad (3.16)$$

$$2. \quad \limsup_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^\pi(T_n(j))}{\log n} = 0, \text{ για κάθε } j \notin D(\underline{\theta}) \quad (3.17)$$

$$3. \quad nC_0 - E_{\underline{\theta}}^\pi(C_n^\pi) = o(\log n). \quad (3.18)$$

Πρίν ορίσουμε τη πολιτική μας η οποία ικανοποιεί τις παραπάνω σχέσεις, χρειαζόμαστε μερικούς ορισμούς. Θεωρούμε μια κλάση πολιτικών τέτοια ώστε η δειγματοληψία

να πραγματοποιείται σε ομάδες από υπακολουθίες περιόδων οι οποίες λέγονται μπλόκ δειγματοληψίας, πεπερασμένου μήκους. Κάθε μπλόκ είναι ‘αυτοχρηματοδοτούμενο’ το οποίο σημαίνει ότι οι δειγματοληψίες εντός του μπλόκ πραγματοποιούνται με τέτοια σειρά ώστε ποτέ το κόστος εντός ενός μπλόκ να μην ξεπερνά το μηδέν. Οπότε το διαθέσιμο ποσό για δειγματοληψία στο τέλος ενός μπλόκ είναι μεγαλύτερο μεγαλύτερο ή ίσο με το διαθέσιμο ποσό για δειγματοληψία που είχαμε στην αρχή του μπλόκ.

Ορίζουμε τις διαφορές

$$\delta_i \equiv c_i - C_0.$$

Η ποσότητα δ_i εκφράζει την επίδραση μιας παρατήρησης από ένα απλό πληθυσμό i στο διαθέσιμο ποσό για δειγματοληψία. Αυτή η επίδραση είναι κόστος αν $\delta_i > 0$ και όφελος (αποταμίευση) αν $\delta_i < 0$.

Το πραγματικό πρόβλημα είναι ισοδύναμο με το μετασχηματισμένο πρόβλημα όπου $c_i = \delta_i$, $i = 1, \dots, k$, $C_0 = 0$ και ο περιορισμός στην δειγματοληψία είναι

$$\frac{1}{n} \sum_{t=1}^n \delta_{A_t} \leq 0, \quad \forall n.$$

Αφού $\delta_i \in \mathbb{Q}$, $i = 1, \dots, k$ έτσι χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι $\delta_i \in \mathbb{Z}$, $i = 1, \dots, k$ αφού είναι πεπερασμένα στον αριθμό.

Έστω ότι το $J \subseteq \{1, \dots, k\}$ είναι το υποσύνολο των πληθυσμών που έχουν επιλεγεί μέσα σε ένα μπλόκ. Λόγω της ιδιότητας της αυτοχρηματοδότησης πρέπει να ισχύει ότι $\min_{j \in J} j \leq d$, αφού ένα μπλόκ δειγματοληψίας δεν μπορεί να επιλέγει μόνο από ‘ακριβούς’ πληθυσμούς, για τους οποίους $\delta_j > 0$.

Οι ‘φθηνοί’ πληθυσμοί στο J πρέπει να επιλέγονται αρκετά συχνά ώστε να χρηματοδοτούν την δειγματοληψία των ‘ακριβών’. Μαθηματικά αρκεί να βρούμε $\{m_j, j \in J\}$ έτσι ώστε κάθε πληθυσμός $j \in J$ να επιλέγεται m_j φορές, και $\sum_{j \in J} m_j \delta_j \leq 0$, $m_j \in \mathbb{N}$, $\forall j \in J$. Κάθε μπλόκ με m_j που ικανοποιεί τις προηγούμενες ιδιότητες καλείται αποδεκτό. Μια δυνατότητα είναι να θεωρήσουμε το μικρότερο μπλόκ, το οποίο θα είναι κατάλληλο στην περίπτωση ελλιπούς πληροφόρησης παρακάτω. Έτσι το ελάχιστο μήκος του μπλόκ δειγματοληψίας είναι η λύση του ακόλουθου Π.Γ.Π.

$$T_s^J = \min \sum_{j \in J} m_j$$

$$\sum_{j \in J} m_j \delta_j \leq 0$$

$$m_j \in \mathbb{N}, \forall j \in J.$$

Η ιδιότητα της αυτοχρηματοδότησης είναι ένας τρόπος να ενδυναμώσουμε το διαθέσιμο ποσό για δειγματοληψία λόγω του περιορισμού. Στο μοντέλο με το μέσο κόστος μια πολιτική επιλέγει σε κάθε περίοδο, σύμφωνα με την λύση του (3.1), έναν πληθυσμό για να κάνει δειγματοληψία. Εδώ στο τέλος ενός μπλόκ δειγματοληψίας ένα νέο μπλόκ επιλέγεται για να ξεκινήσει. Έτσι για να κατασκευάσουμε μια βέλτιστη πολιτική για τον αυστηρότερο περιορισμό αρκεί να βρούμε ένα σύνολο J ώστε να κατασκευάσουμε το μπλόκ δειγματοληψίας.

Μια βέλτιστη λύση του (3.1) δίνει τις πιθανότητες τυχαιοποίησης οι οποίες εγγυώνται την μεγιστοποίηση της μέσης επιστροφής υπό τον περιορισμό στο κόστος. Οι πληθυσμοί σε αυτή τη βέλτιστη λύση ορίζουν το σύνολο J και μπορούν να κατασκευάσουν το αντίστοιχο μπλόκ δειγματοληψίας. Τώρα αν ξεκινήσουμε με αυτό το μπλόκ δειγματοληψίας και στο τέλος του μπλόκ επαναλάβουμε το ίδιο μπλόκ δειγματοληψίας κτλ., έχουμε μια βέλτιστη πολιτική για το πρόβλημα μας κάτω από πλήρη πληροφόρηση.

Τώρα θεωρούμε δύο τύπους μπλόκ δειγματοληψίας, το Μπλόκ αραιών ακολουθιών (ΜΑΑ) και το μπλόκ γραμμικού προγραμματισμού (ΜΓΠ). Ένα ΜΑΑ είναι έτσι ώστε όλοι οι πληθυσμοί $\{1, \dots, k\}$ να επιλέγονται τουλάχιστον μια φορά με τον αλλά με τον μικρότερο δυνατό αριθμό φορών ώστε να εξασφαλίζουν την αυτοχρηματοδότηση. Ο σκοπός ενός ΜΑΑ είναι να πάρουμε εκτιμήσεις των $\mu_j(\theta_j)$ για όλους τους πληθυσμούς. Τα ΜΑΑ παίζουν τον ίδιο ρόλο με τις εξαναγκαστικές δειγματοληψίες του [8], έτσι δεν μπορούν να χρησιμοποιούνται συχνά. Και από το προηγούμενο Κεφάλαιο όπου μια συνεπής πολιτική κατασκευάστηκε για το πρόβλημα με τον περιορισμό στο μέσο κόστος, έχουμε ότι σκοπός ενός ΜΓΠ είναι να μιμηθεί την τυχαιοποίηση η οποία προτείνεται από το Ισοδύναμο Π.Γ.Π. με αυτοχρηματοδοτούμενο τρόπο. Για κάθε περίοδο n και για κάθε βέλτιστη ΒΕΛ $b \in s(\hat{\theta}_n)$ ορίζουμε το ΜΓΠ(b). Μια βέλτιστη ΒΕΛ μπορεί να είναι $b = \{i, j\}$ ή $b = \{i\}$, έτσι κάθε πληθυσμός στην b επιλέγεται κατά την διάρκεια το ΜΓΠ(b) έτσι ώστε οι συχνότητες να είναι ίσες με τις πιθανότητες τυχαιοποίησης

$$x_i = \frac{|\delta_j|}{|\delta_i| + |\delta_j|}, \quad x_j = \frac{|\delta_i|}{|\delta_i| + |\delta_j|}, \quad \text{αν } b = \{i, j\}$$

ή

$$x_i = 1 \text{ αν } b = \{i\}.$$

Μετά τον ορισμό των αυτοχρηματοδοτούμενων μπλόκ δειγματοληψίας δίνουμε μερικούς ακόμα ορισμούς για να κατασκευάσουμε την απώλεια των μπλόκ δειγματοληψίας. Η απώλεια των μπλόκ έχει τέτοια σχέση με την συνάρτηση απώλειας μας όπου μπορούμε να αποδείξουμε το άνω φράγμα για την αρχική μας συνάρτηση απώλειας μέσω της απώλειας των μπλόκ.

Υποθέτουμε ότι έχουμε l διαδοχικά μπλόκ τότε το $\tilde{T}_l(b)$ είναι ο αριθμός των ΜΓΠ(b) στα πρώτα l μπλόκ. Έτσι $\sum_{b \in K} \tilde{T}_l(b) = l$. Έστω ότι το N_l είναι το συνολικό μήκος των πρώτων l μπλόκ και L_n ο αριθμός των μπλόκ στις n περιόδους. Έτσι μπορούμε εύκολα να δείξουμε ότι

$$T_{N_l}(\alpha) = \sum_{b: \alpha \in b} m_\alpha^b \cdot \tilde{T}_l(b),$$

όπου το m_α^b είναι ο αριθμός των επιλογών από έναν πληθυσμό α μέσα σε ένα ΜΓΠ(b).

Τώρα μπορούμε να ορίσουμε την απώλεια των μπλόκ

$$\tilde{R}_l^\pi(\theta) = E_\theta^\pi N_l \cdot z^*(\theta) - E_\theta^\pi \sum_{j=1}^k \sum_{b \in K} \mu_j(\theta_j) \cdot m_j^b \cdot \tilde{T}_l(b).$$

Από τον ορισμό του $T_n(\alpha)$ μπορούμε να δούμε ότι

$$T_{N_{L_n}}(b) \leq T_n(\alpha) \leq T_{N_{L_n}}(b) + M_\alpha, \quad (3.19)$$

όπου το M_α είναι ο μέγιστος αριθμός των φορών όπου ο πληθυσμός α εμφανίζεται σε όλα τα μπλόκ. Έτσι έχουμε ότι την ακόλουθη σχέση για τους δύο τύπους απωλειών,

$$\begin{aligned} E_\theta^\pi \tilde{R}_{L_n}^\pi(\theta) + (n - E_\theta^\pi N_{L_n}) \cdot z^*(\theta) - \sum_{j=1}^k M_j \cdot \mu_j(\theta_j) \\ \leq R_n^\pi(\theta) \leq E_\theta^\pi \tilde{R}_{L_n}^\pi(\theta) + (n - E_\theta^\pi N_{L_n}) \cdot z^*(\theta). \end{aligned} \quad (3.20)$$

Επομένως, από τη σχέση του n με το L_n και από την (3.20) έχουμε την ακόλουθη σχέση μεταξύ των δύο απωλειών,

$$\limsup_{n \rightarrow \infty} \frac{R_n^\pi(\theta)}{\log n} = \limsup_{n \rightarrow \infty} \frac{\widetilde{R}_{L_n}^\pi(\theta)}{\log L_n}. \quad (3.21)$$

Είναι φανερό ότι αν θέλουμε να βρούμε μια πολιτική η οποία να επιτυγχάνει το κάτω φράγμα της απώλειας μπορούμε να βρούμε μια πολιτική που να επιτυγχάνει το κάτω φράγμα της απώλειας των μπλόκ.

Μπορούμε να ορίσουμε την πολιτική μας π^0 . Το πρώτο μπλόκ είναι ένα ΜΑΑ. Υποθέτουμε ότι στην αρχή του l μπλόκ έχουμε τις εκτιμήσεις $\hat{\theta}^l$ με $\mu_1(\hat{\theta}_1^l), \dots, \mu_k(\hat{\theta}_k^l)$. Στην αρχή του l μπλόκ έχουμε τις εκτιμήσεις των προηγούμενων $l - 1$ μπλόκ και παίρνουμε την λύση του Π.Γ.Π.

$$z^{b(\hat{\theta}^l)} = \max_{\tilde{b}_i(\hat{\theta}^l)} \{z^{\tilde{b}_i(\hat{\theta}^l)} : \widetilde{T}_l(\tilde{b}_i(\hat{\theta}^l)) \geq \delta(l-1)\}$$

όπου \tilde{b}_i είναι όλες οι ΒΕΛ στο K .

Τότε για όλα τα $\alpha = \{1, \dots, k\} \setminus \{\text{πληθυσμοί στο } b(\hat{\theta}^l)\}$, έστω

$$\tilde{\mu}_\alpha = \sup_{\theta'_\alpha} \{\mu_\alpha(\theta'_\alpha) : I(\hat{\theta}_\alpha^l, \theta'_\alpha) \leq \frac{\log N_{l-1}}{T_{N_{l-1}}(\alpha)}\},$$

και

$$\Phi_l^{(\hat{B}, \hat{\theta}^l)} = \{\alpha : \mu_{\alpha, \text{inf}}(\hat{\theta}^l) < \tilde{\mu}_\alpha < \mu_{\alpha, \text{inf}}(\hat{\theta}^l) + \rho(\hat{\theta}^l, \alpha, \hat{B})\}.$$

Έτσι, αν $\Phi_l^{(\hat{B}, \hat{\theta}^l)} = \emptyset$ τότε η πολιτική π^0 επιλέγει την $b(\hat{\theta}^l)$ ($\pi_l^0 = b(\hat{\theta}^l)$), διαφορετικά για όλα τα $\alpha \in \Phi_l^{(\hat{B}, \hat{\theta}^l)}$ ορίζουμε τον δείκτη

$$y_\alpha(\hat{\theta}^l, \theta'_\alpha) = \max_{\theta'_\alpha} \{z^{b_\alpha(\hat{\theta}^l, \theta'_\alpha)} : I(\hat{\theta}_\alpha^l, \theta'_\alpha) \leq \frac{\log N_{l-1}}{T_{N_{l-1}}(\alpha)}\},$$

όπου $b_\alpha(\hat{\theta}^l, \theta'_\alpha)$ είναι η εμφυσημένη λύση όταν αλλάζουμε μόνο την παράμετρο του πληθυσμού α , και η πολιτική π^0 επιλέγει $b_{\alpha^*}(\hat{\theta}^l, \theta'_{\alpha^*})$ το οποίο αντιστοιχεί στο $y_{\alpha^*}(\hat{\theta}^l) = \max_{\alpha \in \Phi_l^{(\hat{B}, \hat{\theta}^l)}} \{y_\alpha(\hat{\theta}^l, \theta'_\alpha)\}$, ($\pi_l^0 = b_{\alpha^*}(\hat{\theta}^l, \theta'_{\alpha^*})$).

Για την ανάλυση μας χρειαζόμαστε τον ορισμό του δείκτη $w_\alpha(\hat{\underline{\theta}}^l, \epsilon)$,

$$w_\alpha(\hat{\underline{\theta}}^l, \epsilon) = \inf_{\theta'_\alpha} \{I(\hat{\underline{\theta}}^l, \theta'_\alpha) : z^{b_\alpha(\hat{\underline{\theta}}^l, \theta'_\alpha)} > z^*(\underline{\theta}) - \epsilon\},$$

όπου $\alpha \in \Phi_l^{(\hat{B}, \hat{\theta}^l)}$. Έτσι, $y_\alpha(\hat{\underline{\theta}}^l, \theta'_\alpha) > z^*(\underline{\theta}) - \epsilon$ αν και μόνο αν $w_\alpha(\hat{\underline{\theta}}^l, \epsilon) < \frac{\log N_{l-1}}{T_{N_{l-1}}(\alpha)}$.

Τώρα υποθέτουμε χ.β.γ. έναν πληθυσμό $i \in D(\underline{\theta})$, $i \in b$ ($b \notin s(\underline{\theta})$) και $c_i < C_0$. Έστω

$$T_{N_{L_n}}(i) = \sum_{b:i \in b} m_i^b \cdot \tilde{T}_{L_n}(b).$$

Έτσι μπορούμε να αναλύσουμε το $\tilde{T}_{L_n}(b)$ όπως παρακάτω

$$\begin{aligned} \tilde{T}_{L_n}(b) &= \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \notin s(\underline{\theta})\} + \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta})\} \\ &\leq \sum_{t=1}^{L_n} 1\{b(\hat{\underline{\theta}}^t) \notin s(\underline{\theta})\} + \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta})\} \end{aligned} \quad (3.22)$$

και έστω $\tilde{T}_{L_n}^1(b)$, $\tilde{T}_{L_n}^2(b)$ να είναι αντίστοιχα τα αθροίσματα στην τελευταία ανισότητα. Τώρα μπορούμε να αποδείξουμε το κύριο Θεώρημα του Κεφαλαίου

Θεώρημα 3.4.1. Έστω η πολιτική π^0 , τότε

$$\limsup_{n \rightarrow \infty} \frac{R_n^{\pi^0}(\underline{\theta})}{\log n} \leq \sum_{j \in D(\underline{\theta})} \frac{\phi_j^B(\underline{\theta})}{K_j(\underline{\theta})}, \text{ για κάθε } \underline{\theta} \in \Theta.$$

Απόδειξη. Χρειάζεται να αποδείξουμε τις (3.16), (3.17) και (3.18). Από τις (3.19), (3.22) και τα Λήμματα 3.4.1 και 3.4.2 έχουμε αποδείξει τις σχέσεις (3.16) και (3.17). Η (3.18) προκύπτει από την εφικτότητα της π^0 και τις μπλόκ πολιτικές. \square

Λήμμα 3.4.1. Για την πολιτική π^0 ισχύει

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{E_{\hat{\underline{\theta}}}^{\pi^0}(\tilde{T}_{L_n}^2(b))}{\log L_n} &\leq \frac{1}{K_i(\underline{\theta})}, \text{ για κάθε } i \in D(\underline{\theta}), i \in b, b \notin s(\underline{\theta}) \text{ και} \\ \limsup_{n \rightarrow \infty} \frac{E_{\hat{\underline{\theta}}}^{\pi^0}(\tilde{T}_{L_n}^2(b))}{\log L_n} &= 0, \text{ για κάθε } i \notin D(\underline{\theta}), i \in b, b \in s(\underline{\theta}). \end{aligned}$$

Απόδειξη. Μπορούμε να αναλύσουμε το άθροισμα $\widetilde{T}_{L_n}^2(b)$ ως εξής

$$\begin{aligned}\widetilde{T}_{L_n}^2(b) &= \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t)\} \\ &= \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), y_i(\hat{\underline{\theta}}^t, \theta'_i) > z^*(\underline{\theta}) - \epsilon\} \\ &\quad + \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), y_i(\hat{\underline{\theta}}^t, \theta'_i) \leq z^*(\underline{\theta}) - \epsilon\}.\end{aligned}$$

Από την σχέση ανάμεσα στους δύο δείκτες y_i και w_i έχουμε ότι

$$\begin{aligned}&\sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), y_i(\hat{\underline{\theta}}^t, \theta'_i) > z^*(\underline{\theta}) - \epsilon\} \\ &\leq \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), w_i(\hat{\underline{\theta}}^t, \epsilon) < \frac{\log N_{t-1}}{T_{N_{t-1}}(i)}\} \\ &= \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), \\ &\quad w_i(\hat{\underline{\theta}}^t, \epsilon) < \frac{\log N_{t-1}}{T_{N_{t-1}}(i)}, w_i(\hat{\underline{\theta}}^t, \epsilon) > w_i(\underline{\theta}, \epsilon) - \delta\} \\ &\quad + \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), \\ &\quad w_i(\hat{\underline{\theta}}^t, \epsilon) < \frac{\log N_{t-1}}{T_{N_{t-1}}(i)}, w_i(\hat{\underline{\theta}}^t, \epsilon) \leq w_i(\underline{\theta}, \epsilon) - \delta\} \\ &\leq \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), T_{N_{t-1}}(i) < \frac{\log L_n}{w_i(\underline{\theta}, \epsilon) - \delta}\} \\ &\quad + \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), w_i(\hat{\underline{\theta}}^t, \epsilon) \leq w_i(\underline{\theta}, \epsilon) - \delta\}.\end{aligned}$$

Τώρα, το πρώτο άθροισμα στην τελευταία ανισότητα για $c = \frac{\log L_n}{w_i(\underline{\theta}, \epsilon) - \delta}$ και s ακέραιος

είναι ίσο με

$$\begin{aligned}
& \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), T_{N_{t-1}}(i) < c\} \\
& \leq \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, T_{N_{t-1}}(i) < c\} \\
& = \sum_{t=1}^{L_n} \sum_{s=0}^{\lfloor c/m_i^b \rfloor} 1\{\pi_t^0 = b, T_{N_{t-1}}(i) = s \cdot m_i^b\} \\
& = \sum_{s=0}^{\lfloor c/m_i^b \rfloor} \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, T_{N_{t-1}}(i) = s \cdot m_i^b\} \\
& \leq \lfloor c/m_i^b \rfloor + 1 \\
& \leq \frac{c}{m_i^b} + 1 = \frac{\log L_n}{m_i^b(w_i(\underline{\theta}, \epsilon) - \delta)} + 1.
\end{aligned}$$

Έτσι,

$$\begin{aligned}
& E_{\underline{\theta}}^{\pi^0} \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), T_{N_{t-1}}(i) < \frac{\log L_n}{w_i(\underline{\theta}, \epsilon) - \delta}\} \\
& \leq \frac{\log L_n}{m_i^b(w_i(\underline{\theta}, \epsilon) - \delta)} + 1. \tag{3.23}
\end{aligned}$$

Επιπλέον,

$$\begin{aligned}
& \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), w_i(\hat{\underline{\theta}}^t, \epsilon) \leq w_i(\underline{\theta}, \epsilon) - \delta\} \\
& \leq \sum_{t=1}^{L_n} 1\{b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), w_i(\hat{\underline{\theta}}^t, \epsilon) \leq w_i(\underline{\theta}, \epsilon) - \delta\} \\
& = \sum_{t=1}^{L_n} 1\{b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), |\hat{\theta}_i^t - \theta_i| > \xi\}, \xi > 0.
\end{aligned}$$

η τελευταία ισότητα ακολουθεί από την εκθετική οικογένεια και τους ορισμούς των w

$$\begin{aligned}
w_i(\hat{\underline{\theta}}^t, \epsilon) &= \inf_{\theta'_i} \{I(\hat{\underline{\theta}}^t, \theta'_i) : z^{b_i(\hat{\underline{\theta}}^t, \theta'_i)} > z^*(\underline{\theta}) - \epsilon\} \leq \\
w_i(\underline{\theta}, \epsilon) &= \inf_{\theta'_i} \{I(\underline{\theta}, \theta'_i) : z^{b_i(\underline{\theta}, \theta'_i)} > z^*(\underline{\theta}) - \epsilon\} - \delta,
\end{aligned}$$

όπου $w_i(\underline{\theta}, \epsilon)$ και $w_i(\hat{\underline{\theta}}^t, \epsilon)$ είναι συνεχείς συναρτήσεις ως προς $\underline{\theta}, \hat{\underline{\theta}}^t$. Έχουμε ότι $\hat{\theta}_i^t$ είναι ο μέσος ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με μέση τιμή θ_i έτσι

$$\begin{aligned} P_{\theta_i}^{\pi^0}(|\hat{\theta}_i^t - \theta_i| > \xi) &\leq P_{\theta_i}^{\pi^0}(|\hat{\theta}_i^l - \theta_i| > \xi, \text{ για κάποιο } l \leq t) \\ &\leq \sum_{l=1}^t P_{\theta_i}^{\pi^0}(|\hat{\theta}_i^l - \theta_i| > \xi) \\ &\leq \sum_{l=1}^t e^{-\gamma l} = o(1/t). \end{aligned}$$

Τότε

$$\begin{aligned} E_{\underline{\theta}}^{\pi^0} \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), w_i(\hat{\underline{\theta}}^t, \epsilon) \leq w_i(\underline{\theta}, \epsilon) - \delta\} \\ \leq o(\log L_n). \end{aligned} \quad (3.24)$$

Τώρα έστω $b(\hat{\underline{\theta}}^t) = (r, s)$ και προφανώς $b = (i, s)$,

$$\begin{aligned} \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), y_i(\hat{\underline{\theta}}^t, \theta'_i) \leq z^*(\underline{\theta}) - \epsilon\} \\ \leq \sum_{t=1}^{L_n} 1\{y_s(\hat{\underline{\theta}}^t, \theta'_s) \leq z^*(\underline{\theta}) - \epsilon\} \\ \leq \sum_{t=1}^{L_n} 1\{y_s(\hat{\underline{\theta}}^j, \theta'_s) \leq z^*(\underline{\theta}) - \epsilon, \text{ για κάποιο } j \leq N_{t-1}\} \\ = \sum_{t=1}^{L_n} 1\{|\hat{\theta}_s^j - \theta_s| > \xi, \text{ για κάποιο } j \leq N_{t-1}\}. \end{aligned}$$

Έτσι

$$\begin{aligned} E_{\underline{\theta}}^{\pi^0} \sum_{t=1}^{L_n} 1\{\pi_t^0 = b, b(\hat{\underline{\theta}}^t) \in s(\underline{\theta}), y_i(\hat{\underline{\theta}}^t, \theta'_i) = y_{\alpha^*}(\hat{\underline{\theta}}^t), y_i(\hat{\underline{\theta}}^t, \theta'_i) \leq z^*(\underline{\theta}) - \epsilon\} \\ \leq o(\log L_n), \end{aligned} \quad (3.25)$$

επειδή

$$\begin{aligned} P_{\theta_s}^{\pi^0}(|\hat{\theta}_s^j - \theta_s| > \xi, \text{ για κάποιο } j \leq t) \\ \leq \sum_{j=1}^t P_{\theta_s}^{\pi^0}(|\hat{\theta}_s^j - \theta_s| > \xi) = o(1/t), \end{aligned}$$

αφού η πολιτική π^0 σε κάθε μπλόκ t επιλέγει $b(\hat{\theta}^t) = (r, s)$ όταν $\tilde{T}_i(b(\hat{\theta}^t)) \geq \delta(t-1)$.

Τελικά, έχουμε από τις (3.23), (3.24) και (3.25) ότι

$$E_{\underline{\theta}}^{\pi^0} \tilde{T}_{L_n}^2(b) \leq \frac{\log L_n}{m_i^b(w_i(\underline{\theta}, \epsilon) - \delta)} + 1 + o(\log L_n) + o(\log L_n).$$

και

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} w_i(\underline{\theta}, \epsilon) &= K_i(\underline{\theta}), \text{ για } i \in D(\underline{\theta}) \text{ και} \\ \lim_{\epsilon \rightarrow 0} w_i(\underline{\theta}, \epsilon) &= \infty, \text{ για } i \notin D(\underline{\theta}), \end{aligned}$$

έτσι

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^{\pi^0}(\tilde{T}_{L_n}^2(b))}{\log L_n} &\leq \frac{1}{K_i(\underline{\theta})}, \text{ για κάθε } i \in D(\underline{\theta}), i \in b, b \notin s(\underline{\theta}) \text{ και} \\ \limsup_{n \rightarrow \infty} \frac{E_{\underline{\theta}}^{\pi^0}(\tilde{T}_{L_n}^2(b))}{\log L_n} &= 0, \text{ για κάθε } i \notin D(\underline{\theta}), i \in b, b \in s(\underline{\theta}). \end{aligned}$$

□

Τώρα, έστω $0 < \epsilon < \{z^*(\underline{\theta}) - \max_{b \notin s(\underline{\theta})} z^{b(\underline{\theta})}\}/2$ και c ένας θετικός ακέραιος, τότε ορίζουμε για $r = 0, 1, 2, \dots$

$$A_r = \bigcap_{1 \leq j \leq m} \left\{ \max_{\delta c^{r-1} \leq l-1 \leq c^{r+1}} |z^{\tilde{b}_j(\hat{\theta}^l)} - z^{\tilde{b}_j(\underline{\theta})}| \leq \epsilon \right\} \text{ και}$$

$$B_r = \bigcap_{\tilde{b}_j \in s(\underline{\theta})} \left\{ z^{\tilde{b}_j(\hat{\theta}^l)} \geq z^*(\underline{\theta}) - \epsilon, \text{ για κάθε } 1 \leq i \leq \delta(l-1) \text{ και } c^{r-1} \leq l-1 \leq c^{r+1} \right\},$$

όπου $0 < \delta < |K|$ είναι το ίδιο όπως στην π^0 .

Λήμμα 3.4.2. (ι) $P_{\underline{\theta}}^{\pi^0}(\bar{A}_r) = o(c^{-r})$, $P_{\underline{\theta}}^{\pi^0}(\bar{B}_r) = o(c^{-r})$.

Επιπλέον, αν $c > 1/(1 - |K|\delta)$ και $r \geq r_0$ τότε

(ii) στο $A_r \cap B_r$, $b(\hat{\theta}^l) \in s(\underline{\theta})$ για κάθε $c^{r-1} \leq l-1 \leq c^{r+1}$.

(iii) $E_{\underline{\theta}}^{\pi^0} \tilde{T}_{L_n}^1(b) = \sum_{t=1}^{L_n} P_{\underline{\theta}}^{\pi^0}(b(\hat{\theta}^t) \notin s(\underline{\theta})) = o(\log L_n)$.

Απόδειξη. Η απόδειξη είναι η ίδια όπως στο [4] στο Λήμμα 1. □

Τέλος όπως αναφέραμε προηγούμενα με βάση τα παραπάνω Λήμματα προκύπτει η απόδειξη του Θεωρήματος 3.4.1 και επομένως η πολιτική που κατασκευάστηκε είναι αποδοτική.

Κεφάλαιο 4

Πολιτικές παραγγελιών για δύο προϊόντα με ανταλλαξιμότητα στη ζήτηση

4.1 Εισαγωγή

Στο Κεφάλαιο αυτό θεωρούμε ένα πρόβλημα διαχείρισης αποθεμάτων με στοχαστική ζήτηση. Συγκεκριμένα θεωρούμε το μοντέλο αποθέματος με δύο προϊόντα με ανταλλαξιμότητα στην ζήτηση το οποίο σημαίνει ότι αν κάποιος από τα δύο προϊόντα δεν είναι διαθέσιμο τότε ο πελάτης μπορεί να επιλέξει το άλλο προϊόν με μια πιθανότητα. Επίσης υποθέτουμε ότι στο μοντέλο μας έχουμε μια μηχανή πωλητή η οποία λειτουργεί σε απλές περιόδους. Έτσι παραγγέλνουμε σταθερές ποσότητες για τα δύο προϊόντα στην αρχή κάθε κύκλου και τα απώλητα προϊόντα πρέπει να επιστραφούν στο τέλος του κύκλου. Στο Κεφάλαιο αυτό μελετάμε δύο διαφορετικούς τύπους του παραπάνω μοντέλου οι οποίοι έχουν να κάνουν με τον χρόνο ανεφοδιασμού. Στην πρώτη περίπτωση έχουμε ένα ντετερμινιστικό χρόνο ανεφοδιασμού όπου ο ανεφοδιαστής έρχεται σε σταθερές και προκαθορισμένες χρονικές περιόδους. Στην δεύτερη περίοδο υποθέτουμε ότι ο ανεφοδιαστής καταφθάνει σύμφωνα με μια διαδικασία Poisson με σταθερό ρυθμό.

Σε αυτή την περιοχή υπάρχουν πολλά άρθρα που έχουν βρει τη βέλτιστη πολιτική

παραγγελίας και απέδειξαν κάποιες σημαντικές ιδιότητες για την συνάρτηση ωφέλειας. Στο [24] γίνεται η υπόθεση ότι οι τιμές των απώλητων ποσοτήτων και των χαμένων πωλήσεων είναι μηδενικές και ότι η ανταλλαξιμότητα συμβαίνει σαν ποσοστό της ζήτησης που δεν ικανοποιείται. Εξέφρασαν την συνάρτηση οφέλους και έδειξαν ότι είναι αυστηρά κοίλη κάτω από μια βασική συνθήκη. Έτσι ανέλυσαν τις αναγκαίες συνθήκες βελτιστότητας και χρησιμοποίησαν μια επαναληπτική τεχνική όπως η μέθοδος Newton και βρήκαν τις βέλτιστες ποσότητες παραγγελίας.

Στο [22] επανεξετάστηκε το πρόβλημα αποθέματος με δύο ανταλλάξιμα προϊόντα. Επίσης υπέθεσαν ότι η ανταλλαξιμότητα εμφανίζεται σαν ποσοστό της ζήτησης. Υπέθεσαν θετικές τιμές για τα περισεύματα και τις χαμένες πωλήσεις και ανέλυσαν τη συνάρτηση ωφέλειας αλλά δεν κατάφεραν να δείξουν ότι είναι κοίλη. Έτσι βρήκαν άνω και κάτω φράγματα για τις βέλτιστες ποσότητες παραγγελίας και τις ταυτοποίησαν χρησιμοποιώντας προσομοίωση Monte Carlo.

Στο [26] βρήκαν τις αναγκαίες συνθήκες βελτιστότητας και παρείχαν την ερμηνεία των συνθηκών αυτών χρησιμοποιώντας επιχειρήματα ανάλυσης οριακού κέρδους. Επίσης απέδειξαν ότι είναι ικανές για την ειδική περίπτωση των ομοιόμορφων ζητήσεων και πρότειναν δύο ευρετικές (heuristic) λύσεις οι οποίες αφορούν ξεχωριστά την τόνωση της ζήτησης και την ανταλλαξιμότητα των προϊόντων και μας βοήθησε να καταλάβουμε καλύτερα την επιροή της χρήσης ανακριβών πληροφοριών για την ζήτηση όταν παίρνουμε αποφάσεις για αποθέματα.

Ένα παρόμοιο πρόβλημα με το δικό μας είναι το μοντέλο στο [25], όπου μελέτησαν το από κοινού ανεφοδιαστικό μοντέλο με ανταλλαξιμότητα με σκοπό να δείξουν ότι η ανταλλαξιμότητα μπορεί να βελτιώνει το παραδοσιακό μοντέλο κάτω από συγκεκριμένες συνθήκες τις οποίες εξέτασαν. Σε αυτή την εργασία έδειξαν ότι η ανταλλαξιμότητα ανάμεσα στα προϊόντα μπορεί να μειώσει το σταθερό κόστος παραγγελιών, το οποίο είναι ανεξάρτητο από τον αριθμό των προϊόντων που περιλαμβάνονται στην παραγγελία, και το κόστος αποθήκευσης το οποίο εξαρτάται από τον αριθμό των προϊόντων που έχουν στο απόθεμα. Τέλος, η ανάλυση έδειξε ότι το μοντέλο με ανταλλαξιμότητα μπορεί να έχει καλύτερα αποτελέσματα στην ουσιαστική ανάπτυξη του συνολικού κόστους σε σχέση με το αρχικό μοντέλο.

Ένα παρόμοιο μοντέλο παρουσιάζεται στο [23], όπου υπέθεσαν δύο παρόμοια ανταλλάξιμα προϊόντα τα οποία παραγγέλνονται ταυτόχρονα από τον ίδιο πωλητή σε κάθε

κύκλο ανεφοδιασμού. Στην περίπτωση που δεν υπάρχει απόθεμα στο ένα προϊόν, ένα γνωστό ποσοστό της ζήτησης του μπορεί να ικανοποιηθεί πηγαίνοντας στο άλλο προϊόν. Η ζήτηση για κάθε προϊόν εξαρτάται από την ποσότητα αποθέματος και των δύο προϊόντων σε μια ορισμένη χρονική στιγμή και τα αποθέματα των προϊόντων αδειάζουν ταυτόχρονα. Ο ανεφοδιαστής πρέπει να αποφασίσει για την ποσότητα της παραγγελίας του κάθε προϊόντος με σκοπό να βελτιστοποιήσει το όφελος του. Τέλος, σε αυτό το μοντέλο υπέθεσαν ότι όλες οι παράμετροι του κόστους και του κέρδους είναι διαφορετικές για τα δύο προϊόντα.

Όπως αναφέραμε προηγούμενα η βασική διαφορά στο δικό μας προτεινόμενο μοντέλο είναι ότι υποθέτουμε πιο ρεαλιστική διαδικασία για τις κατανομές της ζήτησης. Οι στοχαστικά αφικνούμενοι πελάτες στο πρόβλημά μας μας βοηθούν να μοντελοποιήσουμε πιο ρεαλιστικά το πρόβλημα αποθέματος με δύο ανταλλάξιμα προϊόντα. Στο πρόβλημα που μελετήθηκε εδώ έγινε η πιο ρεαλιστική υπόθεση ότι η ζήτηση για τα δύο προϊόντα ακολουθεί δύο ανεξάρτητες διαδικασίες Poisson. Με τον τρόπο αυτό ενσωματώνεται φυσιολογικά στην ανάλυση το σημαντικό χαρακτηριστικό του προβλήματος ότι οι ελλείψεις που δημιουργούνται εξαρτώνται όχι μόνο από τη συνολική ζήτηση των δύο προϊόντων αλλά και από τις συγκεκριμένες χρονικές στιγμές που έρχεται η ζήτηση. Σε αυτό το Κεφάλαιο ορίζουμε την συνάρτηση ωφέλειας για τις δύο περιπτώσεις ανεφοδιασμού και χρησιμοποιούμε επιχειρήματα σύζευξης (coupling) για να αποδείξουμε ότι είναι submodular συνάρτηση. Αυτή η ιδιότητα μας επιτρέπει να κατασκευάσουμε έναν αλγόριθμο για να βρούμε τις βέλτιστες ποσότητες παραγγελίας.

Το Κεφάλαιο έχει την εξής δομή. Στη δεύτερη Ενότητα παρουσιάζουμε τις δύο προσεγγίσεις σύμφωνα με τον τύπο ανεφοδιασμού. Στην τρίτη Ενότητα βρίσκουμε τη μεταβατική κατανομή του αριθμού των εναπομείναντων προϊόντων μετά από ντετερμινιστικό χρόνο και αποδεικνύουμε ότι η συνάρτηση ωφέλειας είναι submodular. Στην τέταρτη Ενότητα βρίσκουμε τη στάσιμη κατανομή του αριθμού των εναπομείναντων προϊόντων, όταν ο χρόνος ανεφοδιασμού είναι εκθετικά κατανεμημένος και δείχνουμε ότι η συνάρτηση ωφέλειας είναι submodular. Στην πέμπτη Ενότητα περιγράφουμε έναν αλγόριθμο ο οποίος βρίσκει τις βέλτιστες ποσότητες παραγγελιών. Τέλος στην έκτη Ενότητα έχουμε κάποια πολύ ενδιαφέροντα αριθμητικά αποτελέσματα.

4.2 Περιγραφή μοντέλου

Θεωρούμε έναν έμπορο ο οποίος παραγγέλνει δύο μερικώς ανταλλάξιμα προϊόντα στο περιβάλλον του εφημεριδοπώλη. Έστω r_i, c_i, h_i η εμπορική τιμή, η χονδρική τιμή και το κόστος ενός απώλητου προϊόντος, αντίστοιχα, για το προϊόν $i, i = 1, 2$. Παρατηρούμε ότι το h_i μπορεί να είναι αρνητικό αν στην πραγματικότητα είναι αξία των απώλητων ποσοτήτων. Παραγγελίες μεγέθους Q_1, Q_2 λαμβάνουν χώρα στην αρχή της περιόδου πωλήσεων. Υποθέτουμε ότι λόγω της περιορισμένης χωρητικότητας ή άλλων περιορισμών οι ποσότητες των παραγγελιών περιορίζονται από μια γραμμική ανισότητα $a_1 Q_1 + a_2 Q_2 \leq C$, όπου a_1, a_2, C μη αρνητικές σταθερές. Για παράδειγμα, για $a_1 = a_2 = 1$, η ανισότητα μπορεί να αντανακλά έναν περιορισμό στην χωρητικότητα, και για $a_i = c_i, i = 1, 2$, έναν περιορισμό στο διαθέσιμο ποσό για παραγγελίες.

Ανάμεσα στα δύο προϊόντα υπάρχει μια δίπλευρη ανταλλαξιμότητα. Ειδικότερα, υποθέτουμε ότι ένας πελάτης που ήθελε πραγματικά να αγοράσει το προϊόν i θα μεταπηδήσει στο προϊόν j με πιθανότητα $p_{ij}, i \neq j$, αν το αρχικό προϊόν δεν είναι διαθέσιμο. Οι πιθανότητες ανταλλαγής p_{12}, p_{21} είναι εξωγενείς παράμετροι.

Τώρα θεωρούμε τη διαδικασία της ζήτησης. Λόγω της πιθανότητας ανταλλαγής, οι πραγματικές πωλήσεις, οι απώλητες ποσότητες και ελλείψεις και από τα δύο προϊόντα, και σαν αποτέλεσμα τα οφέλη/κόστη δεν εξαρτώνται μόνο από την ζήτηση του κάθε προϊόντος, αλλά επίσης από τις ανεξάρτητες χρονικές αφίξεις των πελατών. Στην πραγματικότητα αυτός είναι ένας λόγος που περιπλέκει την αξιολόγηση και την βελτιστοποίηση των πολιτικών παραγγελίας, και ως τέτοια έχει προσεγγιστεί με διάφορους τρόπους στην βιβλιογραφία. Σε αυτό το Κεφάλαιο μοντελοποιούμε σαφέστατα τη διαδικασία της ζήτησης, υποθέτωντας ότι οι πελάτες που πραγματικά θέλουν να αγοράσουν ένα προϊόν i , φθάνουν σύμφωνα με μια διαδικασία Poisson με ρυθμό $\lambda_i, i = 1, 2$. Όταν η ζήτηση μιας μονάδας προϊόντος i έρχεται και δεν υπάρχει προϊόν i διαθέσιμο, ο πελάτης δέχεται να αγοράσει μια μονάδα προϊόντος $j, j \neq i$, με πιθανότητα $p_{ij}, i, j = 1, 2$. Το όφελος αυτής της μοντελοποίησης είναι ότι η επίδραση των χρονικών στιγμών των αφίξεων πάνω στη διαδικασία ανταλλαγής είναι ενδογενής και έτσι παρουσιάζεται ακριβέστερα.

Όσον αφορά την περίοδο ανεφοδιασμού, υποθέτουμε δύο προσεγγίσεις, οι οποίες οδηγούν σε διαφορετικά στοχαστικά μοντέλα. Στην πρώτη προσέγγιση υποθέτουμε ότι η διάρκεια είναι ντετερμινιστική και ίση με T . Αυτό αντιστοιχεί στην πιο συνηθισμένη

κατάσταση όπου ο ανεφοδιασμός συμβαίνει σε σταθερού μήκους διαστήματα, όπως σε ημερήσια ή εβδομαδιαία βάση. Στη δεύτερη προσέγγιση θεωρούμε την περίπτωση όπου ο χρόνος ανεφοδιασμού είναι εκθετικά κατανομημένος με παράμετρο μ . Η υπόθεση της εκθετικής κατανομής από την μια πλευρά μας επιτρέπει να μοντελοποιήσουμε την διαδικασία αποθέματος σαν ένα συνεχές προσιτό Μαρκοβιανό περιβάλλον. Από την άλλη μεριά περιγράφει καταστάσεις όπου η διαδικασία ανεφοδιασμού δεν είναι τόσο αξιόπιστη και περιστασιακά μπορεί να είναι πολύ καθυστερημένη.

Στόχος μας είναι να καθορίσουμε τις ποσότητες Q_1 και Q_2 οι οποίες μεγιστοποιούν το αναμενόμενο όφελος ανά μονάδα χρόνου. Για αυτό, αναπτύσσουμε ένα στοχαστικό μοντέλο της διαδικασίας αποθέματος για κάθε μια από τις προσεγγίσεις του χρόνου ανεφοδιασμού, και δίνουμε μια αναλυτική έκφραση για τη συνάρτηση ωφελιμότητας για κάθε μοντέλο και δείχνουμε ότι είναι submodular ως προς τις ποσότητες παραγγελιών. Η ιδιότητα του submodularity οδηγεί σε έναν γρηγορότερο αλγόριθμο υπολογισμού των βέλτιστων ποσοτήτων Q_1 και Q_2 .

4.2.1 Προσέγγιση 1

Όπως αναφέραμε προηγούμενα, σε αυτή την περίπτωση η διάρκεια μιας περιόδου είναι T . Η κατάσταση του αυτόματου πωλητή τη στιγμή t περιγράφεται από ένα ζεύγος $(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t))$, όπου $n_i^{(Q_1, Q_2)}(t)$ δηλώνει το απόθεμα του προϊόντος i τη στιγμή t , $i = 1, 2$, $t \in [0, T]$. Η στοχαστική διαδικασία $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$ είναι μια Μαρκοβιανή αλυσίδα συνεχούς χρόνου με χώρο καταστάσεων $S = \{(i_1, i_2) : i_1 \in \{0, 1, \dots, Q_1\}, i_2 \in \{0, 1, \dots, Q_2\}\}$ και γεννήτορα πίνακα

$$\mathbb{Q}^{(1)} = \begin{pmatrix} Q_0^{(1)} & 0 & 0 & \dots & 0 & 0 \\ Q_1^{(1)} & Q_2^{(1)} & 0 & \dots & 0 & 0 \\ 0 & Q_1^{(1)} & Q_2^{(1)} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & Q_2^{(1)} & 0 \\ 0 & 0 & 0 & \dots & Q_1^{(1)} & Q_2^{(1)} \end{pmatrix}, \quad (4.1)$$

$\mathbb{Q}^{(1)} \in M_{Q_1+1 \times Q_2+1, Q_1+1 \times Q_2+1}$, όπου

$$\mathbb{Q}_0^{(1)} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ s_2 & -s_2 & 0 & \dots & 0 & 0 \\ 0 & s_2 & -s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -s_2 & 0 \\ 0 & 0 & 0 & \dots & s_2 & -s_2 \end{pmatrix}, \quad (4.2)$$

$\mathbb{Q}_0^{(1)} \in M_{Q_2+1, Q_2+1}$, όπου $s_2 = \lambda_2 + \lambda_1 p_{12}$ και οι γραμμές και οι στήλες αντιστοιχούν στις καταστάσεις $(0, j)$, $j = 0, 1, \dots, Q_2$,

$$\mathbb{Q}_1^{(1)} = \begin{pmatrix} s_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_1 & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda_1 \end{pmatrix}, \quad (4.3)$$

$\mathbb{Q}_1^{(1)} \in M_{Q_2+1, Q_2+1}$, όπου $s_1 = \lambda_1 + \lambda_2 p_{21}$, και για σταθερό i , $i = 1, \dots, Q_1$ οι γραμμές αντιστοιχούν στις καταστάσεις (i, j) , $j = 0, 1, \dots, Q_2$ και οι στήλες αντιστοιχούν στις καταστάσεις $(i-1, j)$, $j = 0, 1, \dots, Q_2$,

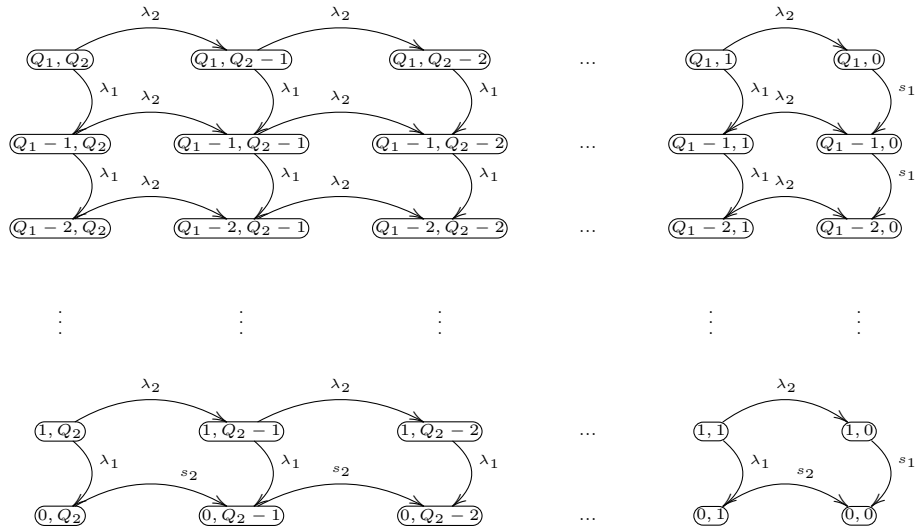
$$\mathbb{Q}_2^{(1)} = \begin{pmatrix} -s_1 & 0 & 0 & \dots & 0 & 0 \\ \lambda_2 & -s & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & -s & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -s & 0 \\ 0 & 0 & 0 & \dots & \lambda_2 & -s \end{pmatrix}, \quad (4.4)$$

$\mathbb{Q}_2^{(1)} \in M_{Q_2+1, Q_2+1}$, όπου $s = \lambda_1 + \lambda_2$, και για σταθερό i , $i = 1, \dots, Q_1$ οι γραμμές και οι στήλες αντιστοιχούν στις καταστάσεις (i, j) , $j = 0, 1, \dots, Q_2$,

$$\mathbb{O} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \quad (4.5)$$

$\mathbb{O} \in M_{Q_2+1, Q_2+1}$.

Οι ρυθμοί μετάβασης της Μαρκοβιανής αλυσίδας συνεχούς χρόνου δείχνονται επίσης στο παρακάτω διάγραμμα



Αφού στην αρχή κάθε περιόδου παραγγέλνουμε τις ποσότητες Q_1 και Q_2 από το προϊόν 1 και 2, αντίστοιχα, το αναμενόμενο όφελος ανά περίοδο για αυτή την προσέγγιση είναι

$$\begin{aligned} \pi^1(Q_1, Q_2) &= (r_1 - c_1)Q_1 + (r_2 - c_2)Q_2 \\ &\quad - (r_1 + h_1)En_1^{(Q_1, Q_2)}(T) - (r_2 + h_2)En_2^{(Q_1, Q_2)}(T). \end{aligned} \quad (4.6)$$

Έτσι, με σκοπό να βρούμε μια αναλυτική έκφραση για το $\pi^1(Q_1, Q_2)$, χρειάζεται να υπολογίσουμε την μεταβατική κατανομή της Μαρκοβιανής αλυσίδας συνεχούς χρόνου

$\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)), t \in [0, T]\}$ τη χρονική στιγμή T .

Στην Ενότητα 4.3 δίνουμε αναλυτικές εκφράσεις για την μεταβατική κατανομή. Μετά, δίνουμε το αναμενόμενο όφελος ανα περίοδο σαν συνάρτηση των Q_1 και Q_2 και τελικά αποδεικνύουμε ότι αυτή η συνάρτηση είναι submodular. Χρησιμοποιώντας την ιδιότητα του submodularity δίνουμε, στην Ενότητα 4.5, έναν αλγόριθμο ο οποίος προσδιορίζει τις βέλτιστες ποσότητες Q_1 και Q_2 .

4.2.2 Προσέγγιση 2

Στην δεύτερη προσέγγιση η διάρκεια της περιόδου είναι εκθετικά κατανομημένη με παράμετρο μ . Σε αυτή την περίπτωση η κατάσταση του αυτόματου πωλητή τη χρονική στιγμή t περιγράφεται από ένα ζεύγος $(m_1^{(Q_1, Q_2)}(t), m_2^{(Q_1, Q_2)}(t))$ όπου $m_i^{(Q_1, Q_2)}(t)$ δηλώνει το απόθεμα του προϊόντος i τη χρονική στιγμή t , $t \in \mathbb{R}$, $i \in \{1, 2\}$. Η стоχαστική διαδικασία $\{(m_1^{(Q_1, Q_2)}(t), m_2^{(Q_1, Q_2)}(t)), t \in \mathbb{R}\}$ είναι συνεχούς χρόνου Μαρκοβιανή αλυσίδα με χώρο καταστάσεων $S = \{(i_1, i_2) : i_1 \in \{0, 1, \dots, Q_1\}, i_2 \in \{0, 1, \dots, Q_2\}\}$ και γεννήτορα πίνακα

$$\mathbb{Q}^{(2)} = \begin{pmatrix} Q_0^{(2)} & 0 & 0 & \dots & 0 & \mathbb{M} \\ Q_1^{(2)} & Q_2^{(2)} & 0 & \dots & 0 & \mathbb{M} \\ 0 & Q_1^{(2)} & Q_2^{(2)} & \dots & 0 & \mathbb{M} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & Q_2^{(2)} & \mathbb{M} \\ 0 & 0 & 0 & \dots & Q_1^{(2)} & Q_2^{(2)} + \mathbb{M} \end{pmatrix}, \quad (4.7)$$

$\mathbb{Q}^{(2)} \in M_{Q_1+1 \times Q_2+1, Q_1+1 \times Q_2+1}$, όπου

$$\mathbb{M} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & \mu \\ 0 & 0 & 0 & \dots & 0 & \mu \\ 0 & 0 & 0 & \dots & 0 & \mu \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \mu \\ 0 & 0 & 0 & \dots & 0 & \mu \end{pmatrix}, \quad (4.8)$$

$\mathbb{M} \in M_{Q_2+1, Q_2+1}$,

$$Q_0^{(2)} = Q_0^{(1)} - \mu \mathbb{I}, \quad (4.9)$$

$$Q_1^{(2)} = Q_1^{(1)}, \quad (4.10)$$

$$Q_2^{(2)} = Q_2^{(1)} - \mu \mathbb{I}, \quad (4.11)$$

όπου $Q_0^{(1)}$, $Q_1^{(1)}$ και $Q_2^{(1)}$ δίνονται στις (4.2)-(4.4) και $\mathbb{I} \in M_{Q_2+1, Q_2+1}$ είναι ο αντίστοιχος μοναδιαίος πίνακας.

Οι ρυθμοί μετάβασης της Μαρκοβιανής αλυσίδας συνεχούς χρόνου επίσης δίνονται από το προηγούμενο διάγραμμα αλλά πρέπει να παρατηρήσουμε ότι η διαδικασία μπορεί να κάνει μια μετάβαση από οποιαδήποτε κατάσταση στην κατάσταση (Q_1, Q_2) με ρυθμό μ .

Πάλι θέλουμε να υπολογίσουμε το αναμενόμενο όφελος ανα περίοδο, επομένως χρειάζεται να καθορίσουμε το αναμενόμενο απόθεμα των δύο προϊόντων στο τέλος της περιόδου. Αφού η διάρκεια της περιόδου είναι εκθετικά κατανομημένη, χρησιμοποιώντας την ιδιότητα PASTA, το αναμενόμενο απόθεμα στο τέλος της περιόδου είναι ίσο με το αναμενόμενο απόθεμα σε μια τυχαία στιγμή. Συμβολίζοντας με $m_i^{(Q_1, Q_2)}$ το στάσιμο απόθεμα του προϊόντος i , $i \in \{1, 2\}$, έχουμε ότι το αναμενόμενο όφελος ανα περίοδο για αυτή την προσέγγιση δίνεται από τον παρακάτω τύπο

$$\begin{aligned} \pi^2(Q_1, Q_2) &= (r_1 - c_1)Q_1 + (r_2 - c_2)Q_2 \\ &\quad - (r_1 + h_1)Em_1^{(Q_1, Q_2)} - (r_2 + h_2)Em_2^{(Q_1, Q_2)}. \end{aligned} \quad (4.12)$$

Επομένως για να βρούμε μια αναλυτική έκφραση για το $\pi^2(Q_1, Q_2)$, χρειάζεται να υπολογίσουμε τη στάσιμη κατανομή της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(m_1^{(Q_1, Q_2)}(t), m_2^{(Q_1, Q_2)}(t)), t \in \mathbb{R}\}$.

Στην ενότητα 4.4 υπολογίζουμε τη στάσιμη κατανομή, δίνουμε έναν τύπο για το αναμενόμενο όφελος των Q_1 και Q_2 και αποδεικνύουμε ότι αυτή η συνάρτηση είναι submodular.

4.3 Προσέγγιση 1: Μέτρα απόδοσης

Σε αυτή την Ενότητα προσδιορίζουμε την μεταβατική κατανομή τη χρονική στιγμή T για τη Μαρκοβιανή αλυσίδα συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$ χρησιμοποιώντας την μέθοδο της ομοιομορφοποίησης και συνδυαστικά επιχειρήματα. Μετά, χρησιμοποιώντας αυτά τα αποτελέσματα, γράφουμε το αναμενόμενο όφελος ανά περίοδο σαν συνάρτηση των Q_1 και Q_2 και τέλος αποδεικνύουμε ότι αυτή η συνάρτηση είναι submodular.

4.3.1 Μεταβατική κατανομή

Σε αυτή την υποενότητα προσδιορίζουμε την μεταβατική κατανομή τη χρονική στιγμή T για την Μαρκοβιανή αλυσίδα συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$. Εισάγουμε τον ακόλουθο συμβολισμό:

$$p_{(j_1, j_2)}^{(Q_1, Q_2)}(t) = P(n_1^{(Q_1, Q_2)}(t) = j_1, n_2^{(Q_1, Q_2)}(t) = j_2),$$

$j_1 \in \{0, 1, 2, \dots, Q_1\}$, $j_2 \in \{0, 1, 2, \dots, Q_2\}$, $t \in [0, T]$, είναι η πιθανότητα μετάβασης της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$,

$$p^{(Q_1, Q_2)}(t) = [p_{(j_1, j_2)}^{(Q_1, Q_2)}(t)],$$

είναι το διάνυσμα γραμμή το οποίο δίνει τη μεταβατική κατανομή της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$,

$$\begin{aligned} p_{(i_1, i_2), (j_1, j_2)}^{(Q_1, Q_2)}(t) &= P(n_1^{(Q_1, Q_2)}(t) = j_1, n_2^{(Q_1, Q_2)}(t) = j_2 | n_1^{(Q_1, Q_2)}(0) = i_1, \\ &n_2^{(Q_1, Q_2)}(0) = i_2), \end{aligned}$$

$i_1, j_1 \in \{0, 1, 2, \dots, Q_1\}$, $i_2, j_2 \in \{0, 1, 2, \dots, Q_2\}$, $t \in [0, T]$, είναι η πιθανότητα μετάβασης της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$,

$$\mathbb{P}^{(Q_1, Q_2)}(t) = [p_{(i_1, i_2), (j_1, j_2)}^{(Q_1, Q_2)}(t)],$$

είναι ο πίνακας μεταβάσεων της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$.

Έτσι, θέλουμε να προσδιορίσουμε το διάνυσμα γραμμή $p^{(Q_1, Q_2)}(T)$. Έχουμε ότι

$$p^{(Q_1, Q_2)}(T) = p^{(Q_1, Q_2)}(0) \cdot \mathbb{P}^{(Q_1, Q_2)}(T). \quad (4.13)$$

Στην αρχή της περιόδου παραγγέλνουμε Q_1 μονάδες από το i και Q_2 μονάδες από το j προϊόν. Έτσι, η αρχική κατανομή $p^{(Q_1, Q_2)}(0)$ δίνεται από

$$p^{(Q_1, Q_2)}(0) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \quad (4.14)$$

Αφού το μόνο μη μηδενικό στοιχείο της $p^{(Q_1, Q_2)}(0)$ είναι το τελευταίο, το μόνο που χρειάζεται να ξέρουμε είναι η τελευταία γραμμή του πίνακα $\mathbb{P}^{(Q_1, Q_2)}(T)$ έτσι ώστε να υπολογίσουμε $p^{(Q_1, Q_2)}(T)$.

Θα προσδιορίσουμε αυτά τα στοιχεία του $p^{(Q_1, Q_2)}(T)$ χρησιμοποιώντας τη μέθοδο της ομοιομορφοποίησης (uniformization). Η μεγαλύτερη απόλυτη τιμή του γεννήτορα πίνακα $\mathbb{Q}^{(1)}$ είναι s , (υπενθυμίζουμε ότι $s = \lambda_1 + \lambda_2$), έτσι έχουμε ότι

$$p^{(Q_1, Q_2)}(t) = \sum_{k=0}^{\infty} e^{-st} \frac{(st)^k}{k!} \mathbb{P}^k, \quad (4.15)$$

όπου ο πίνακας $\mathbb{P} \in M_{Q_1+1 \times Q_2+1, Q_1+1 \times Q_2+1}$ δίνεται από

$$\mathbb{P} = \frac{1}{s} \mathbb{Q}^{(1)} + \mathbb{I}, \quad (4.16)$$

όπου ο πίνακας $\mathbb{Q}^{(1)}$ δίνεται από την (4.1) και $\mathbb{I} \in M_{Q_1+1 \times Q_2+1, Q_1+1 \times Q_2+1}$ είναι ο διαγώνιος πίνακας με όλα τα διαγώνια στοιχεία ίσα με μονάδα.

Έτσι,

$$\mathbb{P} = \begin{pmatrix} \mathbb{P}_0 & \mathbb{O} & \mathbb{O} & \dots & \mathbb{O} & \mathbb{O} \\ \mathbb{P}_1 & \mathbb{P}_2 & \mathbb{O} & \dots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \mathbb{P}_1 & \mathbb{P}_2 & \dots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \dots & \mathbb{P}_2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \dots & \mathbb{P}_1 & \mathbb{P}_2 \end{pmatrix}, \quad (4.17)$$

όπου

$$\mathbb{P}_0 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ s_2/s & (s-s_2)/s & 0 & \dots & 0 & 0 \\ 0 & s_2/s & (s-s_2)/s & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & (s-s_2)/s & 0 \\ 0 & 0 & 0 & \dots & s_2/s & (s-s_2)/s \end{pmatrix}, \quad (4.18)$$

$\mathbb{P}_0 \in M_{Q_2+1, Q_2+1}$, (υπενθυμίζουμε ότι $s_2 = \lambda_2 + \lambda_1 p_{12}$) όπου οι γραμμές και οι στήλες αντιστοιχούν στις καταστάσεις $(0, j)$, $j = 0, 1, \dots, Q_2$,

$$\mathbb{P}_1 = \begin{pmatrix} s_1/s & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_1/s & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_1/s & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_1/s & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda_1/s \end{pmatrix}, \quad (4.19)$$

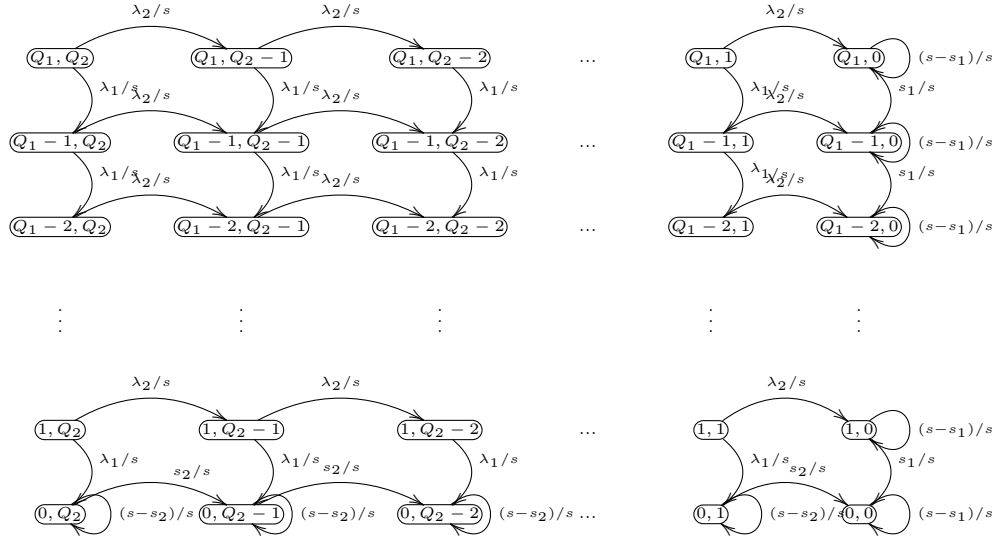
$\mathbb{P}_1 \in M_{Q_2+1, Q_2+1}$, (υπενθυμίζουμε ότι $s_1 = \lambda_1 + \lambda_2 p_{21}$), όπου για σταθερό i , $i = 1, \dots, Q_1$ οι γραμμές αντιστοιχούν στις καταστάσεις (i, j) , $j = 0, 1, \dots, Q_2$ και οι στήλες αντιστοιχούν στις καταστάσεις $(i-1, j)$, $j = 0, 1, \dots, Q_2$,

$$\mathbb{P}_2 = \begin{pmatrix} s-s_1/s & 0 & 0 & \dots & 0 & 0 \\ \lambda_2/s & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2/s & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & \lambda_2/s & 0 \end{pmatrix}, \quad (4.20)$$

$\mathbb{P}_2 \in M_{Q_2+1, Q_2+1}$, όπου για σταθερό i , $i = 1, \dots, Q_1$ οι γραμμές και οι στήλες αντιστοιχούν στις καταστάσεις (i, j) , $j = 0, 1, \dots, Q_2$, και \mathbb{O} δίνεται από την (4.5).

Ο πίνακας \mathbb{P} είναι ένας στοχαστικός πίνακας και μπορεί να θεωρηθεί σαν ένας πίνακας πιθανοτήτων μεταβάσεων ενός βήματος της Μαρκοβιανής αλυσίδας διακριτού

χρόνου $\{(n_1^d(k), n_2^d(k), k \in \{0, 1, \dots\})\}$ με χώρο καταστάσεων $S = \{(i_1, i_2) : i_1 \in \{0, \dots, Q_1\}, i_2 \in \{0, \dots, Q_2\}\}$ και μεταβάσεις όπως φαίνονται στο ακόλουθο διάγραμμα.



Συμβολίζοντας με $p_{(i_1, i_2), (j_1, j_2)}^d(k)$ την k -οστή πιθανότητα μετάβασης της Μαρκοβιανής αλυσίδας διακριτού χρόνου, π.χ.

$$p_{(i_1, i_2), (j_1, j_2)}^d(k) = P(n_1^d(k) = j_1, n_2^d(k) = j_2 | n_1^d(0) = i_1, n_2^d(0) = i_2),$$

έχουμε ότι $\mathbb{P}^k = [p_{(i_1, i_2), (j_1, j_2)}^d(k)]$.

Όπως αναφέραμε προηγούμενα, το μόνο που χρειάζεται να υπολογίσουμε είναι η τελευταία γραμμή του $p^{(Q_1, Q_2)}(T)$. Έτσι χρησιμοποιώντας την (4.15), χρειαζόμαστε μόνο τα στοιχεία της τελευταίας γραμμής του \mathbb{P}^k . Έτσι, οι πιθανότητες μετάβασης $p_{(Q_1, Q_2), (j_1, j_2)}^d(k)$, $j_1 \in \{0, \dots, Q_1\}$, $j_2 \in \{0, \dots, Q_2\}$, $k \in \{0, 1, \dots\}$ πρέπει να υπολογιστούν. Έχουμε την ακόλουθη Πρόταση.

Πρόταση 4.3.1. Οι πιθανότητες μετάβασης $p_{(Q_1, Q_2), (j_1, j_2)}^d(k)$, $j_1 \in \{0, \dots, Q_1\}$, $j_2 \in \{0, \dots, Q_2\}$, $k \in \{0, 1, \dots\}$ της Μαρκοβιανής αλυσίδας διακριτού χρόνου $\{(n_1^d(k), n_2^d(k), k \in \{0, 1, \dots\})\}$ δίνονται από τους ακόλουθους τύπους:

$$p_{(Q_1, Q_2), (j_1, j_2)}^d(k) = \begin{cases} \binom{k}{Q_1 - j_1} \left(\frac{\lambda_1}{s}\right)^{Q_1 - j_1} \left(\frac{\lambda_2}{s}\right)^{Q_2 - j_2}, & \text{αν } j_1, j_2 > 0, \\ 0, & \text{διαφορετικά,} \end{cases} \quad \begin{matrix} k = Q_1 - j_1 + Q_2 - j_2 \\ (4.21) \end{matrix}$$

$$p_{(Q_1, Q_2), (j_1, 0)}^d(k) = \begin{cases} \sum_{l=Q_2}^{Q_2+Q_1-j_1} \binom{l-1}{Q_2-1} \left(\frac{\lambda_2}{s}\right)^{Q_2} \left(\frac{\lambda_1}{s}\right)^{l-Q_2} \\ \cdot \binom{k-l}{Q_1+Q_2-j_1-l} \left(\frac{s_1}{s}\right)^{Q_1+Q_2-j_1-l} \left(\frac{s-s_1}{s}\right)^{k-Q_1-Q_2+j_1}, & \text{αν } j_1 > 0, \\ 0, & \text{διαφορετικά,} \end{cases} \quad \begin{matrix} k \geq Q_1 - j_1 + Q_2 \\ (4.22) \end{matrix}$$

$$p_{(Q_1, Q_2), (0, j_2)}^d(k) = \begin{cases} \sum_{l=Q_1}^{Q_2+Q_1-j_2} \binom{l-1}{Q_1-1} \left(\frac{\lambda_1}{s}\right)^{Q_1} \left(\frac{\lambda_2}{s}\right)^{l-Q_1} \\ \cdot \binom{k-l}{Q_1+Q_2-j_2-l} \left(\frac{s_2}{s}\right)^{Q_1+Q_2-j_2-l} \\ \cdot \left(\frac{s-s_2}{s}\right)^{k-Q_1-Q_2+j_2}, & \text{αν } j_2 > 0, \\ 0, & \text{διαφορετικά,} \end{cases} \quad \begin{matrix} k \geq Q_1 - j_2 + Q_2 \\ (4.23) \end{matrix}$$

$$p_{(Q_1, Q_2), (0, 0)}^d(k) = 1 - \sum_{(j_1, j_2) \in S \setminus (0, 0)} p_{(Q_1, Q_2), (j_1, j_2)}^d(k). \quad (4.24)$$

Απόδειξη. Θέλουμε να υπολογίσουμε την πιθανότητα η διακριτού χρόνου Μαρκοβιανή αλυσίδα $\{(n_1^d(k), n_2^d(k), k \in \{0, 1, \dots\})\}$ να κάνει μια μετάβαση από την κατάσταση

(Q_1, Q_2) στην κατάσταση (j_1, j_2) , $j_1 \in \{0, 1, \dots, Q_1\}$, $j_2 \in \{0, 1, \dots, Q_2\}$, σε k βήματα, $k \in \{0, 1, \dots\}$. Μπορούμε να το κάνουμε αυτό αν προσδιορίσουμε με πόσους τρόπους αυτή η μετάβαση μπορεί να γίνει και με τι πιθανότητα ο κάθε τρόπος.

Για να περιγράψουμε τις μεταβάσεις θα αναφέρουμε ένα βήμα από την κατάσταση (i_1, i_2) στην κατάσταση $(i_1 - 1, i_2)$, $i_1 \in \{1, \dots, Q_1\}$, $i_2 \in \{0, \dots, Q_2\}$ σαν τύπου-1 βήμα και ένα βήμα από την κατάσταση (i_1, i_2) στην κατάσταση $(i_1, i_2 - 1)$, $i_1 \in \{0, \dots, Q_1\}$, $i_2 \in \{1, \dots, Q_2\}$ σαν τύπου-2 βήμα.

Τώρα, αν $j_1, j_2 > 0$ μια μετάβαση από (Q_1, Q_2) στην (j_1, j_2) μπορεί να συμβεί σε ακριβώς $k = Q_1 - j_1 + Q_2 - j_2$ βήματα. Πιο συγκεκριμένα, πρέπει να έχουμε $Q_1 - j_1$ τύπου-1 βήματα και $Q_2 - j_2$ τύπου-2 βήματα. Αυτό μπορεί να γίνει με $\binom{k}{Q_1 - j_1}$ διαφορετικούς τρόπους και κάθε τρόπος έχει πιθανότητα $\left(\frac{\lambda_1}{s}\right)^{Q_1 - j_1} \cdot \left(\frac{\lambda_2}{s}\right)^{Q_2 - j_2}$. Έτσι, έχουμε την (4.21).

Για να βρούμε την πιθανότητα μιας μετάβασης από την κατάσταση (Q_1, Q_2) στην $(j_1, 0)$, $j_1 > 0$, σε k βήματα πρέπει να δεσμεύσουμε ως προς τον αριθμό των βημάτων μέχρι την πρώτη μεταπήδηση στο σύνολο των καταστάσεων $\{(j_1, 0), (j_1 + 1, 0), \dots, (Q_1, 0)\}$. Υποθέτοντας ότι ο αριθμός αυτών των βημάτων είναι l , $l = Q_2, \dots, Q_2 + Q_1 - j_1$, έχουμε ότι μετά από l βήματα το σύστημα είναι στην κατάσταση $(Q_1 + Q_2 - l, 0)$ και η μετάβαση από την (Q_1, Q_2) στην $(Q_1 + Q_2 - l, 0)$ μπορεί να συμβεί με $\binom{l-1}{Q_2 - 1}$ τρόπους (αφού το τελευταίο βήμα είναι τύπου-2) και κάθε τρόπος έχει πιθανότητα $\left(\frac{\lambda_2}{s}\right)^{Q_2} \cdot \left(\frac{\lambda_1}{s}\right)^{l - Q_2}$. Τότε, η μετάβαση από την $(Q_1 + Q_2 - l, 0)$ στην $(j_1, 0)$ σε $k - l$ βήματα μπορεί να συμβεί με $\binom{k-l}{Q_1 + Q_2 - j_1 - l}$ τρόπους, επειδή πρέπει να έχουμε $Q_1 + Q_2 - j_1 - l$ τύπου-1 βήματα, και κάθε τρόπος έχει πιθανότητα $\left(\frac{s_1}{s}\right)^{Q_1 + Q_2 - j_1 - l} \left(\frac{s - s_1}{s}\right)^{k - Q_1 - Q_2 + j_1}$. Έτσι, έχουμε την (4.22).

Χρησιμοποιώντας τα ίδια επιχειρήματα έχουμε την (4.23).

Τέλος, αφού ο πίνακας \mathbb{P}^k είναι στοχαστικός, έχουμε ότι $\sum_{(j_1, j_2) \in S} p_{(Q_1, Q_2), (j_1, j_2)}^d(k) = 1$. Έτσι, έχουμε τον τύπο (4.24).

□

Τώρα, χρησιμοποιώντας τους τύπους (4.13), (4.15) και την Πρόταση 4.3.1 προσδιορίζουμε απευθείας τις πιθανότητες μεταβάσεων $p_{(j_1, j_2)}^{(Q_1, Q_2)}(T)$, $j_1 \in \{0, 1, \dots, Q_1\}$, $j_2 \in$

$\{0, 1, \dots, Q_2\}$ της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$. Έχουμε την ακόλουθη Πρόταση.

Πρόταση 4.3.2. Οι πιθανότητες μετάβασης $p_{(j_1, j_2)}^{(Q_1, Q_2)}(T)$, $j_1 \in \{0, 1, \dots, Q_1\}$, $j_2 \in \{0, 1, \dots, Q_2\}$ της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(n_1^{(Q_1, Q_2)}(t), n_2^{(Q_1, Q_2)}(t)) : t \in [0, T]\}$ δίνονται από τους ακόλουθους τύπους:

$$p_{(j_1, j_2)}^{(Q_1, Q_2)}(T) = e^{-st} \cdot \frac{(st)^{Q_1+Q_2-j_1-j_2}}{(Q_1+Q_2-j_1-j_2)!} \cdot \binom{Q_1+Q_2-j_1-j_2}{Q_1-j_1} \cdot \left(\frac{\lambda_1}{s}\right)^{Q_1-j_1} \cdot \left(\frac{\lambda_2}{s}\right)^{Q_2-j_2}, \quad j_1, j_2 > 0, \quad (4.25)$$

$$p_{(j_1, 0)}^{(Q_1, Q_2)}(T) = \sum_{k=Q_1+Q_2-j_1}^{\infty} e^{-st} \cdot \frac{(st)^k}{k!} \sum_{l=Q_2}^{Q_2+Q_1-j_1} \binom{l-1}{Q_2-1} \cdot \left(\frac{\lambda_2}{s}\right)^{Q_2} \cdot \left(\frac{\lambda_1}{s}\right)^{l-Q_2} \cdot \binom{k-l}{Q_1+Q_2-j_1-l} \cdot \left(\frac{s_1}{s}\right)^{Q_1+Q_2-j_1-l} \cdot \left(\frac{s-s_1}{s}\right)^{k-Q_1-Q_2+j_1}, \quad j_1 > 0, \quad (4.26)$$

$$p_{(0, j_2)}^{(Q_1, Q_2)}(T) = \sum_{k=Q_1+Q_2-j_2}^{\infty} e^{-st} \cdot \frac{(st)^k}{k!} \sum_{l=Q_2}^{Q_2+Q_1-j_2} \binom{l-1}{Q_1-1} \cdot \left(\frac{\lambda_1}{s}\right)^{Q_1} \cdot \left(\frac{\lambda_2}{s}\right)^{l-Q_1} \cdot \binom{k-l}{Q_1+Q_2-j_2-l} \cdot \left(\frac{s_2}{s}\right)^{Q_1+Q_2-j_2-l} \cdot \left(\frac{s-s_2}{s}\right)^{k-Q_1-Q_2+j_2}, \quad j_2 > 0, \quad (4.27)$$

$$p_{(0,0)}^{(Q_1, Q_2)}(T) = 1 - \sum_{(j_1, j_2) \in S \setminus (0,0)} p_{(j_1, j_2)}^{(Q_1, Q_2)}(T). \quad (4.28)$$

4.3.2 Αναμενόμενο όφελος

Σε αυτή την υποενότητα θα εκφράσουμε το αναμενόμενο όφελος ανά περίοδο ως συνάρτηση των Q_1 και Q_2 . Έχουμε ότι

$$En_1^{(Q_1, Q_2)}(T) = \sum_{j_1=0}^{Q_1} j_1 \sum_{j_2=0}^{Q_2} p_{(j_1, j_2)}^{(Q_1, Q_2)}(T)$$

και

$$En_2^{(Q_1, Q_2)}(T) = \sum_{j_2=0}^{Q_2} j_2 \sum_{j_1=0}^{Q_1} p_{(j_1, j_2)}^{(Q_1, Q_2)}(T),$$

όπου οι πιθανότητες μετάβασης $p_{(j_1, j_2)}^{(Q_1, Q_2)}(T)$, $j_1 \in \{0, 1, \dots, Q_1\}$, $j_2 \in \{0, 1, \dots, Q_2\}$ δίνονται από τις (4.25)-(4.28).

Έτσι, έχουμε ότι

$$En_1^{(Q_1, Q_2)}(T) = \sum_{j_1=1}^{Q_1} j_1 \cdot \left[(4.26) + \sum_{j_2=1}^{Q_2} (4.25) \right] \quad (4.29)$$

και

$$En_2^{(Q_1, Q_2)}(T) = \sum_{j_2=1}^{Q_2} j_2 \cdot \left[(4.27) + \sum_{j_1=1}^{Q_1} (4.25) \right]. \quad (4.30)$$

Τώρα, αντικαθιστώντας την (4.29) και την (4.30) στην (4.6), βρίσκουμε το αναμενόμενο όφελος ανά περίοδο

$$\pi^{(1)}(Q_1, Q_2) = (r_1 - c_1)Q_1 + (r_2 - c_2)Q_2 - (r_1 + h_1) \cdot (4.29) - (r_2 + h_2) \cdot (4.30). \quad (4.31)$$

4.3.3 Submodularity του αναμενόμενου οφέλους

Σε αυτή την υποενότητα θα αποδείξουμε ότι το αναμενόμενο όφελος ανά περίοδο είναι submodular για την πρώτη προσέγγιση. Έχουμε το ακόλουθο Θεώρημα.

Θεώρημα 4.3.1. Το $\pi^{(1)}(Q_1, Q_2)$ είναι submodular συνάρτηση.

Απόδειξη. Για να αποδείξουμε ότι το $\pi^{(1)}(Q_1, Q_2)$ είναι submodular πρέπει να αποδείξουμε την ανισότητα

$$\pi^{(1)}(Q_1 + 1, Q_2 + 1) - \pi^{(1)}(Q_1 + 1, Q_2) - \pi^{(1)}(Q_1, Q_2 + 1) + \pi^{(1)}(Q_1, Q_2) \leq 0. \quad (4.32)$$

Χρησιμοποιώντας την (4.6), έχουμε ότι η (4.32) ισχύει αν και μόνο αν

$$\begin{aligned} & -(r_1 + h_1) \{En_1^{(Q_1+1, Q_2+1)}(T) - En_1^{(Q_1+1, Q_2)}(T) \\ & \quad - En_1^{(Q_1, Q_2+1)}(T) + En_1^{(Q_1, Q_2)}(T)\} \\ & -(r_2 + h_2) \{En_2^{(Q_1+1, Q_2+1)}(T) - En_2^{(Q_1+1, Q_2)}(T) \\ & \quad - En_2^{(Q_1, Q_2+1)}(T) + En_2^{(Q_1, Q_2)}(T)\} \leq 0. \end{aligned}$$

Αφού $r_i, h_i \geq 0$, $i = 1, 2$, για να αποδείξουμε την (4.32) αρκεί να δείξουμε ότι

$$En_1^{(Q_1+1, Q_2+1)}(T) - En_1^{(Q_1+1, Q_2)}(T) \geq En_1^{(Q_1, Q_2+1)}(T) - En_1^{(Q_1, Q_2)}(T), \quad (4.33)$$

και

$$En_2^{(Q_1+1, Q_2+1)}(T) - En_2^{(Q_1+1, Q_2)}(T) \geq En_2^{(Q_1, Q_2+1)}(T) - En_2^{(Q_1, Q_2)}(T). \quad (4.34)$$

Θα αποδείξουμε τις ανισότητες (4.33) και (4.34) χρησιμοποιώντας τη μέθοδο σύζευξης (coupling).

Υποθέτουμε ότι στην αρχή μιας περιόδου το σύστημα Α είναι στην κατάσταση $(Q_1 + 1, Q_2 + 1)$, το σύστημα Β στην κατάσταση $(Q_1 + 1, Q_2)$, το σύστημα Γ στην κατάσταση $(Q_1, Q_2 + 1)$ και το σύστημα Δ στην κατάσταση (Q_1, Q_2) . Τρέχουμε ταυτόχρονα αυτά τα τέσσερα συστήματα μέσω των αφίξεων και των ανταλλαγών. Μπορούμε εύκολα να δούμε ότι για κάθε ακολουθία από γεγονότα το απόθεμα του προϊόντος 1 στο σύστημα Α είναι το ίδιο με αυτό στο σύστημα Β και αυτό στο σύστημα Γ είναι ίδιο με αυτό στο σύστημα Δ (έτσι $n_1^{(Q_1+1, Q_2+1)}(T) = n_1^{(Q_1+1, Q_2)}(T)$ και $n_1^{(Q_1, Q_2+1)}(T) = n_1^{(Q_1, Q_2)}(T)$) εκτός από τις ακόλουθες περιπτώσεις.

Αρχικά υποθέτουμε ότι μετά από κάποια ακολουθία ζητήσεων το σύστημα Α βρίσκεται στην κατάσταση $(i, 1)$, $i \geq 1$ (έτσι τα συστήματα Β, Γ και Δ είναι στις καταστάσεις $(i, 0)$, $(i - 1, 1)$ και $(i - 1, 0)$, αντίστοιχα) και έρχεται ζήτηση του προϊόντος 2 όπου μεταπηδά (για τα συστήματα Β και Δ) στο προϊόν 1. Τότε έχουμε ότι $n_1^{(Q_1+1, Q_2+1)}(T) = 1 + n_1^{(Q_1+1, Q_2)}(T)$ και $n_1^{(Q_1, Q_2+1)}(T) = 1 + n_1^{(Q_1, Q_2)}(T)$. Αυτές οι ισότητες ισχύουν για κάθε ακολουθία γεγονότων μέχρι τα συστήματα Α, Β, Γ και Δ βρεθούν στις καταστάσεις $(2, 0)$, $(1, 0)$, $(1, 0)$ και $(0, 0)$, αντίστοιχα. Τότε, αν μια ζήτηση του προϊόντος 1 έρθει ή έρθει του προϊόντος 2 και μεταπηδήσει στο προϊόν 1, τα συστήματα θα πάνε στις καταστάσεις $(1, 0)$, $(0, 0)$, $(0, 0)$ και $(0, 0)$, αντίστοιχα.

Έτσι, $n_1^{(Q_1+1, Q_2+1)}(T) = 1 + n_1^{(Q_1+1, Q_2)}(T)$ και $n_1^{(Q_1, Q_2+1)}(T) = n_1^{(Q_1, Q_2)}(T)$. Τα τέσσερα συστήματα θα παραμείνουν σε αυτές τις καταστάσεις μέχρι μια ζήτηση του προϊόντος 1 να έρθει ή να έρθει μια ζήτηση του προϊόντος 2 και να μεταπηδήσει στο προϊόν 1. Τότε, όλα τα συστήματα θα είναι στην κατάσταση $(0, 0)$ και πάλι οι ισότητες $n_1^{(Q_1+1, Q_2+1)}(T) = n_1^{(Q_1+1, Q_2)}(T)$ και $n_1^{(Q_1, Q_2+1)}(T) = n_1^{(Q_1, Q_2)}(T)$ θα ισχύουν.

Τώρα, υποθέτουμε ότι μετά από μια ακολουθία γεγονότων το σύστημα A είναι στην κατάσταση $(1, 1)$ (έτσι τα συστήματα B, Γ και Δ είναι στις καταστάσεις $(1, 0)$, $(0, 1)$ και $(0, 0)$, αντίστοιχα) και συμβαίνει μια ζήτηση προϊόντος 2 η οποία (για τα συστήματα B και Δ) μεταπηδά στο προϊόν 1. Τότε, τα τέσσερα συστήματα θα πάνε απευθείας στις καταστάσεις $(1, 0)$, $(0, 0)$, $(0, 0)$ και $(0, 0)$ και συνεχίζουμε όπως περιγράψαμε παραπάνω.

Έτσι, για κάθε ακολουθία γεγονότων έχουμε

$$n_1^{(Q_1+1, Q_2+1)}(T) - n_1^{(Q_1+1, Q_2)}(T) \geq n_1^{(Q_1, Q_2+1)}(T) - n_1^{(Q_1, Q_2)}(T).$$

Έτσι,

$$En_1^{(Q_1+1, Q_2+1)}(T) - En_1^{(Q_1+1, Q_2)}(T) \geq En_1^{(Q_1, Q_2+1)}(T) - En_1^{(Q_1, Q_2)}(T).$$

Έτσι, έχουμε αποδείξει την (4.33).

Χρησιμοποιώντας παρόμοια επιχειρήματα αποδεικνύουμε την (4.34). \square

4.4 Προσέγγιση 2: Μέτρα απόδοσης

Σε αυτή την Ενότητα βρίσκουμε τη στάσιμη κατανομή της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(m_1^{(Q_1, Q_2)}(t), m_2^{(Q_1, Q_2)}(t)), t \in \mathbb{R}\}$. Μετά δίνουμε μια αναλυτική έκφραση του αναμενόμενου οφέλους ανά περίοδο ως συνάρτηση των Q_1 και Q_2 . Τέλος, αποδεικνύουμε ότι το αναμενόμενο όφελος είναι submodular συνάρτηση.

4.4.1 Στάσιμη κατανομή

Σε αυτή την υποενότητα προσδιορίζουμε την στάσιμη κατανομή του συστήματος όταν υποθέσουμε ότι η διάρκεια μιας περιόδου είναι εκθετικά κατανομημένη. Έχουμε την ακόλουθη Πρόταση.

Πρόταση 4.4.1. Η στάσιμη κατανομή $(\pi_{(i_1, i_2)}^{(Q_1, Q_2)} : (i_1, i_2) \in S)$ της Μαρκοβιανής αλυσίδας συνεχούς χρόνου $\{(m_1^{(Q_1, Q_2)}(t), m_2^{(Q_1, Q_2)}(t)), t \in \mathbb{R}\}$ δίνεται από τους τύπους:

$$\pi_{(Q_1 - i_1, Q_2 - i_2)}^{(Q_1, Q_2)} = \binom{i_1 + i_2}{i_1} q_1^{i_1} q_2^{i_2} q_3, \quad i_1 = 0, 1, \dots, Q_1 - 1, \quad i_2 = 0, 1, \dots, Q_2 - 1, \quad (4.35)$$

$$\pi_{(Q_1 - i_1, 0)}^{(Q_1, Q_2)} = B_2 q_2^{Q_2 - 1} q_3 \sum_{k=0}^{i_1} A_1^{i_1 - k} q_1^k \binom{Q_2 + k - 1}{k}, \quad i_1 = 0, 1, \dots, Q_1 - 1, \quad (4.36)$$

$$\pi_{(0, Q_2 - i_2)}^{(Q_1, Q_2)} = B_1 q_1^{Q_1 - 1} q_3 \sum_{k=0}^{i_2} A_2^{i_2 - k} q_2^k \binom{Q_1 + k - 1}{k}, \quad i_2 = 0, 1, \dots, Q_2 - 1, \quad (4.37)$$

$$\begin{aligned} \pi_{(0,0)}^{(Q_1, Q_2)} &= \frac{s_1}{\mu} B_2 q_2^{Q_2 - 1} q_3 \sum_{k=0}^{Q_1 - 1} A_1^{Q_1 - 1 - k} q_1^k \binom{Q_2 + k - 1}{k} \\ &\quad + \frac{s_2}{\mu} B_1 q_1^{Q_1 - 1} q_3 \sum_{k=0}^{Q_2 - 1} A_2^{Q_2 - 1 - k} q_2^k \binom{Q_1 + k - 1}{k}, \end{aligned} \quad (4.38)$$

όπου

$$q_i = \frac{\lambda_i}{s + \mu}, \quad i = 1, 2, \quad (4.39)$$

$$q_3 = \frac{\mu}{s + \mu}, \quad (4.40)$$

$$A_1 = \frac{s_1}{s_1 + \mu}, \quad (4.41)$$

$$A_2 = \frac{s_2}{s_2 + \mu}, \quad (4.42)$$

$$B_1 = \frac{\lambda_1}{s_2 + \mu}, \quad (4.43)$$

και

$$B_2 = \frac{\lambda_2}{s_1 + \mu}. \quad (4.44)$$

Απόδειξη. Η στάσιμη κατανομή $(\pi_{(i_1, i_2)}^{(Q_1, Q_2)} : (i_1, i_2) \in S)$ είναι η μοναδική θετική κανονικοποιημένη λύση του ακόλουθου συστήματος εξισώσεων ισορροπίας:

$$\pi_{(Q_1, Q_2)}^{(Q_1, Q_2)} \cdot (s + \mu) = \mu \sum_{i_1=0}^{Q_1} \sum_{i_2=0}^{Q_2} \pi_{(Q_1-i_1, Q_2-i_2)}^{(Q_1, Q_2)}, \quad (4.45)$$

$$\pi_{(Q_1, Q_2-i_2)}^{(Q_1, Q_2)} \cdot (s + \mu) = \lambda_2 \cdot \pi_{(Q_1, Q_2-i_2+1)}^{(Q_1, Q_2)}, \quad i_2 = 1, \dots, Q_2 - 1, \quad (4.46)$$

$$\pi_{(Q_1-i_1, Q_2)}^{(Q_1, Q_2)} \cdot (s + \mu) = \lambda_1 \cdot \pi_{(Q_1-i_1+1, Q_2)}^{(Q_1, Q_2)}, \quad i_1 = 1, \dots, Q_1 - 1, \quad (4.47)$$

$$\begin{aligned} \pi_{(Q_1-i_1, Q_2-i_2)}^{(Q_1, Q_2)} \cdot (s + \mu) &= \lambda_1 \cdot \pi_{(Q_1-i_1+1, Q_2-i_2)}^{(Q_1, Q_2)} + \lambda_2 \cdot \pi_{(Q_1-i_1, Q_2-i_2+1)}^{(Q_1, Q_2)}, \\ & \quad i_1 = 1, \dots, Q_1 - 1, \quad i_2 = 1, \dots, Q_2 - 1, \end{aligned} \quad (4.48)$$

$$\pi_{(Q_1, 0)}^{(Q_1, Q_2)} \cdot (s_1 + \mu) = \lambda_2 \cdot \pi_{(Q_1, 1)}^{(Q_1, Q_2)}, \quad (4.49)$$

$$\pi_{(0, Q_2)}^{(Q_1, Q_2)} \cdot (s_2 + \mu) = \lambda_1 \cdot \pi_{(1, Q_2)}^{(Q_1, Q_2)}, \quad (4.50)$$

$$\begin{aligned} \pi_{(Q_1-i_1, 0)}^{(Q_1, Q_2)} \cdot (s_1 + \mu) &= s_1 \cdot \pi_{(Q_1-i_1+1, 0)}^{(Q_1, Q_2)} + \lambda_2 \cdot \pi_{(Q_1-i_1, 1)}^{(Q_1, Q_2)}, \\ & \quad i_1 = 1, \dots, Q_1 - 1, \end{aligned} \quad (4.51)$$

$$\begin{aligned} \pi_{(0, Q_2-i_2)}^{(Q_1, Q_2)} \cdot (s_2 + \mu) &= s_2 \cdot \pi_{(0, Q_2-i_2+1)}^{(Q_1, Q_2)} + \lambda_1 \cdot \pi_{(1, Q_2-i_2)}^{(Q_1, Q_2)}, \\ & \quad i_2 = 1, \dots, Q_2 - 1, \end{aligned} \quad (4.52)$$

$$\pi_{(0, 0)}^{(Q_1, Q_2)} \cdot \mu = s_1 \cdot \pi_{(1, 0)}^{(Q_1, Q_2)} + s_2 \cdot \pi_{(0, 1)}^{(Q_1, Q_2)}, \quad (4.53)$$

όπου στην (4.45) έχουμε συμπεριλάβει τις ψευδομεταβάσεις από την κατάσταση (Q_1, Q_2) στην (Q_1, Q_2) με ρυθμό μ , όπου δεν υπήρχε ζήτηση. Χρησιμοποιώντας την εξίσωση κανονικοποίησης

$$\sum_{i_1=0}^{Q_1} \sum_{i_2=0}^{Q_2} \pi_{(Q_1-i_1, Q_2-i_2)}^{(Q_1, Q_2)} = 1, \quad (4.54)$$

η (4.45) μπορεί να γραφτεί ως

$$\pi_{(Q_1, Q_2)}^{(Q_1, Q_2)} \cdot (s + \mu) = \mu. \quad (4.55)$$

Σημειώνουμε ότι η Μαρκοβιανή αλυσίδα συνεχούς χρόνου είναι πάντα θετικά επαναληπτική. Χρησιμοποιώντας τις (4.45)-(4.53) μετά από λίγη άλγεβρα δείχνουμε τις εκφράσεις (4.35)-(4.38) για τις στάσιμες πιθανότητες. \square

4.4.2 Αναμενόμενο όφελος

Σε αυτή την υποενότητα θα δείξουμε μια αναλυτική έκφραση του αναμενόμενου οφέλους ανά περίοδο ως συνάρτηση των Q_1 και Q_2 . Έχουμε ότι

$$Em_1^{(Q_1, Q_2)} = \sum_{i_1=0}^{Q_1} (Q_1 - i_1) \sum_{i_2=0}^{Q_2} \pi_{(Q_1-i_1, Q_2-i_2)}^{(Q_1, Q_2)},$$

και

$$Em_2^{(Q_1, Q_2)} = \sum_{i_2=0}^{Q_2} (Q_2 - i_2) \sum_{i_1=0}^{Q_1} \pi_{(Q_1-i_1, Q_2-i_2)}^{(Q_1, Q_2)},$$

όπου οι στάσιμες πιθανότητες $\pi_{(Q_1-i_1, Q_2-i_2)}^{(Q_1, Q_2)}$, $i_1 \in \{0, \dots, Q_1\}$, $i_2 \in \{0, \dots, Q_2\}$ δίνονται από τις (4.35)-(4.38).

Έτσι, έχουμε ότι

$$\begin{aligned} Em_1^{(Q_1, Q_2)} &= \sum_{i_1=0}^{Q_1-1} (Q_1 - i_1) \left[\sum_{i_2=0}^{Q_2-1} (4.35) + (4.36) \right] \\ &\quad \sum_{i_1=0}^{Q_1-1} \sum_{i_2=0}^{Q_2-1} (Q_1 - i_1) \binom{i_1 + i_2}{i_1} q_1^{i_1} q_2^{i_2} q_3 + B_2 \cdot \\ &\quad q_2^{Q_2-1} q_3 \sum_{i_1=0}^{Q_1-1} \sum_{k=0}^{i_1} (Q_1 - i_1) A_1^{i_1-k} q_1^k \binom{Q_2 + k - 1}{k} \end{aligned} \quad (4.56)$$

και

$$\begin{aligned} Em_2^{(Q_1, Q_2)} &= \sum_{i_2=0}^{Q_2-1} (Q_2 - i_2) \left[\sum_{i_1=0}^{Q_1-1} (4.35) + (4.37) \right] \\ &\quad \sum_{i_2=0}^{Q_2-1} \sum_{i_1=0}^{Q_1-1} (Q_2 - i_2) \binom{i_1 + i_2}{i_1} q_1^{i_1} q_2^{i_2} q_3 + B_1 \cdot \\ &\quad q_1^{Q_1-1} q_3 \sum_{i_2=0}^{Q_2-1} \sum_{k=0}^{i_2} (Q_2 - i_2) A_2^{i_2-k} q_2^k \binom{Q_1 + k - 1}{k} \end{aligned} \quad (4.57)$$

Αντικαθιστώντας την (4.56) και την (4.57) στην (4.12) έχουμε ότι το αναμενόμενο όφελος ανα περίοδο είναι

$$\pi^{(2)}(Q_1, Q_2) = (r_1 - c_1)Q_1 + (r_2 - c_2)Q_2 - (r_1 + h_1)(4.56) - (r_2 + h_2)(4.57). \quad (4.58)$$

4.4.3 Submodularity του αναμενόμενου οφέλους

Σε αυτή την υποενότητα θα αποδείξουμε ότι το αναμενόμενο όφελος ανά περίοδο είναι submodular συνάρτηση για τη δεύτερη προσέγγιση. Έχουμε το ακόλουθο Θεώρημα.

Θεώρημα 4.4.1. Το $\pi^{(2)}(Q_1, Q_2)$ είναι submodular συνάρτηση.

Απόδειξη. Για να αποδείξουμε ότι το $\pi^{(2)}(Q_1, Q_2)$ είναι submodular πρέπει να αποδείξουμε την ανισότητα

$$\pi^{(2)}(Q_1 + 1, Q_2 + 1) - \pi^{(2)}(Q_1 + 1, Q_2) - \pi^{(2)}(Q_1, Q_2 + 1) + \pi^{(2)}(Q_1, Q_2) \leq 0. \quad (4.59)$$

Χρησιμοποιώντας την (4.12), έχουμε ότι η (4.59) ισχύει αν και μόνο αν

$$\begin{aligned} & -(r_1 + h_1) \{Em_1^{(Q_1+1, Q_2+1)} - Em_1^{(Q_1+1, Q_2)} \\ & \quad - Em_1^{(Q_1, Q_2+1)} + Em_1^{(Q_1, Q_2)}\} \\ & -(r_2 + h_2) \{Em_2^{(Q_1+1, Q_2+1)} - Em_2^{(Q_1+1, Q_2)} \\ & \quad - Em_2^{(Q_1, Q_2+1)} + Em_2^{(Q_1, Q_2)}\} \leq 0. \end{aligned}$$

Αφού $r_i, h_i \geq 0$, $i = 1, 2$, για να αποδείξουμε την (4.59) αρκεί να δείξουμε ότι

$$Em_1^{(Q_1+1, Q_2+1)} - Em_1^{(Q_1+1, Q_2)} \geq Em_1^{(Q_1, Q_2+1)} - Em_1^{(Q_1, Q_2)}, \quad (4.60)$$

και

$$Em_2^{(Q_1+1, Q_2+1)} - Em_2^{(Q_1+1, Q_2)} \geq Em_2^{(Q_1, Q_2+1)} - Em_2^{(Q_1, Q_2)}. \quad (4.61)$$

Οι ανισότητες (4.60) και (4.61) αποδεικνύονται χρησιμοποιώντας ακριβώς την ίδια ανάλυση όπως στο Θεώρημα 4.3.1. \square

4.5 Αλγόριθμος

Ο σκοπός αυτού του Κεφαλαίου είναι να προσδιορίσει τις βέλτιστες ποσότητες Q_1 και Q_2 οι οποίες θα παραγγέλονται στην αρχή κάθε περιόδου. Έτσι, για την πρώτη προσέγγιση έχουμε να λύσουμε το ακόλουθο πρόβλημα

$$\max_{\substack{Q_1, Q_2 \\ Q_1 + Q_2 \leq C}} \pi^{(1)}(Q_1, Q_2),$$

όπου σε ότι ακολουθεί $a_1 = a_2 = 1$.

Για σταθερό Q_1 συμβολίζουμε με $Q_2^*(Q_1)$ την μεγαλύτερη τιμή του Q_2 στην οποία συμβαίνει $\max_{Q_2 \in \{0, \dots, C - Q_1\}} \pi^{(1)}(Q_1, Q_2)$. Έτσι,

$$Q_2^*(Q_1) = \max\{Q_2^0 \in \{0, \dots, C - Q_1\} : \max_{Q_2 \in \{0, \dots, C - Q_1\}} \pi^{(1)}(Q_1, Q_2) = \pi^{(1)}(Q_1, Q_2^0)\}.$$

Τότε το $\max_{Q_1, Q_2} \pi^{(1)}(Q_1, Q_2)$ μπορεί να γραφτεί ως

$$\begin{aligned} \max_{\substack{Q_1, Q_2 \\ Q_1 + Q_2 \leq C}} \pi^{(1)}(Q_1, Q_2) &= \max_{Q_1 \in \{0, \dots, C\}} \left\{ \max_{Q_2 \in \{0, \dots, C - Q_1\}} \pi^{(1)}(Q_1, Q_2) \right\} \\ &= \max_{Q_1 \in \{0, \dots, C\}} \pi^{(1)}(Q_1, Q_2^*(Q_1)), \end{aligned}$$

και το πρόβλημα παίρνει την ακόλουθη μορφή

$$\max_{Q_1 \in \{0, \dots, C\}} \pi^{(1)}(Q_1, Q_2^*(Q_1)).$$

Αντίστοιχα, για την δεύτερη προσέγγιση έχουμε να λύσουμε το πρόβλημα

$$\max_{Q_1 \in \{0, \dots, C\}} \pi^{(2)}(Q_1, Q_2^*(Q_1)).$$

Θα παρουσιάσουμε έναν αλγόριθμο ο οποίος χρησιμοποιεί το submodularity των $\pi^{(1)}(Q_1, Q_2)$ και $\pi^{(2)}(Q_1, Q_2)$ για να λύσουμε αυτά τα προβλήματα γρήγορα. Αρχικά δίνουμε το Λήμμα 4.5.1.

Λήμμα 4.5.1. Αν $\pi(Q_1, Q_2)$ με $Q_1 \in \{0, \dots, C\}$, $Q_2 \in \{0, \dots, C - Q_1\}$ είναι μια submodular συνάρτηση και $Q_2^*(Q_1) = \max\{Q_2^0 \in \{0, \dots, C - Q_1\} : \max_{Q_2 \in \{0, \dots, C - Q_1\}} \pi(Q_1, Q_2) = \pi(Q_1, Q_2^0)\}$ τότε $Q_2^*(Q_1)$ είναι αύξουσα ως προς Q_1 .

Απόδειξη. Αφού η $\pi(Q_1, Q_2)$ είναι submodular, η ακόλουθη ανισότητα ισχύει

$$\begin{aligned} \pi(Q_1 + 1, Q_2 + 1) - \pi(Q_1 + 1, Q_2) - \pi(Q_1, Q_2 + 1) + \pi(Q_1, Q_2) \leq 0, \\ Q_1 \in \{0, \dots, C\}, Q_2 \in \{0, \dots, C - Q_1\}. \end{aligned} \quad (4.62)$$

Θα αποδείξουμε ότι

$$\pi(Q_1 + 1, Q_2') - \pi(Q_1 + 1, Q_2) - \pi(Q_1, Q_2') + \pi(Q_1, Q_2) \leq 0, \quad Q_2' > Q_2. \quad (4.63)$$

Προσθέτοντας την (4.62) για Q_2 και την (4.62) για $Q_2 + 1$ δείχνουμε

$$\begin{aligned} \pi(Q_1 + 1, Q_2 + 2) - \pi(Q_1 + 1, Q_2) - \pi(Q_1, Q_2 + 2) + \pi(Q_1, Q_2) \leq 0, \\ Q_1 \in \{0, \dots, C - 1\}, Q_2 \in \{0, \dots, C - Q_1 - 2\}. \end{aligned}$$

Αν συνεχίσουμε, μπορούμε να αποδείξουμε με επαγωγή την (4.63).

Τώρα, για $Q_1 = 0$, έχουμε $Q_2^*(0) \in [0, C]$ και

$$\pi(0, Q_2^*(0)) \geq \pi(0, Q_2), \quad \text{για κάθε } Q_2 \in [0, C]. \quad (4.64)$$

Επίσης, από την (4.62) για $Q_1 = 0$ δείχνουμε

$$\begin{aligned} \pi(1, Q_2 + 1) - \pi(1, Q_2) - \pi(0, Q_2 + 1) + \pi(0, Q_2) \leq 0, \\ Q_2 \in \{0, \dots, C - 1\}. \end{aligned} \quad (4.65)$$

Για να αποδείξουμε ότι $Q_2^*(1) \leq Q_2^*(0)$ είναι αρκετό να αποδείξουμε ότι $\pi(1, Q_2^*(0)) \geq \pi(1, Q_2)$, $Q_2 \in [Q_2^*(0), C - 1]$.

Από την (4.63) για $Q_1 = 0$ και $Q_2 = Q_2^*(0)$ έχουμε ότι

$$\pi(1, Q_2') - \pi(1, Q_2^*(0)) - \pi(0, Q_2') + \pi(0, Q_2^*(0)) \leq 0, \quad Q_2' > Q_2^*(0),$$

το οποίο δίνει

$$\pi(1, Q_2') - \pi(1, Q_2^*(0)) \leq \pi(0, Q_2') - \pi(0, Q_2^*(0)), \quad Q_2' > Q_2^*(0).$$

Το δεύτερο μέρος της προηγούμενης ανισότητας είναι μικρότερο ή ίσο με το μηδέν λόγω της (4.64). Έτσι, $\pi(1, Q_2') \leq \pi(1, Q_2^*(0))$, $Q_2' > Q_2^*(0)$. Άρα, $Q_2^*(1) \leq Q_2^*(0)$.

Χρησιμοποιώντας τα ίδια επιχειρήματα επαναλαμβανόμενα, αποδεικνύουμε ότι η $Q_2^*(Q_1)$ είναι φθίνουσα ως προς Q_1 . \square

Τώρα, χρησιμοποιώντας ότι η $Q_2^*(Q_1)$ είναι αύξουσα ως προς Q_1 θα παρουσιάσουμε έναν αλγόριθμο ο οποίος δίνει τις βέλτιστες ποσότητες Q_1 και Q_2 . Αυτός ο αλγόριθμος λύνει το πρόβλημα για την πρώτη προσέγγιση αν $i = 1$ και για την δεύτερη προσέγγιση αν $i = 2$, αντίστοιχα.

Αλγόριθμος

- Υπολόγισε,

$$Q_2^*(0) = \max\{Q_2^0 \in \{0, \dots, C\} : \max_{Q_2 \in \{0, \dots, C\}} \pi^{(i)}(0, Q_2) = \pi^{(i)}(0, Q_2^0)\}.$$

- Για $j = 1 : C$ υπολόγισε,

$$Q_2^*(j) = \max\{Q_2^0 \in \{0, \dots, Q_2^*(j-1)\} : \max_{Q_2 \in \{0, \dots, Q_2^*(j-1)\}} \pi^{(i)}(j, Q_2) = \pi^{(i)}(j, Q_2^0)\}.$$

- Υπολόγισε,

$$\max_{Q_1 \in \{0, \dots, C\}} \pi^{(i)}(Q_1, Q_2^*(Q_1)).$$

4.6 Αριθμητική ανάλυση

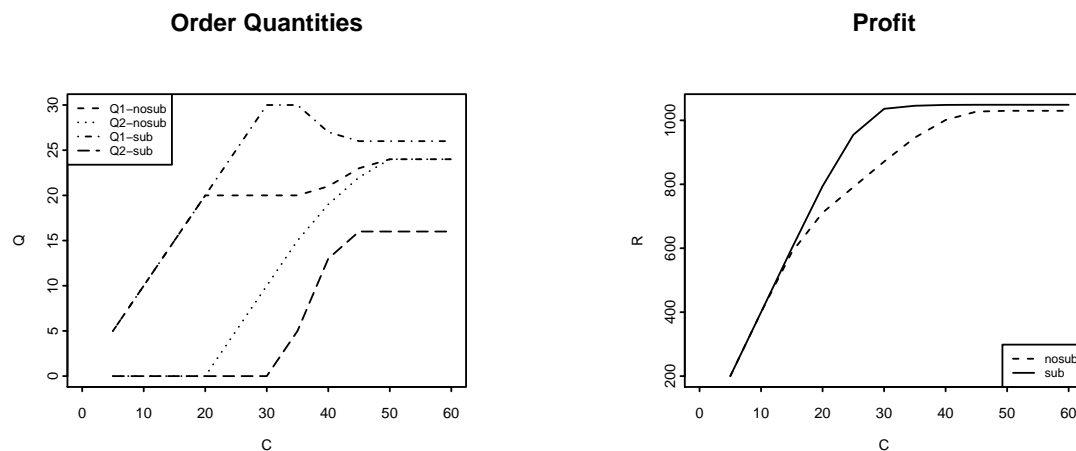
Σε αυτή την ενότητα εφαρμόζουμε τους αλγόριθμους που αναπτύχθηκαν προηγούμενα με σκοπό να αποκτήσουμε γνώσεις για την επίδραση της ανταλλαξιμότητας στις πολιτικές παραγγελίας και στο όφελος. Ενδιαφερόμαστε για δύο γενικές ερωτήσεις. Πρώτη, πόσο επηρεάζονται οι πολιτικές παραγγελίας και το αναμενόμενο όφελος ανά περίοδο όταν υπάρχει ανταλλαξιμότητα ανάμεσα στα δύο προϊόντα συγκρινόμενες με την περίπτωση όπου δεν έχουμε ανταλλαξιμότητα, και δεύτερη πόσο η τυχαιότητα του χρόνου ανεφοδιασμού επηρεάζει τις πολιτικές παραγγελιών και το όφελος.

Πρώτα αναλύουμε την επίδραση της ανταλλαξιμότητας, κάτω από σταθερό χρόνο ανεφοδιασμού. Θεωρούμε δύο προϊόντα για τα οποία οι ρυθμοί της ζήτησης Poisson

είναι ίσες με $\lambda_1 = \lambda_2 = 20$ μονάδες ανα μονάδα χρόνου και κανονικοποιούμε το χρόνο ανεφοδιασμού $T = 1$. Όσον αφορά τις οικονομικές παραμέτρους, μπορούμε να διακρίνουμε τρεις αντιπροσωπευτικές περιπτώσεις που αντιστοιχούν σε διαφορετικούς τύπους προϊόντων. Ειδικότερα, αν η χωρητικότητα αποθήκευσης είναι άπειρη και δεν υπάρχει ανταλλαξιμότητα, τότε το πρόβλημα παραγγελίας είναι γνωστό και για κάθε προϊόν ισχύει ο κανόνας του εφημεριδοπώλη: η βέλτιστη ποσότητα παραγγελίας του προϊόντος j είναι η πρώτη ώστε η πιθανότητα να επαρκέσει το προϊόν να είναι τουλάχιστον ίση με τον λόγο $R_j = \frac{p_j - c_j}{c_j + h_j}$. Βασιζόμενοι σε αυτή την παρατήρηση θεωρούμε τα ακόλουθα σενάρια: (i) και τα δύο προϊόντα έχουν ένα σχετικά μεγάλο λόγο $R_1 = R_2 = 0.8$, (ii) το ένα έχει υψηλό και το άλλο χαμηλό λόγο $R_1 = 0.8, R_2 = 0.4$ και (iii) και οι δύο λόγοι είναι μικροί $R_1 = R_2 = 0.4$. Για κάθε ένα από τα τρία σενάρια υπολογίζουμε την βέλτιστη πολιτική παραγγελιών και το βέλτιστο αναμενόμενο όφελος ανα περίοδο για διάφορες τιμές χωρητικότητας κάτω από δύο επιλογές: η πρώτη είναι όταν δεν έχουμε ανταλλαξιμότητα μεταξύ των δύο προϊόντων, και η δεύτερη όταν υπάρχει ανταλλαξιμότητα και από τις δύο πλευρές με $p_{12} = p_{21} = 0.4$. Τα αποτελέσματα παρουσιάζονται στα διαγράμματα 4.1, 4.2 και 4.3, αντίστοιχα.

Αρκετά συμπεράσματα μπορούν να εξαχθούν από αυτά τα αποτελέσματα. Πρώτα, η καμπύλη του οφέλους δείχνει ότι, όπως αναμενόταν, ο εφημεριδοπώλης είναι σε καλύτερη θέση όταν υπάρχει ανταλλαξιμότητα μεταξύ των προϊόντων. Δεύτερον, αυτό επιτυγχάνεται γενικά με την ανταλλαγή της αναλογίας των προϊόντων που έχουμε παραγγείλει, υπέρ του πιο κερδοφόρου από τα δύο.

Ειδικότερα, κάτω από το σενάριο 1 όπου και τα δύο προϊόντα έχουν υψηλή ποσότητα παραγγελίας, το όφελος της ανταλλαξιμότητας είναι μικρό σε σχέση με την μη ύπαρξη όταν η χωρητικότητα είναι είτε μικρή είτε μεγάλη. Στην πρώτη περίπτωση, οι ποσότητες παραγγελίας είναι πολύ μικρότερες από ότι χρειάζεται, και σαν αποτέλεσμα και τα δύο προϊόντα εξαντλούνται ταυτόχρονα και γρήγορα, χωρίς να δίνουν ευκαιρία να συμβεί ανταλλαγή στα προϊόντα. Σε αυτή την περίπτωση οι ποσότητες παραγγελίας είναι ίδιες ανάμεσα στις δύο πιθανότητες. Από την άλλη μεριά η χωρητικότητα είναι πολύ υψηλή και η ανταλλαξιμότητα προσφέρει μικρό όφελος. Είναι πιθανό να παραγγέλνουμε μικρότερη συνολικά ποσότητα και από τα δύο προϊόντα χωρίς να επηρεάζεται το επίπεδο υπηρεσιών. Αυτό συμβαίνει επειδή η ανταλλαξιμότητα ουσιαστικά επιτρέπει μερική εκμετάλλευση της από κοινού επίδρασης ανάμεσα στις δύο ανεξάρτητες ζητήσεις, και έτσι μειώνεται ο

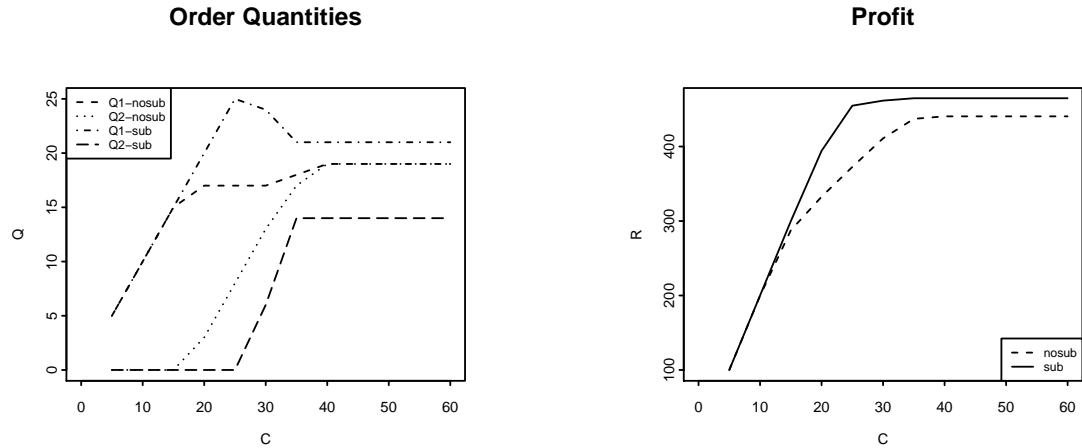


Σχήμα 4.1: Επίδραση ανταλλαξιμότητας, Σενάριο 1: $\lambda_1 = \lambda_2 = 20, T = 1, p_1 = 50, c_1 = 10, h_1 = 0, p_2 = 20, c_2 = 4, h_2 = 0, p_{12} = p_{21} = 0(\text{nosub}), 0.4(\text{sub})$

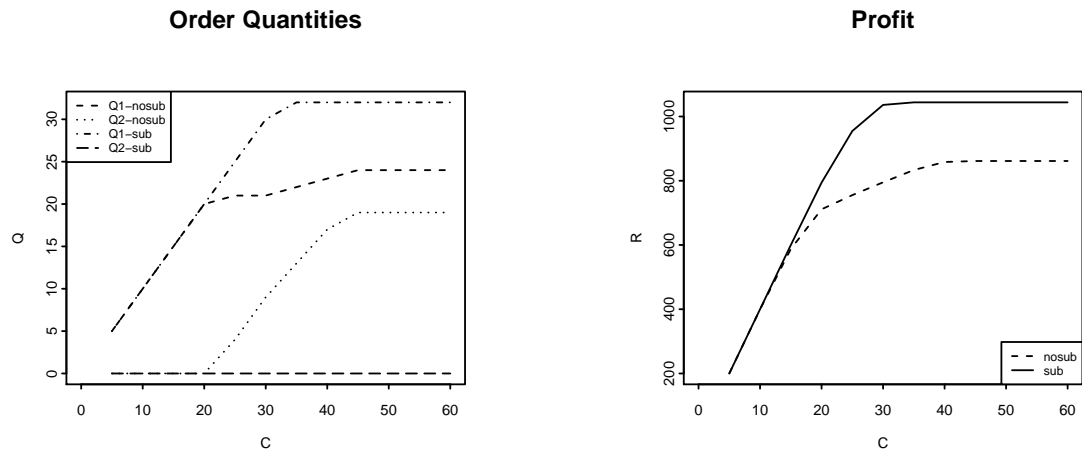
κίνδυνος πολλών απώλητων προϊόντων. Αυτή η ωφέλεια είναι ακόμη πιο έντονη όταν η χωρητικότητα είναι αρκετά χαμηλή και γίνεται περιοριστικός παράγοντας για το κέρδος, αλλά αρκετά υψηλή ώστε να επιτρέψει να εμφανιστεί η ανταλλαξιμότητα. Σε αυτό το ενδιάμεσο φάσμα το σχετικό όφελος αυξάνει λόγω του ότι η ανταλλαξιμότητα είναι συχνότερη.

Κάτω από το σενάριο 2, όπου και τα δύο προϊόντα έχουν σχετικά χαμηλές ποσότητες παραγγελίας, δείχνουμε ποιοτικά όμοια συμπεράσματα με το σενάριο 1. Από την άλλη μεριά, όταν τα δύο προϊόντα διαφέρουν σημαντικά όσο αφορά τους λόγους, που είναι το τρίτο σενάριο, τότε η επίδραση του της ανταλλαξιμότητας μπορεί να είναι σε τέτοιο βαθμό που το προϊόν με τον μικρότερο λόγο να τελειώσει εντελώς και η ζήτηση του να ικανοποιείται μερικώς μέσω της ανταλλαξιμότητας από το άλλο προϊόν. Σε αυτή την περίπτωση, το όφελος της ανταλλαξιμότητας από την άποψη της κερδοφορίας παραμένει ακόμα και όταν η χωρητικότητα είναι μη δεσμευτική.

Στο δεύτερο σύνολο αριθμητικών πειραμάτων, στοχεύουμε να βρούμε την επίδραση της μεταβλητότητας του χρόνου ανεφοδιασμού στην πολιτική παραγγελιών και τα οφέλη κάτω από την ανταλλαξιμότητα. Για αυτό το λόγο πρώτα παρουσιάζουμε ένα σύνολο πειραμάτων ανάλογο των σεναρίων 1-3 παραπάνω, αλλά τώρα ο χρόνος ανεφοδιασμού



Σχήμα 4.2: Επίδραση ανταλλαξιμότητας, Σενάριο 2: $\lambda_1 = \lambda_2 = 20, T = 1, p_1 = 50, c_1 = 30, h_1 = 0, p_2 = 20, c_2 = 12, h_2 = 0, p_{12} = p_{21} = 0(\text{nosub}), 0.4(\text{sub})$

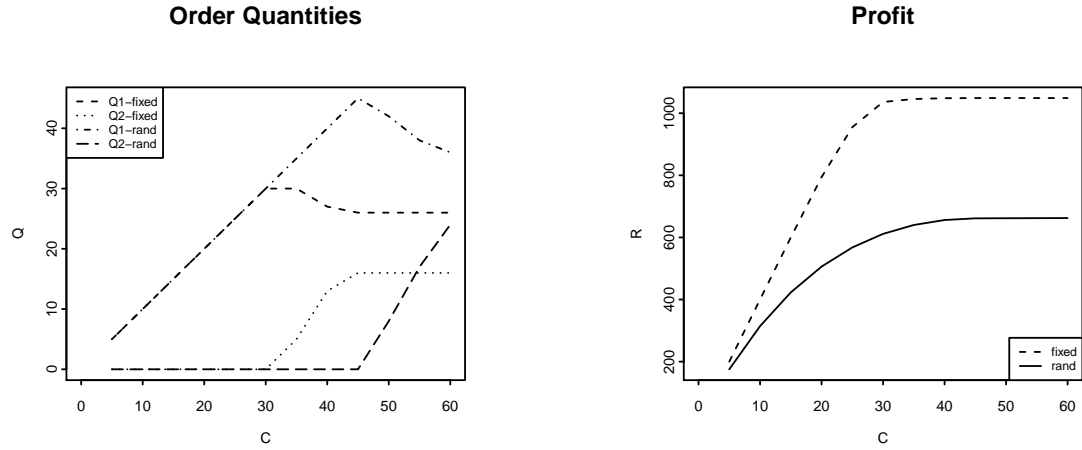


Σχήμα 4.3: Επίδραση ανταλλαξιμότητας, Σενάριο 3: $\lambda_1 = \lambda_2 = 20, T = 1, p_1 = 50, c_1 = 10, h_1 = 0, p_2 = 20, c_2 = 12, h_2 = 0, p_{12} = p_{21} = 0(\text{nosub}), 0.4(\text{sub})$

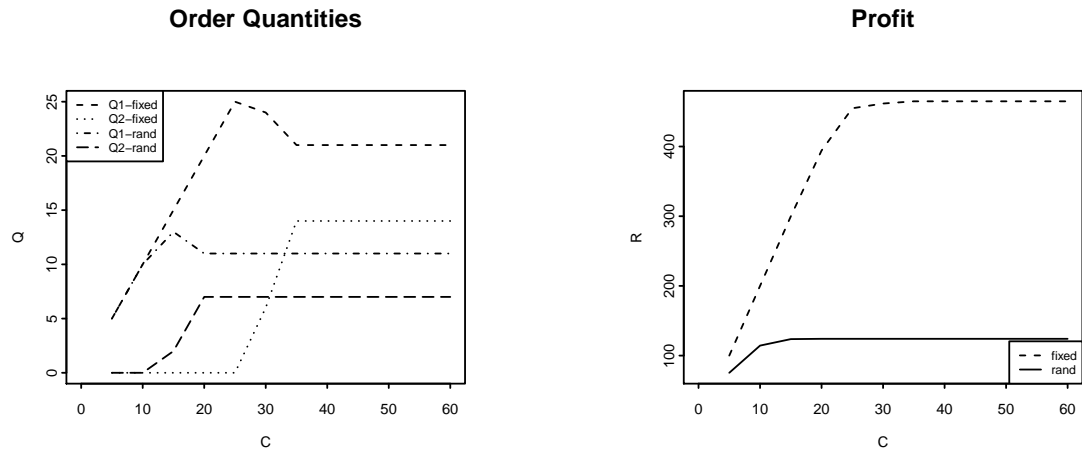
είναι εκθετικά κατανεμημένος με μέση τιμή 1. Η συμπεριφορά της πολιτικής παραγγελίας και η επίδραση της ανταλλαξιμότητας ήταν όμοια με τον σταθερό χρόνο ανεφοδιασμού, ωστόσο το αναμενόμενο όφελος ήταν ομοιόμορφα χαμηλότερο, δείχνοντας ότι η ανταλλαξιμότητα στον τυχαίο χρόνο ανεφοδιασμού έχει επιζήμια επίδραση. Για μεγαλύτερη ανάλυση της επίδρασης, δείχνουμε στα διαγράμματα 4.4 και 4.5 τη σύγκριση των πολιτικών παραγγελίας και των κερδών υπό την ανταλλαξιμότητα ανάμεσα στις περιπτώσεις του σταθερού και εκθετικά κατανεμημένου χρόνου ανεφοδιασμού. Από αυτά τα αποτελέσματα αποδεικνύεται ότι όταν ο χρόνος ανεφοδιασμού είναι τυχαίος υπάρχει μια σημαντική μείωση στο όφελος. Όσο αφορά την επίδραση στην πολιτική παραγγελίας, αυτό εξαρτάται από τους τύπους των προϊόντων. Ειδικότερα, κάτω από το σενάριο 1, ο κίνδυνος ελλείψεων έχει πιο μεγάλο κόστος από το να έχουμε απώλητα προϊόντα. Επομένως, στην τυχαία περίπτωση το πιο επιζήμιο γεγονός είναι ο χρόνος ανεφοδιασμού να είναι πολύ μεγάλος και σαν αποτέλεσμα να έχουμε πάρα πολλές ελλείψεις. Για να μετριάσουμε αυτή την πιθανότητα οι ποσότητες παραγγελίας είναι σημαντικά μεγαλύτερες από αυτές στη σταθερή περίπτωση. Από την άλλη μεριά, κάτω από το σενάριο 2, όπου και οι δύο λόγοι είναι χαμηλοί, το ρίσκο του να έχουμε ελλείψεις έχει μικρότερο κόστος και σαν αποτέλεσμα οι πολιτικές κλίνουν σε μικρότερες πολιτικές παραγγελίας, για να μετριάσουμε την πιθανότητα μικρών χρόνων ανεφοδιασμού και μεγάλες ποσότητες από απώλητα προϊόντα.

4.7 Ανασκόπηση και επεκτάσεις

Οι δύο βασικές περιοχές της διατριβής Προσαρμοστική δειγματοληψία υπό ελλιπή πληροφόρηση και επιπλέον περιορισμούς και Βέλτιστες πολιτικές παραγγελιών σε σύστημα αποθεμάτων με πεπερασμένη χωρητικότητα και ανταλλαξιμότητα στη ζήτηση έχουν όπως αναφέραμε λογική συσχέτιση. Συνεπώς τα προβλήματα που μελετήθηκαν στην πρώτη περιοχή είναι γενικά και περιλαμβάνουν εφαρμογές εκτός των άλλων και στην διαχείριση αποθεμάτων με άγνωστη κατανομή ζήτησης, στα οποία ζητείται μια προσαρμοστική πολιτική παραγγελιών που μεγιστοποιεί το αναμενόμενο κέρδος ανά μονάδα χρόνου. Επικεντρωθήκαμε λοιπόν σε ένα συγκεκριμένο πρόβλημα αυτής της περιοχής, στο οποίο γίνονται παραγγελίες για δύο προϊόντα ταυτόχρονα σε περιβάλλον εφημεριδοπώλη (τα προϊόντα διατηρούνται για μια περίοδο μόνο) και μερική αμφίδρομη υποκατάσταση



Σχήμα 4.4: Επίδραση ανεφοδιασμού τυχαίου χρόνου, Σενάριο 1: $\lambda_1 = \lambda_2 = 20, T = 1, p_1 = 50, c_1 = 10, h_1 = 0, p_2 = 20, c_2 = 4, h_2 = 0, p_{12} = p_{21} = 0.4, T = 1(\text{fixed})/T \sim \text{Exp}(1)(\text{rand})$



Σχήμα 4.5: Επίδραση ανταλλαξιμότητας, Σενάριο 2: $\lambda_1 = \lambda_2 = 20, T = 1, p_1 = 50, c_1 = 30, h_1 = 0, p_2 = 20, c_2 = 12, h_2 = 0, p_{12} = p_{21} = 0.4, T = 1(\text{fixed})/T \sim \text{Exp}(1)(\text{rand})$

ζήτησης. Για το πρόβλημα αυτό αποδείξαμε ιδιότητες της βέλτιστης πολιτικής για την περίπτωση που οι παράμετροι του προβλήματος (κατανομές ζήτησης και πιθανότητες υποκατάστασης) είναι πλήρως γνωστές. Τα αποτελέσματα αυτά παρέχουν το απαραίτητο υπόβαθρο για την εφαρμογή και των προσαρμοστικών μεθόδων δειγματοληψίας, όταν κάποιες ή όλες από τις παραπάνω παραμέτρους δεν είναι γνωστές. Συνεπώς μια επέκταση του προβλήματος μας είναι να έχουμε άγνωστες παραμέτρους. Για τη μελέτη του προβλήματος αυτού υπό ελλιπή πληροφόρηση, είναι απαραίτητο πρώτα να προσδιοριστεί η βέλτιστη πολιτική παραγγελιών όταν όλες οι παράμετροι του προβλήματος είναι γνωστές. Πράγμα που προσδιορίσαμε σε αυτό το Κεφάλαιο και επιπλέον χρειάζεται να χρησιμοποιήσουμε προσαρμοστικές πολιτικές δειγματοληψίας ώστε να καταφέρουμε να εκτιμήσουμε τις άγνωστες παραμέτρους του προβλήματος.

Βιβλιογραφία

- [1] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multi-armed bandit. *Machine Learning*, **22(1)**, 222-255, (1997).
- [2] M. N. Katehakis and H. Robbins. Sequential choice from several populations. *Proc.Natl.Acad.Sci. USA*, **92**, 8584-8585, (1995).
- [3] S. R. Kulkarni and G. Lugosi. Finite-time lower bounds for the two-armed bandit problem. *IEEE Transactions on Automatic Control*, **45**, 711-714, (2000).
- [4] T. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Adv. App. Math.*, **6**, 4-22, (1985).
- [5] O. Madani, D. Lizotte, and R. Greiner. The budgeted multi-armed bandit problem. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, **3120**, 643-645, (2004).
- [6] H. Pezeshk and J. Gittins. Sample size determination in clinical trials. *Student*, **3(1)**, 19-26, (1999).
- [7] A. Poznyak, K. Nazim, and E. Gomez. Self-Learning Control of Finite Markov Chains. *CRC Press, New York*, (2000).
- [8] H. Robbins. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Monthly*, **58**, 527-536, (1952).
- [9] Y. G. Wang. Gittins indices and constrained allocation in clinical trials. *Biometrika*, **78**, 101-111, (1991).

- [10] R. Agrawal, D. Teneketzis, and V. Anantharam. Asymptotically efficient adaptive allocation schemes for controlled i.i.d. processes : Finite parameter space. *IEEE Trans. AC*, **34**, 258-267, (1989).
- [11] O. Avner and S. Mannor. Stochastic bandits with pathwise constraints. (2012).
- [12] R. Bellman. Dynamic Programming. *Princeton University Press*, (1957).
- [13] A. N. Burnetas and O. A. Kanavetas. Adaptive policies for sequential sampling under incomplete information and a cost constraint. N.J. Daras(ed.), *Applications of Mathematics and Informatics in Military Science*, Springer, 97-112, (2012).
- [14] A. N. Burnetas and M. N. Katehakis. Optimal adaptive policies for sequential allocation problems. *Adv. App. Math.*, **17**, 122-142, (1996).
- [15] S. Dayanik, W. Powell, and K. Yamazaki. Index policies for discounted bandit problems with availability constraints. *Advances in Applied Probability*, **40(2)**, 377-400, (2008).
- [16] E. Denardo, E. Feinberg, and U. Rothblum. The multi-armed bandit, with constraints. *Annals of Operations Research*, **208(1)**, 37-62, (2013).
- [17] J. C. Gittins. Bandit processes and dynamic allocation indices (with discussion). *J. Roy. Stat. Soc. Ser. B*, **41**, 335-340, (1979).
- [18] J. C. Gittins and K. D. Glazebrook. On Bayesian models in stochastic scheduling. *J. App. Prob.*, **14**, 556-565, (1977).
- [19] S. Guha and K. Munagala. Multi-armed bandits with metric switching costs. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, **5556 LNCS (PART 2)**, 496-507, (2009).
- [20] T. He, A. Anandkumar, and D. Agrawal. Index-based sampling policies for tracking dynamic networks under sampling constraints. 1233-1241, (2011).

- [21] M. N. Katehakis and A. F. Veinott Jr. The multi- armed bandit problem: Decomposition and computation. *Math. Oper. Res.*, **12**, 262-268, (1987).
- [22] M. Khouja, A. Mehrez, and G. Rabinowitz. A two-item newsboy problem with substitutability. *International Journal of Production Economics*, **44 (3)**, 267-275, (1996).
- [23] I. Krommyda, K. Skouri, and I. Konstantaras. Optimal ordering quantities for substitutable products with stock-dependent demand. *Applied Mathematical Modelling*, (2014).
- [24] M. Parlur and S. Goyal. Optimal ordering decisions for two substitutable products with stochastic demands. *OPSEARCH*, **21**, 1-15, (1984).
- [25] M. Salameh, A. Yassine, B. Maddah, and L. Ghaddar. Joint replenishment model with substitution. *Applied Mathematical Modelling*, **38**, 3662-3671, (2014).
- [26] E. Stavroulaki. Inventory decisions for substitutable products with stock dependent demand. *Int. J. Production Economics*, **129**, 65-78, (2011).
- [27] Kitaev, Mikhail Yu, Rykov, V. Vladimir . Controlled queueing systems. *CRC press*, **book** (1995).