

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΣΥΝΕΧΕΙΑ ΥΓΕΙΟΝΟΜΙΚΗΣ
ΠΕΡΙΘΑΛΨΗΣ**

Στέλλα Β. Αντωνίου

Διπλωματική εργασία

που υποβλήθηκε στο τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
Απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική

Πειραιάς
Απρίλιος 2006

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΣΥΝΕΧΕΙΑ ΥΓΕΙΟΝΟΜΙΚΗΣ
ΠΕΡΙΘΑΛΨΗΣ**

Στέλλα Β. Αντωνίου

Διπλωματική εργασία

που υποβλήθηκε στο τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
Απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική

Πειραιάς
Απρίλιος 2006

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από την ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Μάρκος Κούτρας (επιβλέπων)
- Μιλτιάδης Νεκτάριος
- Μαρία Κατέρη

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

UNIVERSITY OF PIRAEUS



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

CONTINUITY OF MEDICAL CARE

By

Stella V. Antoniou

MSc Dissertation

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

Piraeus, Greece
April 2006

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

*Στον Κώστα, που αποτέλεσε
και θα αποτελεί πάντα πηγή έμπνευσης*

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Ευχαριστίες

Θα ήθελα να ευχαριστήσω στο σημείο αυτό τον επιβλέποντα καθηγητή μου κ. Κούτρα Μάρκο για τη συνεργασία που είχαμε όλο το διάστημα συγγραφής της παρούσας εργασίας, καθώς και τα μέλη της Τριμελούς Επιτροπής κ. Νεκτάριο Μιλτιάδη και κ. Κατέρη Μαρία για τις πολύτιμες και εύστοχες παρατηρήσεις τους. Τέλος, αισθάνομαι την ανάγκη να ευχαριστήσω όλους όσους μου συμπαραστάθηκαν από το οικογενειακό, φιλικό και εργασιακό μου περιβάλλον.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Περίληψη

Η συνέχεια της υγειονομικής περίθαλψης είναι ένας σημαντικός παράγοντας για την ποιότητα της περίθαλψης που παρέχεται στους ασθενείς και τις τελευταίες δεκαετίες το ενδιαφέρον γύρω από αυτήν έχει αυξηθεί σημαντικά, κυρίως στις ανεπτυγμένες χώρες. Στην εργασία αυτή, αφού δοθεί ένας ορισμός για τη συνέχεια της περίθαλψης και αναπτυχθούν διάφορες παράμετροί της, θα παρουσιαστούν οι κυριότεροι δείκτες που χρησιμοποιούνται για τη μέτρησή της (*UPC, SECON, NOP, COC*), καθώς και πώς ο καθένας από αυτούς μας βοηθάει να εξετάσουμε κατά πόσο η περίθαλψη ενός ασθενούς είναι συνεχής ή όχι. Στη συνέχεια, θα μελετηθούν οι ακριβείς κατανομές τριών δεικτών χρησιμοποιώντας την τεχνική εμφύτευσης Μαρκοβιανής αλυσίδας των Fu and Koutras (1994). Τέλος, θα επιχειρηθεί η εφαρμογή όσων παρουσιάστηκαν σε δεδομένα που συγκεντρώθηκαν από πραγματικές περιπτώσεις ασθενών και θα εξεταστεί κατά πόσο οι ακριβείς κατανομές των δεικτών αντανακλούν την πραγματικότητα ενός συστήματος υγείας.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Abstract

Continuity of medical care is an important factor for the quality of care provided to patients. During the past decades, there is an increasing interest around continuity of care, particularly in developed countries. In this dissertation, at first, we shall provide a definition of continuity of care, along with its various parameters. Then, the most common indices (*UPC*, *SECON*, *NOP*, *COC*) which are used to measure continuity of care will be presented, as well as in which way each one of them helps us investigate whether medical care provided to a patient is continuous or not. Furthermore, we shall examine the exact distributions of three indices using the Markov chain imbedding technique by Fu and Koutras (1994). Finally, we shall attempt to implement what was previously presented to data derived from actual patient cases and we will examine at which extent the exact distributions of the indices reflect the reality of a medical care system.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Περιεχόμενα

Περίληψη	ix
Abstract	xi
Πρόλογος	xv
1. Εισαγωγή	1
1.1 Εισαγωγικές έννοιες	1
1.2 Ιστορική αναδρομή - είδη της συνέχειας της περιθάλψης	3
1.3 Οι συνιστώσες ενός μοντέλου συνέχειας	5
2. Τα κυριότερα μέτρα της συνέχειας της περιθάλψης	11
2.1 Εισαγωγή	11
2.2 Τα κυριότερα μέτρα της συνέχειας της περιθάλψης	12
2.3 Παράδειγμα	17
3. Η ακριβής κατανομή του <i>SECON</i>	21
3.1 Εισαγωγή	21
3.2 Η ακριβής κατανομή	22
3.3 Εκτίμηση του μέτρου ακολουθιακής συνέχειας	29
4. Η ακριβής κατανομή του <i>UPC</i>	35
4.1 Εισαγωγή	35
4.2 Η ακριβής κατανομή	36
4.3 Παράδειγμα	40
4.4 Παρατηρήσεις	44
5. Η ακριβής κατανομή του <i>NOP</i>	47
5.1 Εισαγωγή	47
5.2 Προκαταρκτικές έννοιες	47
5.3 Η ακριβής κατανομή	50
5.4 Παράδειγμα	53
5.5 Παρατηρήσεις	57

6. Case study σε πραγματικά δεδομένα	59
6.1 Εισαγωγή	59
6.2 Περιγραφική ανάλυση	60
6.3 Σύγκριση εμπειρικών τιμών με τις ακριβείς κατανομές	65
6.4 Συμπεράσματα	81
Βιβλιογραφία	83

Πρόλογος

Το ζήτημα της συνέχειας της υγειονομικής περίθαλψης αποτελεί ζήτημα των τελευταίων δεκαετιών στις ανεπτυγμένες χώρες επειδή ο κατακερματισμός και η πολυσυλλεκτικότητα της περίθαλψης είναι χαρακτηριστικά της άκρας εξειδίκευσης της εποχής μας. Λόγω του μεγάλου αριθμού των παροχών υγειονομικής περίθαλψης, καθώς και των διαφορετικών ειδών συστημάτων υγείας, σε αντίθεση με τους προηγούμενους αιώνες όπου υπήρχε ένας γιατρός για όλες τις ασθένειες, σήμερα γίνεται πολλή συζήτηση γύρω από τη διατήρηση της ποιότητας της περίθαλψης. Η συνέχεια της περίθαλψης είναι ένας από τους παράγοντες που συμβάλλουν προς την κατεύθυνση αυτή.

Για τους παραπάνω λόγους προέκυψε η ανάγκη μέτρησης της συνέχειας της περίθαλψης. Έχουν προταθεί κατά καιρούς πολλοί δείκτες και ο καθένας από αυτούς φανερώνει και μια διαφορετική άποψη για τη μέτρηση της συνέχειας, αναλόγως με το τι θεωρεί ο κάθε επιστήμονας καθοριστικό παράγοντας για το αν μια περίθαλψη είναι συνεχής ή όχι. Στο Κεφάλαιο 2 παρουσιάζονται οι 5 δείκτες συνέχειας που χρησιμοποιούνται περισσότερο: η μαθηματικής τους έκφραση, το πεδίο τιμών τους, η διαφορετική πληροφορία που μας δίνει ο καθένας.

Τα Κεφάλαια 3, 4 και 5 αναφέρονται στις ακριβείς κατανομές τριών από τους δείκτες που παρουσιάστηκαν στο Κεφάλαιο 2 και βασίζονται και τα τρία κεφάλαια στις μελέτες της W.Y. Wendy Lou, η οποία τα τελευταία χρόνια ασχολείται συστηματικά με το αντικείμενο. Οι ακριβείς κατανομές εξάγονται με την τεχνική εμφύτευσης Μαρκοβιανής αλυσίδας των Fu and Koutras (1994). Σε κάθε κεφάλαιο, εκτός από τους μαθηματικούς τύπους και τις αποδείξεις τους, υπάρχουν και μερικά παραδείγματα, καθώς και παρατηρήσεις για τα αποτελέσματα.

Στο 6^ο και τελευταίο Κεφάλαιο έχουμε συγκεντρώσει στοιχεία από επισκέψεις 472 ασθενών σε 5 παθολόγους του Ιδρύματος Κοινωνικών Ασφαλίσεων. Αρχικά, γίνεται μια περιγραφική ανάλυση των δεδομένων με τη βοήθεια των δεικτών που παρουσιάστηκαν στο Κεφάλαιο 2 και εξετάζεται κατά πόσο η περίθαλψη των ασθενών είναι συνεχής ή διεσπαρμένη σε πολλούς γιατρούς. Στη συνέχεια, και με βάση τις ακριβείς κατανομές των τριών δεικτών, όπως αυτές παρουσιάστηκαν στα Κεφάλαια 3, 4 και 5, θεωρούμε 6 διαφορετικά μοντέλα που τα καθένα περιγράφει μια διαφορετική κινητικότητα των ασθενών από τον ένα γιατρό στον άλλο. Αφού πάρουμε τις τιμές των ακριβών κατανομών, τις

συγκρίνουμε με τα εμπειρικά δεδομένα που συγκεντρώσαμε και εξετάζουμε ποιο από τα 6 μοντέλα προσεγγίζει περισσότερο την πραγματικότητα. Στην τελευταία ενότητα του Κεφαλαίου 6 υπάρχουν τα συμπεράσματα.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κ Ε Φ Α Λ Α Ι Ο 1

Εισαγωγή

1.1 Εισαγωγικές έννοιες

Η ανάπτυξη της μεσαιάς τάξης τον 18^ο αιώνα επέτρεψε στους γιατρούς να κερδίζουν αρκετά χρήματα για να ζουν και να συντηρούν τα δικά τους ιατρεία σε μια περιοχή για μια μεγάλη χρονική περίοδο. Αυτή η σταθερότητα των γιατρών στην ίδια περιοχή μεγάλωσε το αίσθημα ευθύνης τους για τους ασθενείς τους και η άσκηση της πολιτικής από τους πλούσιους που είχε ως στόχο τον έλεγχο και την κυριαρχία επί των πολιτών οδήγησε στην ευρεία εξάπλωση των νοσοκομείων. Η συνέχεια της περίθαλψης τότε ενισχύθηκε από τη σταθερότητα του πληθυσμού και του γιατρού αλλά βασιζόταν στην καλή μνήμη του γιατρού. Οι τελευταίοι σπάνια διατηρούσαν αρχεία των ασθενών και ακόμα και αν διατηρούσαν, τα χρησιμοποιούσαν περισσότερο ως βοηθήματα για τη θεραπευτική αγωγή παρά ως ιατρικά αρχεία, κάτι που συνεχίζει να υπάρχει ακόμη και σήμερα σε πολλά μέρη της γης. Το ατομικό ιατρικό αρχείο έγινε γενικά αποδεκτό μόνο στα τέλη του 19^{ου} αιώνα.

Τι είναι όμως η συνέχεια της περίθαλψης; Στα περισσότερα λεξικά η συνέχεια ορίζεται ως «μία αδιάκοπη σειρά», «μια συνεχής πορεία» ή «ένα αδιάσπαστο σύνολο». Όσον αφορά στην ιατρική περίθαλψη, η συνέχεια εμφανίζεται όχι μόνο ως μια μετρήσιμη ακολουθία συναντήσεων, αλλά και ως μια στάση γενικότερα.

Με τον όρο «συνέχεια της περίθαλψης» εννοούμε μία σχέση μεταξύ ενός γιατρού ή άλλου παροχέα υγείας και ενός ασθενούς, η οποία διατηρείται με το πέρασμα του χρόνου, σε περιόδους ασθένειας και υγείας και η οποία παραμένει όταν το άτομο νοσηλεύεται, επισκέπτεται το ιατρείο ή το νοσοκομείο ή δέχεται ιατρική περίθαλψη σε κάποιο άλλο ίδρυμα (π.χ. ιδιωτική κλινική). Η συνέχεια της περίθαλψης έχει ως συνέπεια, όταν κάποιος ασθενήσει, ο παροχέας να γνωρίζει ήδη το ιατρικό ιστορικό του, καθώς και τις συνθήκες της καθημερινής του ζωής.

Στις μέρες μας, το ενδιαφέρον για τη συνέχεια της περίθαλψης μοιάζει να έχει αυξηθεί και αυτό φαίνεται να έχει άμεση σχέση με τον κατακερματισμό της ιατρικής περίθαλψης, τη βιομηχανοποίηση και την κινητικότητα τόσο των ασθενών όσο και των γιατρών.

Φαίνεται να υπάρχει μια ποικιλία απόψεων όσον αφορά τη φύση και τον ορισμό της συνέχειας της περίθαλψης με επακόλουθες δυσκολίες στην ερμηνεία της σχετικής σπουδαιότητας των μελετών που ασχολήθηκαν με το αντικείμενο. Οι περισσότεροι στον τομέα της Οικογενειακής Ιατρικής και μερικοί στον τομέα της Παιδιατρικής ισχυρίζονται ότι η συνέχεια εξηγείται από το αίσθημα συνεχούς ευθύνης του γιατρού για το άτομο και την οικογένεια και όχι τόσο από ενδιαφέρον για συγκεκριμένα προβλήματα αυτά καθαυτά. Φαίνεται ότι αυτό το αίσθημα προσωπικής ευθύνης αυξάνεται καθώς βελτιώνεται η συνέχεια της περίθαλψης. Έχει επίσης αναφερθεί ότι ο ασθενής θα πρέπει να δεχτεί ορισμένες ευθύνες έτσι ώστε να διατηρηθεί το «συμβόλαιο της συνέχειας», δηλαδή η συνεχής σχέση γιατρού-ασθενούς. Η φύση του συμβολαίου αυτού είναι ένα βασικό ζήτημα για την οικογενειακή ιατρική και λήγει μόνο με αμοιβαία συμφωνία ή κατόπιν αποφάσεως μίας εκ των δύο πλευρών.

Η συνέχεια της περίθαλψης είναι ένας σημαντικός παράγοντας για:

- την ποιότητα της περίθαλψης
- τον έλεγχο του κλινικού κόστους
- την ικανοποίηση του ασθενούς και
- τη θέσπιση εθνικής ασφάλισης υγείας.

Είναι επίσης σημαντική για τους ασθενείς, το παραϊατρικό προσωπικό και τους κοινωνικούς επιστήμονες.

Ποια είναι όμως τα πλεονεκτήματα που έχουν αποδοθεί στη συνέχεια της περίθαλψης για να υποστηρίξουν ότι αυτή είναι σημαντική για την περίθαλψη του ασθενούς; Τα σημαντικότερα εξ αυτών παρατίθενται παρακάτω:

- Βελτίωση της σχέσης γιατρού – ασθενούς
- Ευρύτερη γνώση και αυξημένο ενδιαφέρον για τον ασθενή από τον γιατρό και από το ιατρικό προσωπικό.
- Ενίσχυση της εμπιστοσύνης και της συμπάθειας προς τον γιατρό και ευκολότερη «αποκάλυψη» προσωπικών πληροφοριών από τον ασθενή.
- Μεγαλύτερη συμμόρφωση με τις ιατρικές οδηγίες.

- Μειωμένα ποσοστά εισαγωγής σε νοσοκομεία, επεισοδίων ασθένειας και λιγότερες εξετάσεις εργαστηρίου.
- Αυξημένη ικανοποίηση του ασθενούς.
- Λιγότερα χαμένα ραντεβού και λιγότερη ώρα αναμονής για τον ασθενή.
- Καλύτερη αναγνώριση του υπάρχοντος ή των προηγούμενων ιατρικών προβλημάτων.

Λόγω των προαναφερθέντων πλεονεκτημάτων της, η συνέχεια της περίθαλψης έχει γίνει μια αρκετά κρίσιμη έννοια στη συζήτηση της πολιτικής του συστήματος της ιατρικής περίθαλψης. Το ενδιαφέρον αυτό σε μεγάλο βαθμό έχει προωθηθεί από την κοινή γνώμη. Στη Σουηδία για παράδειγμα, το ζήτημα της συνέχειας ήταν ιδιαίτερος έντονο, καθώς ενισχύθηκε από τα ευρήματα αρκετών μελετών σχετικά με τις κοινωνικές συμπεριφορές. Έτσι, η συνέχεια έγινε ένας από τους εξέχοντες σκοπούς του σχεδίου για τις ιατρικές υπηρεσίες της Σουηδίας κατά την δεκαετία του 1980.

1.2 Ιστορική αναδρομή – Είδη της συνέχειας της περίθαλψης

Για να εδραιώσουμε την αξία της συνέχειας στο σύστημα ιατρικής περίθαλψης, είναι απαραίτητο πρώτα να την ορίσουμε και ύστερα να τη μετρήσουμε. Ο πιο περιεκτικός θεωρητικός ορισμός δόθηκε από τον Hennen (1975) ο οποίος όρισε πέντε διαστάσεις της συνέχειας : χρονολογική, γεωγραφική, διεπιστημονική, διαπροσωπική και ενημερωτική. Σε ένα πιο λειτουργικό επίπεδο, αρκετοί συγγραφείς έχουν ορίσει τη συνέχεια της περίθαλψης έτσι ώστε να μπορεί να εφαρμοστεί μια αποτελεσματική μέτρησή της.

Η συνέχεια της περίθαλψης ορίζεται σύμφωνα με τους Bass και Windle (1972) ως η σχέση μεταξύ παρελθοντικής και παρούσας θεραπείας σύμφωνα με τις θεραπευτικές ανάγκες του ασθενούς.

Οι Gordis και Markovitz (1971) μελέτησαν τη συνέχεια με την έννοια της διαθεσιμότητας του παιδίατρου όλες τις ώρες, με κάποια συνδρομή από τις νοσοκόμες και τους κοινωνικούς λειτουργούς.

Σύμφωνα με τους Breslau και Reeb (1975), η συνέχεια είναι ο βαθμός στον οποίο ένας μόνο γιατρός διαχειρίζεται τις ανάγκες ενός ασθενούς. Όσο περισσότερες επισκέψεις

πραγματοποιεί ο ασθενής σε ένα μόνο γιατρό, τόσο περισσότερο συνεχής θεωρείται η περίθαλψη.

Από την άλλη πλευρά, οι Shorr και Nutting (1976) ορίζουν τη συνέχεια ως το ρυθμό μετάβασης από το ένα σημαντικό κλινικό στοιχείο στο άλλο: εξέταση, τελική διαγνωστική αξιολόγηση, θεραπεία και υποστήριξη.

Ο Shortell (1976) δηλώνει ότι η συνέχεια της υγειονομικής περίθαλψης είναι ο βαθμός στον οποίο οι ιατρικές υπηρεσίες λαμβάνονται σε μια συντονισμένη και αδιάκοπη ακολουθία γεγονότων σύμφωνα με τις ανάγκες υγειονομικής περίθαλψης του ασθενούς.

Βασιζόμενοι στον παραπάνω ορισμό του Shortell μπορούμε να δώσουμε έναν πιο λειτουργικό και πιο περιοριστικό ορισμό, δηλαδή ότι η συνέχεια είναι ο βαθμός στον οποίο ο ίδιος παροχέας χρησιμοποιείται κατά την διάρκεια μιας σειράς επισκέψεων.

Με βάση κάποιους από τους παραπάνω ορισμούς και μεθόδους μέτρησης, οι μελέτες της συνέχειας προς το παρόν εμπίπτουν σε δύο κατηγορίες. Η πρώτη είναι εκείνη στην οποία μια κατάσταση «ασυνέχειας» εντοπίζεται και μετατρέπεται σε μια κατάσταση «συνέχειας». Απόδειξη της αλλαγής αυτής είναι ότι ορισμένες μεταβλητές που σχετίζονται με την περίθαλψη έχουν βελτιωθεί, πολλές φορές σε σχέση με μια ομάδα ελέγχου (*control group*). Η δεύτερη κατηγορία περιλαμβάνει μελέτες που προσπαθούν να τεκμηριώσουν την ύπαρξη και την πληρότητα μιας αναγνωρισμένης κατάστασης συνέχειας, συχνά με τη σιωπηρή υπόθεση ότι τα αποτελέσματα για την υγεία είναι καλύτερα στο σύστημα αυτό. Γενικά, οι μετρήσεις έχουν στραφεί προς τις παρακάτω καταστάσεις:

1. Συνέχεια περίθαλψης που δίνεται στον ασθενή και την οικογένεια από τον παροχέα ή από την ιατρική ομάδα. Αυτό μετριέται ως επισκέψεις ασθενών, περιπτώσεις νόσου ή ως μια αναλογία των προγραμματισμένων ή των μη προγραμματισμένων επισκέψεων.

2. Συνέχεια περίθαλψης ολόκληρου του συστήματος υγείας, δηλ διερεύνηση των ποσοστών προγραμματισμένων αλλά μη πραγματοποιηθέντων ραντεβού, πλήθος των εξετάσεων που πραγματοποιούνται πολλές φορές, εξετάσεις από γιατρό, πλήθος των πηγών περίθαλψης, ποσοστό της απόσυρσης ασθενών από προγράμματα σκοπός των οποίων είναι η προστασία ή η πρόληψη.

3. Εκτίμηση των στάσεων απέναντι στη συνέχεια της περίθαλψης. Η στάση των ασθενών, του προσωπικού και των γιατρών έχουν μελετηθεί κατόπιν της εφαρμογής των σχεδίων συνέχειας.

1.3 Οι συνιστώσες ενός μοντέλου συνέχειας

Η συνέχεια εξαρτάται από την ποσότητα της προηγούμενης γνώσης που διαθέτουν τα **στοιχεία** που εμπλέκονται στην υγειονομική περίθαλψη. Αυτά τα στοιχεία είναι δύο ειδών:

α. οι καταναλωτές, που αποτελούνται από τους ασθενείς, τις οικογένειες ή ακόμη και από μεγαλύτερες ομάδες και

β. οι παροχείς, οι οποίοι προσφέρουν περίθαλψη.

Στην περίπτωση που η συνέχεια έχει εξασφαλιστεί, είτε ο καταναλωτής είτε και οι δύο πλευρές διαθέτουν κάποια προηγούμενη γνώση ο ένας για τον άλλο και θα χρησιμοποιήσουν τη γνώση αυτή σε μια μεταξύ τους συνάντηση.

Η συνάντηση μεταξύ των δύο πλευρών μπορεί να είναι είτε **άμεση** είτε **έμμεση**. Η άμεση συνάντηση μπορεί να γίνει οπτικά, λεκτικά ή με χρήση γραπτού λόγου, π.χ. πρόσωπο με πρόσωπο, μέσω τηλεφώνου ή μέσω ιατρικών αρχείων κλπ. Για την έμμεση συνάντηση χρησιμοποιούνται οι ίδιες μέθοδοι επικοινωνίας αλλά τη συνάντηση τώρα την αναλαμβάνει ένας μεσάζοντας, π.χ. μία μητέρα που συζητά με μια νοσοκόμα το πρόβλημα του παιδιού της το οποίο δεν είναι παρόν ή μια νοσοκόμα που δίνει στο γιατρό ένα μήνυμα από κάποιον ασθενή.

Η βάση της γνώσης που αφορά τους καταναλωτές και τους παροχείς μπορεί να είναι **καταγεγραμμένη** ή μη **καταγεγραμμένη**. Η πρώτη μπορεί να αποκτηθεί άμεσα και από τις δύο πλευρές ή να αποκτηθεί από «δεύτερο χέρι» (π.χ. από την αναφορά τρίτων). Οι γιατροί γνωρίζουν ότι σημαντικό μέρος μη καταγεγραμμένης γνώσης είναι διαθέσιμο και διαδίδεται από την εφαρμογή της παθολογίας, παρόλο που δεν υπάρχει τεκμηρίωση του γεγονότος αυτού. Ο καταναλωτής συνήθως δε διαθέτει καταγεγραμμένη γνώση για τον παροχέα, αλλά μπορεί να έχει μη καταγεγραμμένη άμεση ή από δεύτερο χέρι γνώση για το άτομο αυτό (π.χ. η φήμη για την ευγένεια, την ικανότητα και την προθυμία για κατ'οίκον επισκέψεις ενός γιατρού κλπ). Αντίθετα, ο παροχέας έχει συνήθως καταγεγραμμένη γνώση για τον καταναλωτή – ιατρικά αρχεία, εργαστηριακές εξετάσεις. Επιπλέον, ο παροχέας είναι δυνατό να έχει σημαντική μη καταγεγραμμένη γνώση από πολλές πηγές, άμεσες και μη. Η πληρότητα και η χρήση της γνώσης αυτής έχει άμεσες συνέπειες στην ποιότητα και το επίπεδο της συνέχειας. Θα μπορούσαμε να θεωρήσουμε δεδομένο ότι όσο μεγαλύτερη είναι η βάση της γνώσης, τόσο μεγαλύτερη είναι η συνέχεια της περίθαλψης για τον καταναλωτή.

Οι συναντήσεις πραγματοποιούνται μέσα στο «περιβάλλον συνέχειας» (*continuity environment*). Το περιβάλλον αυτό αποτελείται από ένα σύνολο παραγόντων που επηρεάζουν την ποιότητα και το επίπεδο της συνέχειας ανάμεσα στα δύο στοιχεία. Βασιζόμενοι στην κεντρική ιδέα του Hennen (1975) το περιβάλλον αυτό απαρτίζεται από τις παρακάτω διαστάσεις:

Τη χρονολογική διάσταση, η οποία περιλαμβάνει τις ιατρικές υπηρεσίες που παρέχονται στους ασθενείς και την οικογένεια κατά τη διάρκεια χρονικών περιόδων που καλύπτουν κύκλους αλλαγής της ζωής και κρίσεις της ζωής. Αυτό προϋποθέτει περίθαλψη για άτομα όλων των ηλικιών. Πληροφορίες για το άτομο και την οικογένεια που συγκεντρώνονται με την πάροδο του χρόνου, προσφέρουν μεστή γνώση απαραίτητη τόσο για τη διάγνωση και τη διαχείριση του ασθενούς όσο και για τη δημιουργία μιας βάσης επιστημονικής μελέτης.

Τη γεωγραφική διάσταση, η οποία υποδεικνύει την τοποθεσία όπου παρέχεται η περίθαλψη. Ο Hennen υποστηρίζει ότι η συνέχεια εδραιώνεται όταν ένας γιατρός προσφέρει μια ολοκληρωμένη ιατρική επικοινωνία ανεξαρτήτως τοποθεσίας.

Την διεπιστημονική διάσταση, η οποία περιγράφεται ως η ικανότητα του γιατρού να υπερβεί τα σύνορα του κύριου συστήματος, να διαχειριστεί και να ενώσει ένα ευρύ φάσμα νόσων, καθώς και κοινωνικών προβλημάτων ή συμπεριφοράς. Επίσης η διάσταση αυτή υποδεικνύει την προθυμία των καταναλωτών να αναζητήσουν ολιστική περίθαλψη σε μία μόνο πηγή.

Τη διαπροσωπική διάσταση, η οποία περιλαμβάνει τη συνέχεια στη σχέση γιατρού-ασθενούς και εκείνους τους συνδέσμους που βοηθούν τον ασθενή στις διάφορες πλευρές του συστήματος ιατρικής περίθαλψης (π.χ. παραϊατρικό προσωπικό, υπαλλήλους υποδοχής και νοσοκομειακό προσωπικό). Μέσα σε αυτή τη διάσταση, ο παροχέας ίσως νιώθει ένα αίσθημα συνεχούς ευθύνης απέναντι στον ασθενή, την οικογένεια και την κοινότητα, το οποίο αίσθημα υπερβαίνει τις συμφωνημένες δεσμεύσεις και το ωράριο εργασίας.

Την πληροφοριακή διάσταση, η οποία υποστηρίζει τη συνέχεια που παρέχεται στον ασθενή με πληροφοριακά συστήματα. Αυτό στηρίζεται σε επαρκή ιατρικά αρχεία που περιέχουν επεισόδια ασθένειας, υποστήριξη, σχέδια χειρισμού της υγείας καθώς και αποδοτικές τηλεπικοινωνίες, καλά συστήματα αναφοράς, και ανατροφοδότηση (*feedback*) από θαλάμους επειγόντων περιστατικών και νοσοκομεία. Μεγάλη είναι η προθυμία των καταναλωτών για παροχή πληροφοριών ή την υπόδειξη καταλλήλων πηγών απ' όπου μπορούν να αντληθούν στοιχεία.

Τη διάσταση της διαθεσιμότητας, η οποία εξασφαλίζει τη συνέχεια που εδραιώνεται από τα άνετα ιατρεία, από τα αποτελεσματικά συστήματα κανονισμού των ραντεβού, από την πρόβλεψη για περίθαλψη εκτός ωραρίου και από την ευκολία πρόσβασης στις ιατρικές συμβουλές.

Τη διάσταση της σταθερότητας, η οποία συνδέεται με τη σταθερότητα της κοινότητας μέσα στην οποία παρέχεται η περίθαλψη και τη σχετική σταθερότητα της οικογένειας και του ατόμου. Επίσης σχετίζεται με τη σταθερότητα ή την κινητικότητα του παροχέα.

Στο Σχήμα 1 (βλέπε επόμενη σελίδα) παρατίθεται ένα μοντέλο με το οποίο μπορεί να μετρηθεί η συνέχεια. Περιλαμβάνει το στοιχείο του παροχέα, το στοιχείο του καταναλωτή, τον τύπο της συνάντησης, τη βάση της γνώσης και τη διάσταση του περιβάλλοντος. Το ιδανικό άμεσο μέτρο της συνέχειας θα ήταν το μέγεθος της βάσης της γνώσης, αλλά η ανάλυση των άλλων συνιστωσών του μοντέλου είναι πιο εύκολο να επιτευχθεί και πιθανώς δίνει ένα καλό έμμεσο μέτρο της συνέχειας της περίθαλψης. Οι συνιστώσες αυτές μπορούν να εκτιμηθούν μαθηματικά, π.χ. ο αριθμός των συναντήσεων, ο αριθμός των παροχέων που εμπλέκονται στην περίθαλψη ενός ασθενούς, ο αριθμός των διαφορετικών τοποθεσιών των συναντήσεων, κλπ.

Καθώς εξετάζουμε την επίσκεψη του καταναλωτή στον παροχέα προκύπτει μια επιπλέον δυσκολία, αφού τα δύο αυτά στοιχεία αλληλεπιδρούν σε πολλά και μεταβαλλόμενα επίπεδα έντασης της συνέχειας. Αυτό μπορεί να περιγραφεί από το παράδειγμα όπου το πρώτο και πιο έντονο επίπεδο της συνέχειας εδραιώνεται από τη συνάντηση μεταξύ του ασθενούς και του γιατρού που του ανατίθεται το δεύτερο επίπεδο συνέχειας πραγματοποιείται όταν υπάρχει επαφή μεταξύ του ασθενούς και ενός γιατρού από την ίδια ιατρική ομάδα. Το τρίτο επίπεδο συνέχειας πραγματοποιείται όταν ένας γιατρός από μια άλλη ιατρική ομάδα με την ίδια όμως σύνθεση συναντά τον ασθενή. Το πρώτο και το δεύτερο επίπεδο είναι καταστάσεις όπου «προσωπικοί» παροχείς έχουν μια υψηλή υποχρέωση στη συνέχεια της περίθαλψης για το άτομο. Οι «απρόσωποι» παροχείς είναι εκείνοι που έχουν μια χαμηλή υποχρέωση στη συνέχεια της περίθαλψης για τον ασθενή και τείνουν να είναι άνθρωποι που δεν έχουν αίσθημα συνεχούς ευθύνης. Μπορεί να παρέχουν συνέχεια σε ένα περιορισμένο πεδίο της ιατρικής αλλά συνήθως δεν αναλαμβάνουν εκτεταμένες ή κοινωνικές ευθύνες για τους καταναλωτές.

Συνοψίζοντας, το μοντέλο αποτελείται από πέντε μέρη: το στοιχείο του παροχέα, το στοιχείο του καταναλωτή, τον τύπο της συνάντησης, τη βάση της γνώσης και το περιβάλλον συνέχειας, όλα τροποποιημένα σύμφωνα με το επίπεδο της έντασης της συνέχειας.

ΣΤΟΙΧΕΙΟ ΠΑΡΟΧΕΑ	ΓΙΑΤΡΟΣ, ΝΟΣΟΚΟΜΑ, ΚΟΙΝΩΝΙΚΟΣ ΛΕΙΤΟΥΡΓΟΣ, ΤΕΧΝΙΚΟΣ, ΔΙΟΙΚΗΤΙΚΟ ΠΡΟΣΩΠΙΚΟ
------------------	--

ΣΤΟΙΧΕΙΟ ΚΑΤΑΝΑΛΩΤΗ	ΜΕΜΟΝΩΜΕΝΟΣ ΑΣΘΕΝΗΣ, ΟΙΚΟΓΕΝΕΙΑ, ΕΠΑΓΓΕΛΜΑΤΙΚΕΣ ΟΜΑΔΕΣ, ΚΟΙΝΟΤΙΚΕΣ ΟΜΑΔΕΣ
---------------------	--

ΤΥΠΟΣ ΣΥΝΑΝΤΗΣΗΣ	ΑΜΕΣΗ:	ΠΡΟΣΩΠΟ ΜΕ ΠΡΟΣΩΠΟ (ΟΠΤΙΚΗ ΚΑΙ ΛΕΚΤΙΚΗ) ΤΗΛΕΦΩΝΗΜΑ (ΛΕΚΤΙΚΗ) ΓΡΑΠΤΟ ΜΗΝΥΜΑ
	ΕΜΜΕΣΗ:	ΠΡΟΣΩΠΟ ΜΕ ΠΡΟΣΩΠΟ (ΟΠΤΙΚΗ ΚΑΙ ΛΕΚΤΙΚΗ) ΤΗΛΕΦΩΝΗΜΑ (ΛΕΚΤΙΚΗ) ΓΡΑΠΤΟ ΜΗΝΥΜΑ

ΒΑΣΗ ΓΝΩΣΗΣ	ΚΑΤΑΓΕΓΡΑΜΜΕΝΗ:	ΙΑΤΡΙΚΑ ΑΡΧΕΙΑ, ΜΗΝΥΜΑΤΑ, ΔΙΣΚΕΤΕΣ Η/Υ
	ΜΗ ΚΑΤΑΓΕΓΡΑΜΜΕΝΗ:	ΠΛΗΡΟΦΟΡΙΕΣ ΑΠΟ ΠΡΩΤΟ ΧΕΡΙ (ΜΝΗΜΗ) ΠΛΗΡΟΦΟΡΙΕΣ ΑΠΟ ΔΕΥΤΕΡΟ ΧΕΡΙ

ΠΕΡΙΒΑΛΛΟΝ ΣΥΝΕΧΕΙΑΣ	ΧΡΟΝΙΚΟ – ΠΕΡΙΘΑΛΨΗ ΣΤΟ ΧΡΟΝΟ ΓΕΩΓΡΑΦΙΚΟ – ΜΕΡΟΣ ΠΑΡΟΧΗΣ ΠΕΡΙΘΑΛΨΗΣ ΕΙΔΙΚΟΤΗΤΑ – ΠΕΡΙΟΡΙΣΜΕΝΗ Ή ΟΛΙΣΤΙΚΗ ΣΧΕΣΗ – ΜΕΤΑΞΥ ΑΣΘΕΝΩΝ ΚΑΙ ΠΑΡΟΧΕΩΝ ΣΥΣΤΗΜΑ ΠΛΗΡΟΦΟΡΗΣΗΣ – ΙΑΤΡΙΚΑ ΑΡΧΕΙΑ ΔΙΑΘΕΣΙΜΟΤΗΤΑ – ΤΟΠΟΘΕΣΙΑ ΠΡΟΣΒΑΣΗΣ, ΣΥΣΤΗΜΑ ΡΑΝΤΕΒΟΥ ΚΛΠ ΚΙΝΗΤΙΚΟΤΗΤΑ/ ΣΤΑΘΕΡΟΤΗΤΑ – ΑΣΘΕΝΗΣ Ή ΠΑΡΟΧΕΑΣ
----------------------	---

ΣΧΗΜΑ 1 – ΟΙ ΣΥΝΙΣΤΩΣΕΣ ΤΟΥ ΜΟΝΤΕΛΟΥ ΣΥΝΕΧΕΙΑΣ

Όπως είδαμε προηγουμένως, η συνέχεια της περίθαλψης είναι σημαντικός παράγοντας όσον αφορά την ποιότητα της παρεχόμενης περίθαλψης. Επομένως είναι σημαντικό να εισάγουμε ορισμένους δείκτες για τη μέτρηση της συνέχειας οι οποίοι να μας επιτρέπουν να βρούμε κατά πόσο μια περίθαλψη είναι συνεχής ή όχι και να κάνουμε συγκρίσεις.

Κατά τη μελέτη των μοντέλων συνέχειας μπορούν να χρησιμοποιηθούν διαφορετικά επίπεδα μέτρησης. Μπορούμε να ξεκινήσουμε από το επίπεδο επίσκεψης αποδίδοντας μια τιμή συνέχειας σε κάθε επίσκεψη, αναλόγως με το ποιος παροχέας χρησιμοποιήθηκε σε σχέση με το ιστορικό του ασθενούς, κατά τη διάρκεια κάποιας προκαθορισμένης περιόδου προσδιορισμού της συνέχειας (*continuity determining period, cdp*). Για παράδειγμα, αν χρησιμοποιείται ο οικογενειακός γιατρός δίνουμε την τιμή 1, ενώ αν χρησιμοποιείται οποιοσδήποτε άλλος γιατρός την τιμή 0. Οι περισσότεροι συγγραφείς ξεκινούν με το επίπεδο του ατόμου αποδίδοντας μια τιμή συνέχειας στον κάθε ασθενή για την φροντίδα που εισέπραξε καθ' όλη τη διάρκεια μιας χρονικής περιόδου, της περιόδου μέτρησης (*measurement period, mp*), π.χ. πόσους παροχείς χρησιμοποίησε ο κάθε ασθενής σε μια συγκεκριμένη χρονική περίοδο. Τόσο τα μέτρα που βασίζονται στις επισκέψεις όσο και τα μέτρα που βασίζονται στα άτομα μπορούν να αναχθούν σε επίπεδο πληθυσμού.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κ Ε Φ Α Λ Α Ι Ο 2

Τα κυριότερα μέτρα της συνέχειας της περίθαλψης

2.1 Εισαγωγή

Τις τελευταίες δύο δεκαετίες έχουν προταθεί πολλές στατιστικές συναρτήσεις για τη μέτρηση της συνέχειας της περίθαλψης. Οι ορισμοί των περισσότερων μέτρων της συνέχειας έχουν την εξής κοινή βάση: έχουμε m παροχές υγειονομικής περίθαλψης και μελετάμε ποιον παροχέα επισκέφθηκε ο ασθενής στην i επίσκεψη, μελετάμε δηλαδή μια ακολουθία επισκέψεων ασθενών σε m παροχές (π.χ. γιατρούς). Χρησιμοποιώντας τις έννοιες των μοντέλων και των ροών σε δοκιμές πολλαπλών καταστάσεων, ορίζονται οι επόμενες πέντε στατιστικές συναρτήσεις: *UPC* (*usual provider continuity*, συνέχεια του συνήθους παροχέα), *SECON* (*sequential continuity*, ακολουθιακή συνέχεια), *NOP* (*number of providers*, πλήθος παροχέων), *K index*, *COC* (*continuity of care*, συνέχεια της περίθαλψης).

Σε ένα μοντέλο συνέχειας με m διαθέσιμους παροχείς υγειονομικής περίθαλψης ορίζουμε τη θετική ακέραια τυχαία μεταβλητή $v(T)$ με συνάρτηση πιθανότητας $\mu(\cdot)$, η οποία παριστάνει το πλήθος των επισκέψεων που πραγματοποιήθηκαν κατά τη χρονική περίοδο $[0, T]$. Στη συνέχεια θεωρούμε τις τυχαίες μεταβλητές $\{X_1, \dots, X_{v(T)}\}$ που ορίζονται από μια ακολουθία δοκιμών πολλαπλών καταστάσεων με έναν πεπερασμένο χώρο καταστάσεων $S = \{1, 2, \dots, m\}$. Η κάθε τυχαία μεταβλητή X_i , $i = 1, \dots, v(T)$ αντιπροσωπεύει το αποτέλεσμα της i δοκιμής και καθορίζει ποιος παροχέας χρησιμοποιήθηκε στην i επίσκεψη, δηλαδή

$$X_i = u$$

αν και μόνο αν κατά την επίσκεψη i ο ασθενής χρησιμοποίησε τον παροχέα u ($u \in S$).

2.2 Τα κυριότερα μέτρα της συνέχειας της περίθαλψης

Με βάση τα παραπάνω μπορούμε τώρα να ορίσουμε μαθηματικά τις πέντε στατιστικές συναρτήσεις συνέχειας της περίθαλψης ως εξής:

1. UPC (usual provider continuity, συνέχεια του συνήθους παροχέα)

Έστω u ο συνήθης παροχέας και $P(X_i = u) = p_u$, $i = 1, \dots, v(T)$, $u \in S$. Ορίζουμε

$$I_u(X_i) = \begin{cases} 1 & \text{αν } X_i = u \\ 0 & \text{διαφορετικά} \end{cases}$$

Τότε

$$UPC = \frac{\sum_{i=1}^{v(T)} I_u(X_i)}{v(T)} = \frac{n_u}{v(T)}$$

όπου $n_u = \sum_{i=1}^{v(T)} I_u(X_i)$ είναι το πλήθος των φορών που ο ασθενής χρησιμοποίησε τον συνήθη παροχέα του.

Επομένως ο UPC είναι το ποσοστό των επισκέψεων στον συνήθη παροχέα επί του συνόλου των επισκέψεων.

Δοθέντος ότι $v(T) = n$ και κάτω από την υπόθεση ότι οι επισκέψεις στον παροχέα είναι ανεξάρτητες, το n_u ακολουθεί διωνυμική κατανομή με παράμετρο $p_u = P(I_u(X_i) = 1)$.

Επομένως, η συνάρτηση πυκνότητας πιθανότητας του n_u θα δίνεται από τον τύπο:

$$P(n_u = x) = \binom{n}{x} p_u^x (1 - p_u)^{n-x}, \quad x = 0, 1, \dots, n$$

και ο UPC είναι αμερόληπτος εκτιμητής του p_u , δηλαδή $E(UPC) = p_u$.

Ο δείκτης UPC παίρνει τιμές στο διάστημα $[0, 1]$. Μικρές τιμές του δείκτη φανερώνουν ότι η συνέχεια του συνήθους παροχέα είναι μικρή ενώ μεγάλες τιμές του ότι η συνέχεια του συνήθους παροχέα είναι μεγάλη. Αν ο UPC ισούται με το 0 σημαίνει ότι καμία επίσκεψη δεν έγινε στον συνήθη παροχέα και αν ισούται με 1 σημαίνει ότι όλες οι επισκέψεις έγιναν στον εν λόγω παροχέα.

2. SECON (sequential continuity, ακολουθιακή συνέχεια)

Ορίζουμε αρχικά την παράμετρο

$$\theta = \sum_{i=1}^m \pi_i p_{ii},$$

όπου οι αρχικές πιθανότητες $\pi_i = \lim_{n \rightarrow \infty} P(X_n = i)$, $i = 1, \dots, m$, είναι οι εργοδικές στάσιμες πιθανότητες της μαρκοβιανής αλυσίδας και p_{ii} είναι η πιθανότητα μεταπηδήσεως από την κατάσταση i στον εαυτό της. Η αρχική πιθανότητα $\pi_i = \lim_{n \rightarrow \infty} P(X_n = i)$, $i = 1, \dots, m$ παριστάνει την πιθανότητα, για n μεγάλο, να χρησιμοποιηθεί ο i παροχέας στην n -οστή επίσκεψη και η πιθανότητα μεταπηδήσεως $p_{ii} = P(X_i = i | X_{i-1} = i)$, $i = 1, \dots, m$ είναι η πιθανότητα να χρησιμοποιηθεί ο παροχέας i σε δύο διαδοχικές επισκέψεις. Επομένως η παράμετρος θ μας δίνει την πιθανότητα για μεγάλο n να χρησιμοποιηθεί κάποιος από τους m διαθέσιμους παροχείς σε δύο διαδοχικές επισκέψεις. Για $i = 2, 3, \dots$, ορίζουμε τις τυχαίες μεταβλητές

$$Y_i = \begin{cases} 1 & \text{αν } X_i = X_{i-1} \\ 0 & \text{διαφορετικά} \end{cases}$$

δηλαδή η Y_i παίρνει την τιμή 1 αν ο παροχέας που χρησιμοποιήθηκε στην επίσκεψη i είναι ο ίδιος με εκείνον της επίσκεψης $i-1$.

Τότε

$$SECON = \frac{\sum_{i=2}^{v(T)} Y_i}{v(T)} = \frac{S_{v(T)}}{v(T)},$$

όπου $S_{v(T)} = \sum_{i=2}^{v(T)} Y_i$ είναι το πλήθος των φορών που για δυο διαδοχικές επισκέψεις χρησιμοποιήθηκε ο ίδιος παροχέας.

Αν ονομάσουμε την τυχαία μεταβλητή Y_i ακολουθιακή συνέχεια της i επίσκεψης, μπορούμε να πούμε ότι ο δείκτης $SECON$ είναι η μέση ακολουθιακή συνέχεια των επισκέψεων, δηλαδή μας δείχνει πόσες φορές χρησιμοποιήθηκε ο ίδιος παροχέας σε δύο διαδοχικές επισκέψεις επί του συνόλου των επισκέψεων.

Δοθέντος ότι $v(T) = n$, η αναμενόμενη τιμή του δείκτη $SECON$ δίνεται από τον τύπο

$$E(SECON | v(T) = n) = E\left(\frac{S_n}{n}\right) = \theta,$$

δηλαδή ο *SECON* είναι αμερόληπτος εκτιμητής για την παράμετρο ακολουθιακής συνέχειας θ για κάθε δεδομένο $v(T) = n$. Ως εκ τούτου, ο *SECON* είναι επίσης αμερόληπτος αν η κατανομή του $v(T)$ είναι ανεξάρτητη της ακολουθίας $\{X_i\}$.

Το διάστημα στο οποίο παίρνει τιμές ο δείκτης *SECON* είναι το $[0,1)$. Αν πάρει την τιμή 0 σημαίνει ότι δε χρησιμοποιήθηκε ο ίδιος παροχέας σε 2 συνεχόμενες επισκέψεις, ενώ τιμές του δείκτη κοντά στη μονάδα φανερώνουν μια αρκετά καλή συνέχεια ως προς κάποιον από τους διαθέσιμους παροχείς, χωρίς να μας ενδιαφέρει ποιος είναι αυτός όπως στην περίπτωση του *UPC*.

3. *NOP* (number of providers, πλήθος παροχέων)

Θέτουμε

$$I_N(X_1) \equiv 1, \quad I_N(X_2) = \begin{cases} 1 & \text{αν } X_2 \neq X_1 \\ 0 & \text{διαφορετικά,} \end{cases}$$

και για $i = 3, \dots, n$

$$I_N(X_i) = \begin{cases} 1 & \text{αν } X_i \neq X_j, \forall j = 1, \mathbf{K}, i-1 \\ 0 & \text{διαφορετικά.} \end{cases}$$

Αν ορίσουμε $W_{v(T)} = \sum_{i=1}^{v(T)} I_N(X_i)$ ως το πλήθος των διαφορετικών παροχέων που χρησιμοποιήθηκαν από τον ασθενή κατά την περίοδο μέτρησης, ο δείκτης *NOP* ορίζεται ως εξής:

$$NOP = \sum_{i=1}^{v(T)} I_N(X_i) / m = W_{v(T)} / m,$$

όπου $1 \leq W_{v(T)} \leq \min\{v(T), m\}$.

Από το πεδίο τιμών του $W_{v(T)}$ εύκολα διαπιστώνουμε ότι ο δείκτης *NOP* παίρνει τιμές μεταξύ του $1/m$ και του 1. Αν η τιμή του είναι $1/m$ σημαίνει ότι μόνο ένας παροχέας χρησιμοποιήθηκε στη χρονική περίοδο $[0, T]$ ενώ αν ο *NOP* πάρει την τιμή 1 σημαίνει ότι χρησιμοποιήθηκαν όλοι οι διαθέσιμοι παροχείς. Επομένως, αν η τιμή του *NOP* είναι κοντά στο $1/m$, τότε έχουμε μια καλή συνέχεια στην υγειονομική περίθαλψη του ασθενούς.

4. K index (δείκτης K)

Ακολουθώντας την παραπάνω λογική, ο K ορίζεται ως

$$K = \frac{v(T) - W_{v(T)}}{v(T) - 1} = 1 - \frac{W_{v(T)} - 1}{v(T) - 1} = 1 - \frac{NOP - 1/m}{v(T)/m - 1/m}$$

και επομένως συνδέεται στενά με τον NOP .

Από τον ορισμό του δείκτη K , εύκολα διαπιστώνουμε ότι ο δείκτης αυτός παίρνει τιμές στο διάστημα $\left[\frac{v(T) - m}{v(T) - 1}, 1 \right]$. Η τιμή 1 του δείκτη K δηλώνει ότι έχει χρησιμοποιηθεί ένας

μόνο παροχέας ενώ η τιμή $\frac{v(T) - m}{v(T) - 1}$ δηλώνει ότι έχουν χρησιμοποιηθεί όλοι οι διαθέσιμοι παροχείς. Εάν μάλιστα σε κάθε επίσκεψη έχει χρησιμοποιηθεί διαφορετικός παροχέας, δηλαδή αν $v(T) = m$, τότε ο δείκτης K παίρνει την τιμή 0.

Από τον τύπο του δείκτη K μπορούμε να πούμε ότι η διαφορά που ορίζεται στον αριθμητή, δηλαδή το πλήθος των επισκέψεων μείον το πλήθος των παροχέων που δέχθηκαν επίσκεψη, είναι αυτή που καθορίζει το πεδίο τιμών του δείκτη. Αν η διαφορά αυτή είναι μικρή, τότε και η τιμή του δείκτη είναι μικρή, επομένως έχουν χρησιμοποιηθεί πολλοί παροχείς και η συνέχεια της περίθαλψης δεν είναι καλή. Αντιθέτως, αν έχουν χρησιμοποιηθεί λίγοι παροχείς, η διαφορά είναι μεγάλη και επομένως και η τιμή που παίρνει ο δείκτης είναι μεγάλη και η συνέχεια της περίθαλψης καλή. Επομένως τιμές του δείκτη κοντά στη μονάδα φανερώνουν μια καλή συνέχεια στην περίθαλψη του ασθενούς.

5. COC (continuity of care, συνέχεια της περίθαλψης)

Με την ίδια λογική του ορισμού του UPC , ορίζουμε $P(X_i = u) = p_u$, $u \in S = \{1, 2, \dots, m\}$ και

$$I_u(X_i) = \begin{cases} 1 & \text{αν } X_i = u \\ 0 & \text{διαφορετικά.} \end{cases}$$

Τότε ο δείκτης COC δίνεται από τον τύπο

$$COC = \frac{HH - 1/v(T)}{1 - 1/v(T)},$$

όπου

$$HH = \sum_{u=1}^m \left\{ \frac{\sum_{i=1}^{v(T)} I_u(X_i)}{v(T)} \right\}^2.$$

Ο δείκτης HH (Herfindahl-Hirschman index) ορίζεται ως το άθροισμα των τετραγώνων των αναλογιών των επισκέψεων για τον κάθε παροχέα και δείχνει κατά πόσο η περίθαλψη είναι συγκεντρωμένη σε έναν παροχέα ή είναι διεσπαρμένη σε πολλούς παροχείς. Ο δείκτης HH παίρνει τιμές στο διάστημα $\left[\frac{1}{m}, 1\right]$. Όταν παίρνει την τιμή $1/m$, σε κάθε επίσκεψη έχει χρησιμοποιηθεί διαφορετικός παροχέας, αφού $v(T) = m$ (δηλαδή το πλήθος των επισκέψεων είναι ίσο με το πλήθος των παροχέων που χρησιμοποιήθηκαν) και επομένως η περίθαλψη είναι διεσπαρμένη σε όλους τους παροχείς. Όταν παίρνει την τιμή 1, έχει χρησιμοποιηθεί ένας μόνο παροχέας και επομένως η περίθαλψη είναι συγκεντρωμένη μόνο σε έναν παροχέα.

Ο δείκτης HH χρησιμοποιήθηκε για πρώτη φορά από τον Hirschman σε μια έρευνα για το εμπόριο στη δεκαετία του '40. Το 1950, ο Herfindahl στη διδακτορική του διατριβή χρησιμοποίησε τον τύπο του δείκτη HH όπως χρησιμοποιείται σήμερα. Επίσης, το 1959 έγινε χρήση του ίδιου δείκτη και πάλι από τον Herfindahl σε μια έρευνα για τη διεθνή βιομηχανία χαλκού. Ο δείκτης έγινε γνωστός ως *Herfindahl index* ύστερα από μελέτες του Rosenbluth (1955 και 1957). Το 1964 ο Hirschman δημοσίευσε ένα μικρό άρθρο στο *American Economic Review* διεκδικώντας την «πατρότητα» του δείκτη. Από τότε, ο δείκτης αναφέρεται στη βιβλιογραφία ως δείκτης HH (*Herfindahl-Hirschman Index*).

Όπως και ο δείκτης HH , ο COC είναι ένας δείκτης μέτρησης του κατά πόσο η περίθαλψη είναι διεσπαρμένη σε πολλούς παροχείς ή όχι. Είναι μια μεταβλητή με πεδίο τιμών το $[0,1]$. Η τιμή 0 (η οποία προκύπτει αν $HH = 1/m = 1/v(T)$) υποδηλώνει μέγιστη διασπορά και εμφανίζεται όταν σε κάθε επίσκεψη χρησιμοποιείται διαφορετικός παροχέας. Ο COC παίρνει την τιμή 1 όταν $HH = 1$, που σημαίνει ότι σε όλες τις επισκέψεις χρησιμοποιείται ο ίδιος παροχέας.

$$\text{Εύκολα μπορεί ναδειχθεί ότι } E(COC) = \sum_{j=1}^m p_j^2.$$

Λαμβάνοντας υπόψη τη σχέση μεταξύ COC και HH , αν η $v(T)$ είναι ανεξάρτητη της $\{X_i\}$, τότε η ακριβής κατανομή του COC μπορεί να καθοριστεί από την κατανομή του HH .

2.3 Παράδειγμα

Για να γίνουν απολύτως κατανοητοί οι δείκτες που παρατέθηκαν στην προηγούμενη ενότητα, παραθέτουμε στη συνέχεια ένα απλό παράδειγμα.

Έστω $m = 3$ (δηλ $S = \{A, B, C\}$) και ότι η ακολουθία των επισκέψεων ενός ασθενούς ήταν $\{ABAACCCABB\}$. Σύμφωνα με τα προηγούμενα, $v(T) = n = 10$ και για τους πέντε δείκτες συνέχειας θα έχουμε τα επόμενα:

UPC: Βλέπουμε από την ακολουθία των 10 επισκέψεων ότι ο συνήθης παροχέας είναι ο A . Επομένως η τυχαία μεταβλητή $I_u(X_i)$ θα παίρνει την τιμή 1 αν στην επίσκεψη i ο παροχέας είναι ο A και 0 σε οποιαδήποτε άλλη περίπτωση. Επομένως με βάση την ακολουθία των επισκέψεων του ασθενούς έχουμε:

$$\sum_{i=1}^{v(T)} I_u(X_i) = n_u = 4 \quad \text{και} \quad UPC = \frac{\sum_{i=1}^{v(T)} I_u(X_i)}{v(T)} = \frac{4}{10}.$$

SECON: Εδώ ορίζουμε την μεταβλητή Y_i η οποία παίρνει την τιμή 1 αν ο παροχέας που χρησιμοποιείται στην επίσκεψη i είναι ο ίδιος που είχε χρησιμοποιηθεί στην επίσκεψη $i-1$ και 0 διαφορετικά. Συνεπώς θα έχουμε τον πίνακα:

Επίσκεψη i	Παροχέας στον οποίο έγινε επίσκεψη	Y_i
1	A	0
2	B	0
3	A	0
4	A	1
5	C	0
6	C	1
7	C	1
8	A	0
9	B	0
10	B	1

απ' όπου προκύπτει

$$\sum_{i=1}^{v(T)} Y_i = S_{v(T)} = 4$$

και ο δείκτης ακολουθιακής συνέχειας θα ισούται με:

$$SECON = \frac{\sum_{i=1}^{v(T)} Y_i}{v(T)} = \frac{4}{10}.$$

NOP: Για τον υπολογισμό του δείκτη αυτού χρησιμοποιούμε τις μεταβλητές $I_N(X_i), i=1,2,\dots,n$ για τις οποίες ισχύει: η $I_N(X_1)$ ισούται με 1 εξ' ορισμού, η $I_N(X_2)$ ισούται με 1 αν ο παροχέας που χρησιμοποιήθηκε στην 1^η επίσκεψη είναι διαφορετικός από αυτόν της 2^{ης} και 0 διαφορετικά και τέλος η $I_N(X_i), i=3,\dots,n$ ισούται με 1 αν ο παροχέας της επίσκεψης i είναι διαφορετικός από εκείνον της επίσκεψης j , για όλα τα $j=1,\dots,i-1$ και 0 διαφορετικά (π.χ. η $I_N(X_3)$ ισούται με 1 αν ο παροχέας της 3^{ης} επίσκεψης είναι διαφορετικός από εκείνους των επισκέψεων 1 και 2 και 0 διαφορετικά κοκ). Επίσης το m που είναι ο αριθμός των διαθέσιμων παροχέων ισούται με 3. Επομένως:

$$NOP = \frac{\sum_{i=1}^{v(T)} I_N(X_i)}{m} = \frac{W_{v(T)}}{m} = \frac{3}{3}$$

Αναλυτικά οι τιμές των τυχαίων μεταβλητών $I_N(X_i)$ φαίνονται στον παρακάτω πίνακα:

Επίσκεψη i	Παροχέας στον οποίο έγινε επίσκεψη	$I_N(X_i)$
1	A	1
2	B	1
3	A	0
4	A	0
5	C	1
6	C	0
7	C	0
8	A	0
9	B	0
10	B	0

K index: Από τον ορισμό του δείκτη K έχουμε:

$$K = \frac{v(T) - W_{v(T)}}{v(T) - 1} = 1 - \frac{NOP - 1/m}{v(T)/m - 1/m} = 1 - \frac{3/3 - 1/3}{10/3 - 1/3} = \frac{7}{9}.$$

COC: Για τον υπολογισμό του COC ορίζουμε τη μεταβλητή $I_j(X_i)$, $j \in S = \{A, B, C\}$ η οποία ισούται με 1 αν στην επίσκεψη i ο παροχέας που χρησιμοποιείται είναι ο j και 0 διαφορετικά. Επίσης πρέπει να υπολογιστεί και ο δείκτης HH :

$$\begin{aligned} HH &= \sum_{j=1}^m \left\{ \sum_{i=1}^{v(T)} I_j(X_i) / v(T) \right\}^2 = \sum_{j=1}^3 \left\{ \sum_{i=1}^{10} I_j(X_i) / 10 \right\}^2 = \sum_{j=1}^3 \left\{ \frac{I_j(X_1) + \dots + I_j(X_{10})}{10} \right\}^2 = \\ &= \left\{ \frac{I_1(X_1) + \dots + I_1(X_{10})}{10} \right\}^2 + \left\{ \frac{I_2(X_1) + \dots + I_2(X_{10})}{10} \right\}^2 + \left\{ \frac{I_3(X_1) + \dots + I_3(X_{10})}{10} \right\}^2 = \\ &= \left[\frac{4}{10} \right]^2 + \left[\frac{3}{10} \right]^2 + \left[\frac{3}{10} \right]^2 = \frac{34}{100}. \end{aligned}$$

Με τη βοήθεια της ποσότητας HH υπολογίζουμε και τον COC :

$$COC = \frac{HH - 1/v(T)}{1 - 1/v(T)} = \frac{34/100 - 1/10}{1 - 1/10} = \frac{8}{30}.$$

Αξίζει εδώ να σημειωθεί ότι αναφορικά προς τη διάκριση που είχαμε κάνει παραπάνω μεταξύ των μέτρων της συνέχειας της περιθάλψης, δηλαδή σε μέτρα που βασίζονται στα άτομα και σε μέτρα που βασίζονται στις επισκέψεις, οι πέντε παραπάνω δείκτες είναι μέτρα που βασίζονται στις επισκέψεις, παράγονται όμως από αντίστοιχα μέτρα που βασίζονται στα άτομα.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΚΕΦΑΛΑΙΟ 3

Η ακριβής κατανομή του *SECON*

3.1 Εισαγωγή

Υπενθυμίζουμε αρχικά τον ορισμό του δείκτη ακολουθιακής συνέχειας *SECON*. Έστω $\{X_k\}$ μια στάσιμη ακολουθία τυχαίων μεταβλητών που σχηματίζουν μια ομογενή μαρκοβιανή αλυσίδα, η οποία αντιπροσωπεύει το αποτελέσματα των k επισκέψεων και ορίζεται σε έναν πεπερασμένο χώρο καταστάσεων $S = \{1, 2, \dots, m\}$. Υποθέτουμε ότι η αλυσίδα είναι μη υποβιβάσιμη και μη περιοδική (*irreducible and aperiodic*) με πίνακα πιθανοτήτων μεταπηδήσεως

$$A = (p_{ij})_{m \times m} \quad (3.1)$$

όπου $p_{ij} = P(X_t = j | X_{t-1} = i)$.

Έστω $\boldsymbol{\pi} = (\pi_1, \dots, \pi_m)$ η εργοδική (στάσιμη) κατανομή που σχετίζεται με τη μαρκοβιανή αλυσίδα, δηλαδή $\pi_i = \lim_{n \rightarrow \infty} P(X_n = i)$, για $i = 1, \dots, m$. Ορίζουμε μια παράμετρο

$$\theta = \sum_{i=1}^m \pi_i p_{ii}, \quad (3.2)$$

που είναι το μέτρο της ακολουθιακής συνέχειας που σχετίζεται με τη μαρκοβιανή αλυσίδα. Η παράμετρος θ εκφράζει, όπως αναφέρθηκε και στο Κεφάλαιο 2, την πιθανότητα να χρησιμοποιηθεί κάποιος από τους m διαθέσιμους παροχείς σε δύο διαδοχικές επισκέψεις.

Έστω $\nu(T)$ μια θετική ακέραια τυχαία μεταβλητή που ορίζεται στο $J^+ = \{1, 2, \dots\}$ και έχει συνάρτηση πιθανότητας $\mu(\cdot)$. Η $\nu(T)$ είναι ο αριθμός των δοκιμών, δηλαδή των επισκέψεων που έχουν συμβεί μέχρι το χρόνο T . Ορίζουμε μια ακολουθία τυχαίων μεταβλητών Y_i , $i = 1, 2, \dots$, ως εξής:

$$Y_i = \begin{cases} 1 & \text{αν } X_i = X_{i-1} \\ 0, & \text{διαφορετικά.} \end{cases} \quad (3.3)$$

Ο Steinwachs (1979) πρότεινε τη χρήση του μέσου αποτελέσματος (outcome average) της ακολουθίας $\{Y_i\}, i = 1, \dots, v(T)$ και το ονόμασε *SECON*, ως έναν εκτιμητή της παραμέτρου της ακολουθιακής συνέχειας θ , όπου

$$SECON = \frac{S_{v(T)}}{v(T)} \quad (3.4)$$

και

$$S_{v(T)} = Y_1 + \dots + Y_{v(T)} \quad (3.5)$$

Αν η τυχαία μεταβλητή $v(T)$ είναι ανεξάρτητη της μαρκοβιανής αλυσίδας $\{X_k\}$, ο *SECON* είναι αμερόληπτος εκτιμητής για την παράμετρο ακολουθιακής συνέχειας θ . Πράγματι, για κάθε δεδομένο $v(T) = n, n \in J^+$,

$$E(SECON | v(T) = n) = E\left(\frac{S_n}{n}\right) = \theta. \quad (3.6)$$

Εύκολα βλέπουμε ότι η ακολουθία των τυχαίων μεταβλητών $\{Y_i\}$ που προκύπτει από τη μαρκοβιανή αλυσίδα $\{X_k\}$ μέσω της σχέσης (3.3) είναι στάσιμη αλλά όχι μαρκοβιανή η ίδια. Επομένως οι συνήθεις τεχνικές μαρκοβιανών αλυσίδων δεν μπορούν να εφαρμοστούν απευθείας για τη μελέτη των ακριβών κατανομών των $S_{v(T)}$ και *SECON*.

3.2 Η ακριβής κατανομή

Έστω $\Lambda_1 = 11, \Lambda_2 = 22, \dots, \Lambda_m = mm$, m απλά σχέδια και $\Lambda = \mathbf{U}_{i=1}^m \Lambda_i$ ένα σύνθετο σχέδιο που προέρχεται από τα Λ_i . Η ερμηνεία του σχεδίου $\Lambda_i = ii$ είναι ότι ο παροχέας i έχει χρησιμοποιηθεί σε δύο συνεχόμενες επισκέψεις. Ορίζουμε $X_n(\Lambda)$ ως το συνολικό αριθμό σχεδίων του Λ (με επικαλυπτόμενη μέτρηση, *overlap counting*) που συνέβησαν σε μια ακολουθία δοκιμών m καταστάσεων $\{X_0, X_1, \dots, X_n\}$, στις οποίες υπάρχει μαρκοβιανή εξάρτηση (Markov dependent m -state trials). Δηλαδή $X_n(\Lambda)$ είναι ο συνολικός αριθμός σχεδίων είτε του Λ_1, \dots , είτε του Λ_m στην ακολουθία. Για παράδειγμα, αν $m = 3$ και η ακολουθία είναι $\{X_0 = 2, X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 2, X_5 = 3, X_6 = 3\}$, τότε ο συνολικός αριθμός σχεδίων $\Lambda = \Lambda_1 \cup \Lambda_2 \cup \Lambda_3$ είναι 3, δηλαδή $X_6(\Lambda) = 3$. Σύμφωνα με τις έννοιες που

ορίστηκαν στην ενότητα 3.1, για δοσμένο $v(T) = n$, $S_6 = \sum_{i=1}^6 Y_i = 3$. Στην πραγματικότητα, απορρέει από τον ορισμό του $X_n(\Lambda)$, με επικαλυπτόμενη μέτρηση, ότι $S_n = X_n(\Lambda)$ για κάθε ακέραιο $n \in J^+$.

Για δοσμένο n θετικό και ακέραιο, έστω $\Gamma_n = \{0, 1, \dots, n\}$ ένα πεπερασμένο σύνολο δεικτών και $\Omega_n = \{a_1, \dots, a_h\}$ ένας πεπερασμένος χώρος καταστάσεων με h καταστάσεις.

Ορισμός 3.1 Μια ακέραια τυχαία μεταβλητή $X_n(\Lambda) = \varphi(X_0, X_1, \dots, X_n)$ μπορεί να εμφυτευτεί (*imbeddable*) σε μια πεπερασμένη μαρκοβιανή αλυσίδα αν

- (i) υπάρχει μια μαρκοβιανή αλυσίδα $\{Z_t : t \in \Gamma_n\}$ που ορίζεται σε ένα πεπερασμένο χώρο καταστάσεων Ω_n με πίνακες πιθανοτήτων μεταπηδήσεως M_t και διάνυσμα αρχικών πιθανοτήτων ξ , και
- (ii) υπάρχει μια διαμέριση $\{C_s : s = 0, 1, \dots, l\}$ στον χώρο καταστάσεων Ω_n (όπου τα C_s και l μπορεί να εξαρτώνται από το n) έτσι ώστε για κάθε $s = 0, 1, \dots, l$, να ισχύει

$$P(X_n(\Lambda) = s) = P(Z_n \in C_s | \xi).$$

Για δεδομένο n , ας θεωρήσουμε μια ομογενή μαρκοβιανή αλυσίδα $\{X_t : t \in \Gamma_n\}$ όπως ορίστηκε στην Ενότητα 3.1. με πίνακα πιθανοτήτων μεταπηδήσεως A . Ο A μπορεί να χωριστεί σε 2 πίνακες:

$$A_{m \times m} = (p_{ij})_{m \times m} = B_{m \times m} + D_{m \times m} \quad (3.7)$$

όπου ο $D_{m \times m}$ είναι ένας διαγώνιος πίνακας, δηλαδή

$$D_{m \times m} = \begin{pmatrix} p_{11} & & 0 \\ & \mathbf{O} & \\ 0 & & p_{mm} \end{pmatrix}.$$

Στην περίπτωση αυτή μπορούμε να γράψουμε $B_{m \times m} = A_{m \times m} - D_{m \times m}$.

Έστω $\Omega_n = \{(u, v) : u = 0, 1, \dots, n, \text{ και } v = 1, 2, \dots, m\}$ ο χώρος καταστάσεων που περιέχει συνολικά $(n+1)m$ καταστάσεις και $\{Z_t : t \in \Gamma_n\}$ μια μαρκοβιανή αλυσίδα πάνω στο χώρο καταστάσεων Ω_n που ορίζεται ως εξής:

$$Z_t(X) = \begin{cases} \left(\sum_{i=1}^t Y_i, X_t \right) & 1 \leq t \leq n \\ (0, X_0) & t = 0. \end{cases} \quad (3.8)$$

Τότε ισχύει το ακόλουθο αποτέλεσμα, το οποίο δόθηκε από τους Fu and Lou (1996):

Θεώρημα 3.1 Για δεδομένο θετικό ακέραιο n , η τυχαία μεταβλητή $S_n = \sum_{i=1}^n Y_i$ είναι εμφυτεύσιμη σε πεπερασμένη μαρκοβιανή αλυσίδα (FMCI, Finite Markov Chain Imbeddable) και επομένως

- Υπάρχει μια ομογενής πεπερασμένη μαρκοβιανή αλυσίδα $\{Z_t : t=0,1,\dots,n\}$ που ορίζεται στο χώρο καταστάσεων Ω_n με διάνυσμα αρχικών πιθανοτήτων $\xi = (\boldsymbol{\pi}, \mathbf{0}, \mathbf{K}, \mathbf{0}, \mathbf{K}, \mathbf{0})_{1 \times (n+1)m}$ και πίνακα πιθανοτήτων μεταπηδήσεως, $t=1,2,\mathbf{K},n$

$$\mathbf{M}_t = \mathbf{M}_{(n+1)m \times (n+1)m} = \begin{pmatrix} \mathbf{B} & \mathbf{D} & \mathbf{0} & \mathbf{0} & \mathbf{L} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \mathbf{D} & \mathbf{0} & \mathbf{L} & \mathbf{0} & \mathbf{0} \\ & \mathbf{L} & \mathbf{O} & \mathbf{K} & & & \\ \mathbf{0} & & \mathbf{L} & \mathbf{B} & \mathbf{D} & & \\ \mathbf{0} & & \mathbf{L} & \mathbf{0} & \mathbf{B}^* & & \end{pmatrix}, \quad (3.9)$$

όπου $\boldsymbol{\pi} = (\pi_1, \dots, \pi_m)_{1 \times m}$, $\mathbf{0} = (\mathbf{0})_{m \times m}$ και $\mathbf{B}^* = \mathbf{I}_{m \times m}$,

- Υπάρχει μια διαμέριση

$$C = \{C_s = [(s,1), \dots, (s,m)] : s = 0,1,\dots,n\} \quad (3.10)$$

στο χώρο καταστάσεων Ω_n έτσι ώστε

$$P(S_n = s) = P(Z_n \in C_s | Z_0) = \xi \left(\prod_{t=1}^n \mathbf{M}_t \right) \mathbf{U}'(C_s), \quad 0 \leq s \leq n, \quad (3.11)$$

Όπου $\mathbf{U}(C_s) = (\mathbf{0}, \mathbf{K}, \mathbf{0}, \mathbf{1}, \mathbf{K}, \mathbf{1}, \mathbf{0}, \mathbf{K}, \mathbf{0})$ είναι ένα $1 \times (n+1)m$ διάνυσμα-γραμμή με μονάδες στις συντεταγμένες που σχετίζονται με το C_s και μηδενικά οπουδήποτε αλλού και

- Η συνάρτηση πιθανότητας του S_n ικανοποιεί την αναδρομική εξίσωση:

$$P(S_n = s) = \begin{cases} \sum_{i=1}^m (1 - p_{ii}) \xi \mathbf{M}^{n-1} \mathbf{U}'((0, i)) & s = 0 \\ \sum_{i=1}^m p_{ii} \xi \mathbf{M}^{n-1} \mathbf{U}'((s-1, i)) \\ + \sum_{i=1}^m (1 - p_{ii}) \xi \mathbf{M}^{n-1} \mathbf{U}'((s, i)) & 1 \leq s \leq n-1 \\ \sum_{i=1}^m p_{ii} \xi \mathbf{M}^{n-1} \mathbf{U}'((n-1, i)) & s = n \end{cases} \quad (3.12)$$

Ο χρόνος που απαιτείται για τον υπολογισμό της πιθανότητας $P(S_n = s)$ στις εξισώσεις (3.11) και (3.12) παραμένει ελάχιστος, ακόμα και όταν το n είναι μεγάλο. Επίσης, όταν $p_{11} = \mathbf{K} = p_{mm} = \theta$, οι υπολογισμοί μπορούν να απλοποιηθούν σημαντικά ως εξής:

Θεώρημα 3.2 Αν η μαρκοβιανή αλυσίδα $\{X_n\}$ είναι μη υποβιβάσιμη με $p_{11} = \mathbf{K} = p_{mm} = \theta$, τότε η συνάρτηση πιθανότητας της τ.μ. S_n δίνεται από τον τύπο

$$P(S_n = s) = \binom{n}{s} \boldsymbol{\pi} \mathbf{B}^{n-s} \mathbf{D}^s \mathbf{1}', \quad s = 0, 1, \mathbf{K}, n \quad (3.13)$$

όπου $\mathbf{1} = (1, 1, \mathbf{K}, 1)$ είναι ένα $1 \times m$ διάνυσμα-γραμμή με μονάδες σε όλες τις συντεταγμένες.

Για την επεξήγηση των παραπάνω θεωρημάτων, παρατίθεται το επόμενο παράδειγμα.

Παράδειγμα 3.1 Έστω $m = 2, n = 3$, και ότι ο πίνακας πιθανοτήτων μεταπηδήσεως της ομογενούς μαρκοβιανής αλυσίδας $\{X_0, X_1, X_2, X_3\}$ είναι ίσος με

$$\mathbf{A} = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}.$$

Βλέπουμε ότι $\mathbf{B} = \mathbf{B}'$ και $p_{11} = p_{22} = 1/3$. Η εργοδική κατανομή που σχετίζεται με την ομογενή μαρκοβιανή αλυσίδα είναι $\boldsymbol{\pi} = (1/2, 1/2)$. Αν $\boldsymbol{\pi}$ είναι η αρχική πιθανότητα του X_0 , τότε οι πιθανότητες $P(S_n = s)$ μπορούν να βρεθούν μέσω της εξίσωσης (3.13) αντί των εξισώσεων (3.11) ή (3.12), δηλαδή για $s = 0, 1, 2, 3$, έχουμε τις εκφράσεις

$$P(S_3 = s) = \binom{3}{s} (1/2, 1/2) \begin{pmatrix} 0 & 2/3 \\ 2/3 & 0 \end{pmatrix}^{3-s} \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix}^s \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Ο παραπάνω τύπος δίνει:

$$P(S_3 = 0) = 8/27,$$

$$P(S_3 = 1) = 12/27,$$

$$P(S_3 = 2) = 6/27 \text{ και}$$

$$P(S_3 = 3) = 1/27.$$

Τα αποτελέσματα αυτά μπορούν εύκολα να επιβεβαιωθούν γράφοντας και τις 16 πιθανές ακολουθίες από την μαρκοβιανή αλυσίδα $\{X_0, X_1, X_2, X_3\}$.

Προκειμένου να προχωρήσουμε στις αποδείξεις των Θεωρημάτων 3.1 και 3.2, δίνουμε αρχικά τα επόμενα δύο λήμματα.

Λήμμα 3.1 Έστω $e_j = (0, \mathbf{K}, 0, 1, 0, \mathbf{K}, 0)_{\times m}$, $j=1, \dots, m$ ένα μοναδιαίο διάνυσμα. Τότε για κάθε j

(i) $D e'_j = p_{jj}$, και

(ii) $B e'_j = \sum_{i=1}^m p_{ij} e'_j - p_{jj} e'_j$.

Απόδειξη. Τα αποτελέσματα (i) και (ii) απορρέουν απευθείας από τους ορισμούς των πινάκων D και B . □

Λήμμα 3.2 Αν $\{X_t\}$ είναι μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα με πιθανότητες μεταπηδήσεως $p_{ij} > 0$ για όλα τα $i, j=1, \dots, m$ τότε $BD=DB$ αν και μόνο αν $p_{11} = p_{22} = \mathbf{K} = p_{mm} = \theta$.

Απόδειξη. Αν $p_{11} = p_{22} = \mathbf{K} = p_{mm} = \theta > 0$, τότε $D = \theta I$ και είναι φανερό ότι $BD=DB$. Αφού ο D είναι διαγώνιος πίνακας και $p_{ii} > 0$, υπάρχει ο αντίστροφός του $D^{-1} = (d_{ij})$ και έχει στοιχεία $d_{ii} = p_{ii}^{-1}$ και $d_{ij} = 0$ αν $i \neq j$. Αρκεί να αποδείξουμε ότι αν $D^{-1}BD = B$, τότε $p_{11} = p_{22} = \mathbf{K} = p_{mm}$. Αφού

$$p_{ii}^{-1} p_{ij} p_{jj} = p_{ij}, i, j = 1, \mathbf{L}, m,$$

για όλα τα $p_{ij} > 0$, τότε $p_{ii} = p_{jj}$ για όλα τα i και j . Αυτό ολοκληρώνει την απόδειξη. □

Απόδειξη Θεωρήματος 3.1 Αφού $\{X_t : t \in \Gamma_n\}$ είναι μια μαρκοβιανή αλυσίδα με αρχικές πιθανότητες $\pi = (\pi_1, \mathbf{K}, \pi_m)$, προκύπτει από τον Ορισμό 3.1 ότι η αρχικές πιθανότητες που αντιστοιχούν στο Z_0 είναι $\xi = (\pi_1, \mathbf{K}, \pi_m, 0, \mathbf{K}, 0)_{\times (n+1)m}$. Επιπλέον, αφού $\{X_t\}$ είναι μια μαρκοβιανή αλυσίδα με πίνακα πιθανοτήτων μετάβασης A , προκύπτει από τον Ορισμό 3.1 και την εξίσωση (3.8) ότι $Z_t(X)$ είναι μια μαρκοβιανή αλυσίδα με πιθανότητες μετάβασης που ορίζονται από τις παρακάτω εξισώσεις : για $0 \leq s \leq n-1$, $i = 1, \mathbf{K}, m$,

$$P(Z_t(X) = (u, v) | Z_{t-1}(X) = (s, i)) = \begin{cases} p_{ii}, & \text{αν } u = s+1, v = i \\ p_{ij}, & \text{αν } u = s, v = j, \text{ με } i \neq j \\ 0, & \text{διαφορετικά,} \end{cases} \quad (3.14)$$

και για τις ειδικές περιπτώσεις όπου $s = n$ και $i = 1, 2, \mathbf{K}, m$,

$$P(Z_t(X) = (u, v) | Z_{t-1}(X) = (n, i)) = \begin{cases} 1, & \text{αν } u = n, v = i \\ 0, & \text{διαφορετικά} \end{cases} \quad (3.15)$$

Ο πίνακας πιθανοτήτων μεταπηδήσεως M_t που δίνεται στην εξίσωση (3.9) ορίζεται από τα παραπάνω. Προκύπτει από τον Ορισμό 3.1 ότι για κάθε $s = 0, 1, \mathbf{K}, n$, ισχύει $P(S_n = s) = P(Z_n(X) \in C_s)$ και η εξίσωση (3.11) είναι άμεση εφαρμογή των εξισώσεων των Charman-Kolmogorov. Η εξίσωση (3.12) προκύπτει άμεσα από την (3.11) και από το γεγονός ότι $M^n U'(C_s) = M^{n-1} [M U'(C_s)]$. Αυτό ολοκληρώνει την απόδειξη. \square

Απόδειξη Θεωρήματος 3.2 Από το Λήμμα 3.2 έχουμε ότι $BD = DB$ και από τη δομή του M διαπιστώνουμε (επαγωγικά) ότι οι πρώτες m γραμμές του πίνακα M^n έχουν την ακόλουθη μορφή:

$$\left[\begin{array}{c} \binom{n}{0} B^n, \binom{n}{1} B^{n-1} D, \mathbf{K}, \binom{n}{n} D^n \end{array} \right]_{m \times (n+1)m}. \quad (3.16)$$

Πολλαπλασιάζοντας με το διάνυσμα αρχικής πιθανότητας $\xi = (\boldsymbol{\pi}, \mathbf{0}, \mathbf{K}, \mathbf{0})$ της μαρκοβιανής αλυσίδας $\{Z_t\}$ παίρνουμε

$$\xi M^n = \left[\begin{array}{c} \binom{n}{0} \boldsymbol{\pi} B^n, \binom{n}{1} \boldsymbol{\pi} B^{n-1} D, \mathbf{K}, \binom{n}{n} \boldsymbol{\pi} D^n \end{array} \right]. \quad (3.17)$$

Σημειωτέον ότι το διάνυσμα $U(C_s) = (\mathbf{0}, \mathbf{K}, \mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{K}, \mathbf{0})$ έχει ένα $1 \times m$ διάνυσμα γραμμή $\mathbf{1} = (\mathbf{1}, \mathbf{K}, \mathbf{1})$ στις συντεταγμένες που σχετίζονται με το C_s . Το αποτέλεσμα της εξίσωσης (3.13) προκύπτει άμεσα από την εξίσωση (3.11). Αυτό ολοκληρώνει την απόδειξη. \square

Εστω ότι η τυχαία μεταβλητή $v(T)$ έχει ένα μέτρο πιθανότητας $\mu(\cdot)$ στο σύνολο $J^+ = \{1, 2, \mathbf{K}, \}$ και ότι είναι ανεξάρτητη της μαρκοβιανής αλυσίδας $\{X_t\}$. Τότε η κατανομή των $S_{v(T)}$ και $SECON$ μπορεί να προσδιοριστεί από το παρακάτω θεώρημα:

Θεώρημα 3.3 Έστω ότι η τ.μ. $v(T)$ και οι $\{X_t\}$ είναι ανεξάρτητες.

(i) Για τη συνάρτηση πιθανότητας της τ.μ. $S_{v(T)}$ ισχύει

$$P(S_{v(T)} = s) = \sum_{n=s}^{\infty} \xi M^n U'(C_s) \mu(n), \quad (3.18)$$

(ii) Αν $p_{11} = \mathbf{K} = p_{mm}$, ισχύει

$$P(S_{v(T)} = s) = \sum_{n=s}^{\infty} \binom{n}{s} \pi \mathbf{B}^{n-s} \mathbf{D}^s \mathbf{1}' \mu(n), \quad (3.19)$$

(iii) Για $\alpha \in [0,1]$, ισχύει

$$P\left(\frac{S_{v(T)}}{v(T)} \leq \alpha\right) = \sum_{n=1}^{\infty} \sum_{s=0}^{[an]} \binom{n}{s} \pi \mathbf{B}^{n-1} \mathbf{D}^s \mathbf{1}' \mu(n), \quad (3.20)$$

όπου $[an]$ είναι το ακέραιο μέρος του an .

Απόδειξη. Αφού η τυχαία μεταβλητή $v(T)$ και η μαρκοβιανή αλυσίδα $\{X_t\}$ είναι ανεξάρτητες, το αποτέλεσμα (i) απορρέει κατευθείαν από το Θεώρημα 3.1 και από την παρακάτω εξίσωση:

$$\begin{aligned} P(S_{v(T)} = s) &= \sum_{n=s}^{\infty} P(S_n = s \mid v(T) = n) P(v(T) = n) = \\ &= \sum_{n=s}^{\infty} P(S_n = s) \mu(n) \\ &= \sum_{n=s}^{\infty} \xi \mathbf{M}^n \mathbf{U}'(C_s) \mu(n). \end{aligned} \quad (3.21)$$

Επιπλέον, αν $p_{11} = \mathbf{K} = p_{mm}$, το αποτέλεσμα (ii) προκύπτει από το αποτέλεσμα (i) και το Θεώρημα 3.2. Το αποτέλεσμα (iii) προκύπτει απευθείας από το αποτέλεσμα (ii) και από την εξής διαπίστωση, η οποία ισχύει για κάθε $\alpha \in [0,1]$,

$$\begin{aligned} P\left(\frac{S_{v(T)}}{v(T)} \leq \alpha\right) &= \sum_{n=1}^{\infty} P(S_n / n \leq \alpha \mid v(T) = n) P(v(T) = n) \\ &= \sum_{n=1}^{\infty} P\left(\frac{S_n}{n} \leq \alpha\right) \mu(n) \\ &= \sum_{n=1}^{\infty} \sum_{s=0}^{[an]} \binom{n}{s} \pi \mathbf{B}^{n-1} \mathbf{D}^s \mathbf{1}' \mu(n). \end{aligned} \quad (3.22)$$

Έτσι ολοκληρώνεται την απόδειξη. □

3.3 Εκτίμηση του μέτρου ακολουθιακής συνέχειας

Η παράμετρος $\theta = \sum_{i=1}^m \pi_i p_{ii}$, που είναι το μέτρο ακολουθιακής συνέχειας και που σχετίζεται με τη μαρκοβιανή αλυσίδα $\{X_k\}$, η οποία είναι ανεξάρτητη της $v(T)$, έχει έναν αμερόληπτο εκτιμητή, τον *SECON*, όπως τον ορίσαμε στην ενότητα 3.1. Εύκολα διαπιστώνουμε ότι $\theta = 0$ αν και μόνο αν $p_{ii} = 0$ για όλα τα $i = 1, \mathbf{K}, m$. Αν $\{X_t : t \in \Gamma\}$ είναι, αντί για ακολουθία μαρκοβιανά εξαρτημένων δοκιμών, μια ακολουθία ανεξάρτητων και ισόνομων (*i.i.d.*) δοκιμών m καταστάσεων με πιθανότητες $p_{ij} = p_j \geq 0$, $i, j = 1, \mathbf{K}, m$, και $\sum_{j=1}^m p_j \equiv 1$, τότε ισχύει το ακόλουθο θεώρημα.

Θεώρημα 3.4 *Ας θεωρήσουμε μια ακολουθία από ανεξάρτητες και ισόνομες δοκιμές πολλαπλών καταστάσεων $\{X_t = i : i = 1, \mathbf{K}, m, t \in \mathbf{G}\}$. Η παράμετρος θ ως μέτρο της ακολουθιακής συνέχειας ικανοποιεί τις σχέσεις:*

1. $\text{Min}_{(p_1, \mathbf{K}, p_m)} \theta(p_1, \mathbf{K}, p_m) = 1/m$, και
2. $\theta(1/m, \mathbf{K}, 1/m) = 1/m$.

Απόδειξη. Αφού τα X_t είναι ανεξάρτητα και ισόνομα, η εργοδική κατανομή $\pi = (p_1, \mathbf{K}, p_m)$ της $\{X_t\}$ είναι ίδια με αυτήν του X_0 . Επομένως, από τον ορισμό της παραμέτρου της ακολουθιακής συνέχειας, έχουμε $\theta = \sum_{i=1}^m p_i^2$. Τα αποτελέσματα 1 και 2 είναι άμεσες συνέπειες της μεθόδου των πολλαπλασιαστών του Lagrange. Αναλυτικότερα το ζητούμενο προκύπτει ως εξής:

Θέλουμε να βρούμε τα τοπικά ακρότατα της συνάρτησης

$$\theta = \sum_{i=1}^m p_i^2$$

με τον περιορισμό

$$p_1 + p_2 + \dots + p_m = 1 \quad \text{ή αλλιώς} \quad \varphi = p_1 + p_1 + \dots + p_1 - 1 = 0.$$

Θεωρούμε τη συνάρτηση

$$F = \theta + \lambda \varphi = p_1^2 + p_2^2 + \mathbf{K} + p_m^2 + \lambda(p_1 + p_2 + \mathbf{K} + p_m - 1)$$

Και το επόμενο σύστημα ως προς $p_1, \mathbf{K}, p_m, \lambda$:

$$\left\{ \begin{array}{l} \frac{\partial F}{\partial p_1} = 2p_1 + \lambda = 0 \\ \frac{\partial F}{\partial p_2} = 2p_2 + \lambda = 0 \\ \mathbf{M} \\ \frac{\partial F}{\partial p_m} = 2p_m + \lambda = 0 \\ \varphi = p_1 + p_2 + \mathbf{K} + p_m - 1 = 0 \end{array} \right. \quad (1)$$

Από τη λύση του συστήματος (1) θα προκύψουν τα σημεία στα οποία η θ ενδέχεται να έχει τοπικό ακρότατο. Από το σύστημα προκύπτει ότι $p_1 = p_2 = \mathbf{K} = p_m$ και λόγω του περιορισμού $\varphi = p_1 + p_2 + \dots + p_m - 1 = 0$ προκύπτει

$$p_1 = p_2 = \mathbf{K} = p_m = \frac{1}{m} \text{ και } \lambda = -\frac{2}{m}.$$

Επομένως στο σημείο $P_0\left(\frac{1}{m}, \frac{1}{m}, \mathbf{L}, \frac{1}{m}\right)$ η θ ενδέχεται να έχει τοπικό ακρότατο.

Για $P_0\left(\frac{1}{m}, \frac{1}{m}, \mathbf{L}, \frac{1}{m}\right)$ ο πίνακας $A_{1 \times m} = \left(\frac{\partial \varphi}{\partial p_1} \quad \frac{\partial \varphi}{\partial p_2} \quad \mathbf{L} \quad \frac{\partial \varphi}{\partial p_m} \right) = (1 \quad 1 \quad \mathbf{K} \quad 1)$ είναι τάξης

1 και επομένως η θ ενδέχεται να έχει τοπικό ακρότατο στο P_0 . Για την τιμή $\lambda = -\frac{2}{m}$ έχουμε:

$$F = \theta + \lambda\varphi = p_1^2 + p_2^2 + \mathbf{K} + p_m^2 - \frac{2}{m}(p_1 + p_2 + \mathbf{K} + p_m - 1),$$

$$\frac{\partial^2 F}{\partial p_1^2} = 2, \quad \frac{\partial^2 F}{\partial p_1 \partial p_2} = 0, \dots, \quad \frac{\partial^2 F}{\partial p_1 \partial p_m} = 0$$

$$\frac{\partial^2 F}{\partial p_2 \partial p_1} = 0, \quad \frac{\partial^2 F}{\partial p_2^2} = 2, \dots, \quad \frac{\partial^2 F}{\partial p_2 \partial p_m} = 0$$

$$\dots \dots \dots$$

$$\frac{\partial^2 F}{\partial p_m \partial p_1} = 0, \quad \frac{\partial^2 F}{\partial p_m \partial p_2} = 0, \dots, \quad \frac{\partial^2 F}{\partial p_m^2} = 2$$

και ο Εσσιανός πίνακας της F στο P_0 είναι ο

$$\mathbf{H}_{m \times m} = \begin{pmatrix} \frac{\partial^2 F(P_0)}{\partial p_1^2} & \frac{\partial^2 F(P_0)}{\partial p_1 \partial p_2} & \mathbf{L} & \frac{\partial^2 F(P_0)}{\partial p_1 \partial p_m} \\ \frac{\partial^2 F(P_0)}{\partial p_2 \partial p_1} & \frac{\partial^2 F(P_0)}{\partial p_2^2} & \mathbf{L} & \frac{\partial^2 F(P_0)}{\partial p_2 \partial p_m} \\ \mathbf{L} & \mathbf{L} & \mathbf{L} & \mathbf{L} \\ \frac{\partial^2 F(P_0)}{\partial p_m \partial p_1} & \frac{\partial^2 F(P_0)}{\partial p_m \partial p_2} & \mathbf{L} & \frac{\partial^2 F(P_0)}{\partial p_m^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 & \mathbf{L} & 0 \\ 0 & 2 & \mathbf{L} & 0 \\ \mathbf{L} & \mathbf{L} & \mathbf{L} & \mathbf{L} \\ 0 & 0 & \mathbf{L} & 2 \end{pmatrix}.$$

Οι πραγματικοί αριθμοί h_1, h_2, \dots, h_m που επαληθεύουν την ισότητα πινάκων

$$A_{1 \times m} \begin{pmatrix} h_1 \\ h_2 \\ \mathbf{M} \\ h_m \end{pmatrix} = (1 \quad 1 \quad \mathbf{K} \quad 1) \begin{pmatrix} h_1 \\ h_2 \\ \mathbf{M} \\ h_m \end{pmatrix}$$

είναι εκείνοι που επαληθεύουν τη σχέση

$$h_1 + h_2 + \dots + h_m = 0. \quad (2)$$

Θεωρούμε την τετραγωνική μορφή

$$\begin{aligned} q &= \sum_{i=1}^m \sum_{j=1}^m a_{ij} h_i h_j = (h_1 \quad h_2 \quad \mathbf{K} \quad h_m) \mathbf{H}_{m \times m} \begin{pmatrix} h_1 \\ h_2 \\ \mathbf{M} \\ h_m \end{pmatrix} = (h_1 \quad h_2 \quad \mathbf{K} \quad h_m) \begin{pmatrix} 2 & 0 & \mathbf{K} & 0 \\ 0 & 2 & \mathbf{L} & 0 \\ \mathbf{K} & \mathbf{K} & \mathbf{K} & \mathbf{L} \\ 0 & 0 & \mathbf{K} & 2 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \mathbf{M} \\ h_m \end{pmatrix} \\ &= 2(h_1 \quad h_2 \quad \mathbf{K} \quad h_m) \begin{pmatrix} h_1 \\ h_2 \\ \mathbf{M} \\ h_m \end{pmatrix} = 2(h_1^2 + h_2^2 + \mathbf{K} + h_m^2), \end{aligned}$$

όπου a_{ij} είναι τα στοιχεία του πίνακα \mathbf{H} .

Επομένως η τετραγωνική μορφή q είναι θετική ($q > 0$) για όλους τους μη μηδενικούς

πίνακες $\begin{pmatrix} h_1 \\ h_2 \\ \mathbf{M} \\ h_m \end{pmatrix}$, οπότε η θ έχει τοπικό ελάχιστο στο σημείο $P_0(\frac{1}{m}, \frac{1}{m}, \mathbf{L}, \frac{1}{m})$ υπό τη συνθήκη

$$\varphi = p_1 + p_1 + \dots + p_1 - 1 = 0$$

Η ελάχιστη τιμή του είναι ίση με $\theta = \sum_{i=1}^m p_i^2 = \sum_{i=1}^m \frac{1}{m^2} = \frac{1}{m}$.

Αυτό ολοκληρώνει την απόδειξη. □

Έστω $\{X_0, X_1, \mathbf{K}, X_{v(T)=n}\}$ μια ομογενής στάσιμη μαρκοβιανή αλυσίδα m καταστάσεων με αρχικές πιθανότητες $\boldsymbol{\pi}$. Οι πιθανότητες μεταπηδήσεως $p_{ij}, i, j = 1, \mathbf{K}, m$ είναι συχνά άγνωστες σε πραγματικές καταστάσεις και το πρόβλημα της εκτίμησής τους έχει μελετηθεί από πολλούς ερευνητές. Μία προσέγγιση περιγράφεται παρακάτω. Δοθέντος του $v(T) = n$, έστω $N_{ij}(n), i, j = 1, \mathbf{K}, m$ ο συνολικός αριθμός των X_t έτσι ώστε $X_{t-1} = i$ και $X_t = j$, για $i, j = 1, \mathbf{K}, m$. Η δεσμευμένη συνάρτηση πιθανότητας των p_{ij} δοθέντων των $v(T) = n$ και $\{x_0, \mathbf{K}, x_n\}$, μπορεί να γραφτεί ως εξής:

$$L(p_{ij} | x_0, \mathbf{K}, x_n) \propto \boldsymbol{\pi}(x_0) \prod_{i,j=1}^m p_{ij}^{N_{ij}(n)}$$

όπου το σύμβολο \propto σημαίνει ότι η ποσότητα $L(p_{ij} | x_0, \mathbf{K}, x_n)$ είναι ανάλογη της ποσότητας

$\boldsymbol{\pi}(x_0) \prod_{i,j=1}^m p_{ij}^{N_{ij}(n)}$, όπου $\boldsymbol{\pi}(x_0)$ είναι το διάνυσμα αρχικών πιθανοτήτων δοθείσης της

πιθανότητας x_0 . Με άλλα λόγια, σημαίνει ότι η ποσότητα $\boldsymbol{\pi}(x_0) \prod_{i,j=1}^m p_{ij}^{N_{ij}(n)}$ αν πολλαπλασιαστεί με κατάλληλη σταθερά, δίνει την ποσότητα $L(p_{ij} | x_0, \mathbf{K}, x_n)$.

Σημειωτέον ότι τα στατιστικά $N_{ij}(n), i, j = 1, \mathbf{K}, m$, αποτελούν επαρκές στατιστικές συναρτήσεις για τις πιθανότητες μεταπηδήσεως p_{ij} . Οι εκτιμητές μεγίστης πιθανοφάνειας των p_{ij} , δοθέντος ότι $v(T) = n$, είναι ίσες με

$$\hat{p}_{ij} = \frac{N_{ij}}{\sum_{j=1}^m N_{ij}}.$$

Αν η τυχαία μεταβλητή $v(T)$ είναι ανεξάρτητη της ακολουθίας $\{X_t\}$, τότε τα $\hat{p}_{ij}, i, j = 1, \mathbf{K}, m$, είναι οι εκτιμητές μεγίστης πιθανοφάνειας για τα p_{ij} και ανεξάρτητα από το $v(T) = n$, δηλαδή ανεξάρτητα του πλήθους των επισκέψεων. Με άλλα λόγια, οι εκτιμητές των πιθανοτήτων μεταπηδήσεως μπορούν να υπολογιστούν χωρίς να είναι απαραίτητως γνωστό το πλήθος των επισκέψεων.

Αφού τα $N_{ij}(n), i, j = 1, \mathbf{K}, m$ είναι επαρκείς στατιστικές συναρτήσεις και η ποσότητα $SECON = S_{v(T)}/v(T)$ είναι ένας αμερόληπτος εκτιμητής για το θ , τότε εφαρμόζοντας το

θεώρημα των Rao-Blackwell για το *SECON*, μπορούμε εύκολα να πάρουμε τον αμερόληπτο εκτιμητή ελάχιστης διακύμανσης (α.ε.ε.δ.) για το θ .

Θεώρημα 3.5 Αν $v(T)$ και $\{X_t\}$ είναι ανεξάρτητες, τότε

$$E\left(\frac{S_{v(T)}}{v(T)} \mid N_{ij}, v(T) = n\right) = \sum_{i=1}^n \frac{N_{ii}(n)}{n} \quad (3.23)$$

Απόδειξη. Δοθέντων των επαρκών στατιστικών $v(T) = n$ και $N_{ij}(n)$, $i, j = 1, \mathbf{K}, n$, η πιθανότητα να συμβεί το ενδεχόμενο $\{X_{k-1} = i \text{ και } X_k = j\}$, για όλα τα $k = 1, \mathbf{K}, n$, είναι ίση με

$$P(X_{k-1} = i, X_k = j \mid N_{ij}(n), v(T) = n) = N_{ij}(n) / n.$$

Από τον ορισμό του Y_k στην εξίσωση (3.3) έπεται ότι

$$\begin{aligned} P(Y_k = 1 \mid N_{ij}(n), v(T) = n) &= \sum_{i=1}^m P(X_{k-1} = i, X_k = i \mid N_{ij}(n), v(T) = n) \\ &= \sum_{i=1}^m N_{ii}(n) / n. \end{aligned} \quad (3.24)$$

Αν η $v(T)$ είναι ανεξάρτητη από την ακολουθία των τυχαίων μεταβλητών $\{X_k\}$, τότε

$$\begin{aligned} E\left(\frac{S_{v(T)}}{v(T)} \mid N_{ij}(n), v(T) = n\right) &= \frac{1}{n} E(S_n \mid N_{ij}(n)) \\ &= \frac{1}{n} \sum_{k=1}^n E(Y_k \mid N_{ij}(n)) \\ &= \frac{1}{n} \sum_{k=1}^n P(Y_k = 1 \mid N_{ij}(n)) \\ &= P(Y_1 = 1 \mid N_{ij}(n)) \\ &= \sum_{i=1}^m N_{ii}(n) / n. \end{aligned} \quad (3.25)$$

Αυτό ολοκληρώνει την απόδειξη. □

Σύμφωνα με το θεώρημα των Rao-Blackwell, ο εκτιμητής

$$\hat{\theta}_{v(T)} = \frac{1}{v(T)} \sum_{i=1}^m N_{ii}(v(T)) \quad (2.26)$$

θα είναι α.ε.ε.δ. του θ . Εναλλακτικά, μπορούμε να γράψουμε τον τελευταίο εκτιμητή στη μορφή

$$\hat{\theta}_{v(T)} = \sum_{i=1}^m [N_i(v(T))/v(T)][N_{ii}(v(T))/N_i(v(T))] = \sum_{i=1}^m \hat{\pi}_i \hat{p}_{ii},$$

όπου $N_i(v(T)) = \sum_{j=1}^m N_{ij}(v(T))$. Προφανώς ο νέος εκτιμητής $\hat{\theta}_{v(T)}$ είναι αμερόληπτος και συνεπής για την παράμετρο θ του μέτρου ακολουθιακής συνέχειας.

Παρατήρηση. Ο $\hat{\theta}_{v(T)}$ είναι αμερόληπτος μόνο αν η αρχική κατανομή του X_0 είναι η εργοδική κατανομή π της μαρκοβιανής αλυσίδας $\{X_t\}$. Διαφορετικά, η αναμενόμενη τιμή του $\hat{\theta}_{v(T)}$ είναι πιθανό να εξαρτάται από την αρχική κατανομή.

Κ Ε Φ Α Λ Α Ι Ο 4

Η ακριβής κατανομή του *UPC*

4.1 Εισαγωγή

Ένας δείκτης συνέχειας μπορεί να θεωρηθεί ως συνάρτηση μιας σειράς τυχαίων μεταβλητών που ορίζονται από δοκιμές πολλών καταστάσεων όπου το πλήθος των καταστάσεων αυτών ισούται με το πλήθος των διαθέσιμων παροχών μέσα σε ένα σύστημα υγειονομικής περίθαλψης και το μήκος της σειράς αντιστοιχεί στο πλήθος των επισκέψεων του ασθενούς εντός ενός συγκεκριμένου χρονικού διαστήματος. Για παράδειγμα, ο δείκτης συνέχειας του συνήθους παροχέα, ο *UPC*, ορίζεται ως ο λόγος των επισκέψεων στο συνήθη παροχέα, έστω στο γιατρό “1”, κατά την περίοδο του follow-up. Στην πράξη, το ποιος παροχέας περιποιείται τον ασθενή σε μια συγκεκριμένη επίσκεψη θα εξαρτάται συνήθως από το ποιοι παροχείς χρησιμοποιήθηκαν στις προηγούμενες επισκέψεις και επομένως είναι σημαντικό να εξεταστούν οι κατανομές των μέτρων αυτών κάτω από κάποια άλλη υπόθεση πλην εκείνης της τυχαίας ανάθεσης του ασθενούς στον παροχέα. Στην επόμενη ενότητα παρουσιάζουμε την ακριβή κατανομή του δείκτη *UPC*, ενός από τους πιο ευρέως χρησιμοποιούμενους δείκτες για τη συνέχεια της περίθαλψης, κάτω από την υπόθεση της εξάρτησης από μαρκοβιανή αλυσίδα μεταξύ των επισκέψεων.

Κατά το σχεδιασμό πολλών προγραμμάτων υγειονομικής περίθαλψης, για παράδειγμα, μια φυσική επιλογή για το συνήθη παροχέα είναι ο γιατρός για τη στοιχειώδη περίθαλψη (παθολόγος). Όταν μια τέτοια φυσική επιλογή δεν υπάρχει a priori, ο παροχέας με τις περισσότερες επισκέψεις κατά τη διάρκεια μιας περιόδου παρατήρησης θα μπορούσε να θεωρηθεί ως ο συνήθης. Για λόγους απλότητας, ο προηγούμενος ορισμός χρησιμοποιείται για τα συμπεράσματά μας στην επόμενη ενότητα. Αξίζει επίσης να αναφερθεί ότι για ένα δεδομένο αριθμό επισκέψεων, ο *UPC* δεν εξαρτάται από την κατανομή των επισκέψεων που έγιναν σε άλλους παροχείς και ότι απλά μετράει τη συγκέντρωση της συνέχειας στον συνήθη παροχέα. Ως εκ τούτου, αυτό το εξαιρετικά απλό μέτρο χρησιμοποιείται συχνά σε συνδυασμό

με άλλους δείκτες, όπως το στατιστικό ακολουθιακής συνέχειας *SECON*, για τη διασφάλιση μιας πιο ισορροπημένης άποψης των επιπέδων συνέχειας.

Στη συνέχεια θα παρουσιαστεί η κατανομή του δείκτη *UPC* για την περίπτωση της εξάρτησης μεταξύ των επισκέψεων. Μια τέτοια κατανομή προσφέρει τη βάση για περαιτέρω μελέτη των στατιστικών ιδιοτήτων, όπως η ισχύς, και για την ανάπτυξη συγκεντρωτικών μεθόδων για τη σύνοψη και σύγκριση των επιπέδων συνέχειας μεταξύ ατόμων ή ομάδων που τους παρέχεται περίθαλψη καλύτερη από εκείνη που γίνεται με τυχαίο τρόπο.

4.2 Η ακριβής κατανομή

Έστω $\{X_t\}_{t=1}^{v(T)}$ μια ακολουθία τυχαίων μεταβλητών που σχηματίζουν μια μη υποβιβάσιμη μη περιοδική ομογενή μαρκοβιανή αλυσίδα η οποία ορίζεται σε ένα χώρο καταστάσεων $S = \{1, \mathbf{K}, m\}$ με πίνακα πιθανοτήτων μεταπηδήσεως $A = (p_{ij})$ όπου $p_{ij} = P(X_t = j | X_{t-1} = i)$ για κάθε t . Έστω $\pi = (\pi_1, \mathbf{K}, \pi_m)$ η εργοδική κατανομή που σχετίζεται με τη μαρκοβιανή αλυσίδα $\{X_t\}$, δηλαδή για $i=1, \dots, m$, $\pi_i = \lim_{n \rightarrow \infty} P(X_n = i)$. Στα πλαίσια της συνέχειας, το $v(T)$ παριστάνει τον αριθμό των επισκέψεων του ασθενούς σε οποιονδήποτε από τους παροχείς κατά την περίοδο παρατήρησης $[0, T]$ και m είναι το πλήθος των παροχέων στο υπό μελέτη σύστημα υγειονομικής περίθαλψης. Για λόγους ευκολίας, έστω ότι ο παροχέας "1" είναι ο συνήθης παροχέας. Τότε ο δείκτης τού συνήθους παροχέα *UPC* μπορεί να οριστεί μαθηματικά ως

$$UPC = \frac{1}{v(T)} \sum_{t=1}^{v(T)} I_1(X_t) = \frac{N_{v(T)}}{v(T)}, \quad (4.1)$$

όπου $N_{v(T)} = \sum_{t=1}^{v(T)} I_1(X_t)$ και $I_1(X_t)$ είναι η δείκτρια συνάρτηση για το αν ο συνήθης παροχέας "1" χρησιμοποιείται στην t επίσκεψη, δηλαδή για $t = 1, \mathbf{K}, v(T)$,

$$I_1(X_t) = \begin{cases} 1 & X_t = 1 \\ 0 & \text{διαφορετικά.} \end{cases}$$

Σημειωτέον ότι αν η $\{X_t\}$ είναι μια ακολουθία ανεξάρτητων και ισόνομων δοκιμών m καταστάσεων με $P(X_t = 1) = p_1$, τότε υπό τη συνθήκη $v(T) = n$, η N_n ακολουθεί διωνυμική κατανομή με παραμέτρους p_1 και n , και ο *UPC* είναι αμερόληπτος εκτιμητής για το p_1 .

Αξίζει επίσης να σημειωθεί ότι, κάτω από την υπόθεση της μαρκοβιανής εξάρτησης, η κατανομή του N_n (και επομένως του UPC) δεν μπορεί να μελετηθεί μέσω μιας απλής προσέγγισης τύπου “επιτυχία-αποτυχία”, όπου επιτυχία σημαίνει επίσκεψη στον συνήθη παροχέα και αποτυχία οτιδήποτε άλλο, επειδή οι πιθανότητες αλλαγής παροχέα p_{ij} εξαρτώνται από τον παροχέα που δέχεται την επίσκεψη σε μια δεδομένη στιγμή και δεν είναι απαραίτητα οι ίδιες (π.χ. $p_{21} \neq p_{31}$). Συνεπώς, η πληροφορία για το ποιος παροχέας χρησιμοποιείται σε μια δεδομένη επίσκεψη είναι απαραίτητη για την εύρεση της κατανομής και αυτή η επιπλέον πληροφορία, μαζί με μια μέτρηση του πλήθους των επισκέψεων στο συνήθη παροχέα ενσωματώνεται σε μια νέα αλυσίδα δύο συνιστωσών ως εξής.

Δοθέντος ότι $v(T) = n$, ορίζουμε για $t = 1, 2, \mathbf{K}, n$ την αλυσίδα

$$Y_t = (N_t, X_t) \quad (4.2)$$

στο χώρο καταστάσεων

$$\Omega = \{(x, y) : x = 0, 1, \mathbf{K}, n, \text{ και } y = 1, 2, \mathbf{K}, m\}, \quad (4.3)$$

όπου $N_t = \sum_{i=1}^t I_1(X_i)$ είναι το πλήθος των επισκέψεων στον συνήθη παροχέα κατά τη διάρκεια των πρώτων t επισκέψεων και το μέγεθος του χώρου καταστάσεων Ω είναι $d = (n+1) \times m$. Η κατανομή του UPC μπορεί να εξαχθεί με την προβολή του χώρου του Y_t πάνω στο N_t . Τότε ισχύει το ακόλουθο θεώρημα, που δόθηκε από την W.Y.W. Lou (2001):

Θεώρημα 4.1 Δοθέντος του $v(T) = n$ και ότι η αρχική κατανομή του X_0 είναι π , τότε ισχύουν τα ακόλουθα αποτελέσματα:

1. Η ακολουθία $\{Y_t\}$ που ορίζεται στο χώρο καταστάσεων Ω σχηματίζει μια ομογενή μαρκοβιανή αλυσίδα με πίνακα πιθανοτήτων μεταπηδήσεως

$$M = \begin{pmatrix} \mathbf{B} & \mathbf{C} & & \mathbf{0} \\ & \mathbf{B} & \mathbf{C} & \\ \mathbf{0} & \mathbf{O} & \mathbf{O} & \\ & & \mathbf{O} & \mathbf{B} & \mathbf{C} \\ \mathbf{0} & & & \mathbf{0} & \mathbf{B}^* \end{pmatrix}_{d \times d}, \quad (4.4)$$

όπου οι υποπίνακες ορίζονται ως

$$\mathbf{B} = \begin{pmatrix} 0 & p_{12} & \mathbf{L} & p_{1m} \\ 0 & p_{22} & \mathbf{L} & p_{2m} \\ \mathbf{L} & \mathbf{L} & \mathbf{L} & \mathbf{L} \\ 0 & p_{m2} & \mathbf{L} & p_{mm} \end{pmatrix}_{m \times m}, \quad \mathbf{C} = \begin{pmatrix} p_{11} & 0 & \mathbf{L} & 0 \\ p_{21} & 0 & \mathbf{L} & 0 \\ \mathbf{L} & \mathbf{L} & \mathbf{L} & \mathbf{L} \\ p_{m1} & 0 & \mathbf{L} & 0 \end{pmatrix}_{m \times m},$$

και $\mathbf{B}^* = \mathbf{I}_{m \times m}$ είναι ο $m \times m$ μοναδιαίος πίνακας.

2. Η δεσμευμένη κατανομή του UPC_n για $0 \leq x \leq n$ δίνεται από τον τύπο

$$P(N_n = x | v(T) = n) = \xi \mathbf{M}^n \mathbf{U}'(C_x), \quad (4.5)$$

όπου $\xi = (\boldsymbol{\pi}, \mathbf{0}, \mathbf{K}, \mathbf{0})_{1 \times d}$ είναι η αρχική κατανομή της μαρκοβιανής αλυσίδας $\{Y_t\}$ και το $\mathbf{U}'(C_x)$ δηλώνει τον ανάστροφο του διανύσματος γραμμής $\mathbf{U}(C_x) = (0, \mathbf{K}, 0, 1, \mathbf{K}, 1, 0, \mathbf{K}, 0)_{1 \times d}$ που έχει μονάδες στις συντεταγμένες που σχετίζονται με τις καταστάσεις στο σύνολο $C_x = \{(x, 1), \mathbf{K}, (x, m)\}$ και μηδενικά αλλού

3. Ο UPC είναι αμερόληπτος εκτιμητής για το $\sum_{i=1}^m \pi_i p_{i1} = \pi_1$, με την έννοια ότι για κάθε δεδομένο $v(T) = n$ ισχύει

$$E(UPC | v(T) = n) = \pi_1. \quad (4.6)$$

Απόδειξη. Δοθέντος ότι $v(T) = n$, έπεται από τους ορισμούς της μαρκοβιανής αλυσίδας $\{Y_t\}$ και του χώρου καταστάσεων Ω στις εξισώσεις (4.2) και (4.3) αντίστοιχα, ότι οι πιθανότητες μεταπηδήσεως $P(Y_t = (y, j) | Y_{t-1} = (n, i))$ δίνονται από τους τύπους:

$$P(Y_t = (y, j) | Y_{t-1} = (x, i)) = \begin{cases} p_{i1} & \text{αν } y = x, j = 1, \\ p_{ij} & \text{αν } y = x, j = 2, 3, \mathbf{K}, m, \\ 0 & \text{διαφορετικά,} \end{cases} \quad (4.7)$$

για $x = 0, 1, \mathbf{K}, n-1$ και $i = 1, 2, \dots, m$. Για λόγους ευκολίας, ορίζουμε $P(Y_t = (n, i) | Y_{t-1} = (n, i)) \equiv 1$. Οι τύποι αυτοί καθορίζουν τον πίνακα πιθανοτήτων μεταπηδήσεως \mathbf{M} που δίνεται στην εξίσωση (4.4). Επομένως η $N_n = \sum_{i=1}^n I_1(X_i)$ μπορεί να εμφυτευτεί (*imbeddable*) σε πεπερασμένη μαρκοβιανή αλυσίδα και το αποτέλεσμα 2 ισχύει ως άμεση συνέπεια του θεωρήματος εμφύτευσης των Fu and Koutras (1994). Αφού $\boldsymbol{\pi}$ είναι η εργοδική κατανομή της ακολουθίας $\{X_t\}$, θα ισχύει $\boldsymbol{\pi} \mathbf{A} = \boldsymbol{\pi}$ (ή $\sum_{i=1}^m \pi_i p_{ij} = \pi_j$ για όλα τα

$j = 1, \mathbf{K}, m$). Ως εκ τούτου, έπεται από τις εξισώσεις των Charman-Kolmogoron ότι για κάθε δεδομένο $v(T) = n$ ισχύει

$$\begin{aligned}
 E(UPC | v(T) = n, \boldsymbol{\pi}) &= E\left(\sum_{i=1}^n I_1(X_i) / n | \boldsymbol{\pi}\right) \\
 &= \frac{1}{n} \sum_{i=1}^n E(I_1(X_i) | \boldsymbol{\pi}) \\
 &= \frac{1}{n} \sum_{i=1}^n P(X_i = 1 | \boldsymbol{\pi}) \\
 &= \frac{1}{n} \sum_{i=1}^n \boldsymbol{\pi} \mathbf{A}^i (1, 0, \mathbf{K}, 0)'_{m \times 1} \\
 &= \pi_1.
 \end{aligned} \tag{4.8}$$

Αυτό ολοκληρώνει την απόδειξη. □

Σημειώνουμε ότι αν η $\{X_t\}_{t=1}^n$ είναι μια ακολουθία ανεξάρτητων και ισόνομων δοκιμών m καταστάσεων, η εξίσωση (4.6) οδηγείται στην ισότητα $E(UPC) = p_1$. Πράγματι, ως αποτέλεσμα του Πορίσματος 4.1 που δίνεται παρακάτω, η απλοποίηση αυτή ισχύει εφόσον όλα τα p_{i1} του πίνακα \mathbf{C} είναι ίσα. Επιπλέον, αν η αρχική κατανομή του X_0 είναι $\boldsymbol{\pi}^*$ και $\boldsymbol{\pi}^* \neq \boldsymbol{\pi}$, τότε, μόνο αν το n είναι πολύ μεγάλο, ισχύει $E(UPC | v(T) = n, \boldsymbol{\pi}^*) \sim \pi_1$. Αυτό μπορούμε να το δούμε από το αριθμητικό παράδειγμα που δίνεται στην ενότητα 4.3.

Πόρισμα 4.1 Δοθέντος ότι $v(T) = n$, αν $p_{i1} = p_1$, $i = 1, \mathbf{K}, m$, το N_n θα ακολουθεί τη διωνυμική κατανομή

$$P(N_n = x) = \binom{n}{x} p_1^x (1 - p_1)^{n-x}, \quad x = 0, 1, \mathbf{K}, n. \tag{4.9}$$

Απόδειξη. Αφού $p_{11} = p_{21} = \mathbf{K} = p_{m1} = p_1$, η σχέση μεταξύ των \mathbf{BC} και \mathbf{CB} μπορεί να απλοποιηθεί σε

$$\boldsymbol{\pi} \mathbf{BC} \mathbf{1}' = \boldsymbol{\pi} \mathbf{CB} \mathbf{1}' = p_1 (1 - p_1).$$

Από την εξίσωση (4.5) και το γεγονός ότι

$$\xi \mathbf{M}^n = [\pi \mathbf{B}, \pi(\mathbf{B}^{n-1} \mathbf{C} + \mathbf{B}^{n-2} \mathbf{C} \mathbf{B} + \mathbf{K} + \mathbf{C} \mathbf{B}^{n-1}), \mathbf{K}, \pi(\mathbf{B}^{n-1} \mathbf{C} + \mathbf{K} + \mathbf{C}^{n-1} \mathbf{B}), \pi \mathbf{C}^n],$$

η κατανομή του N_n , ύστερα από κάποιους χρονοβόρους αλλά εύκολους αλγεβρικούς υπολογισμούς, οδηγεί στην εξίσωση (4.9). \square

Για την ειδική αυτή περίπτωση, το αποτέλεσμα είναι ανεξάρτητο της αρχικής κατανομής π και του πλήθους των διαθέσιμων παροχών m .

Πόρισμα 4.2 Αν η $v(T)$ έχει μια κατανομή $\mu_{v(T)}$ στο $J^+ = \{1, 2, \mathbf{K}, \}$ και είναι ανεξάρτητη της ακολουθίας $\{X_t\}$, τότε

$$(i) \quad P(UPC \leq \alpha) = \sum_{n=1}^{\infty} \sum_{x=0}^{[an]} \xi \mathbf{M}^n \mathbf{U}'(C_x) \mu(n), \quad \text{για } 0 < \alpha < 1 \quad (4.10)$$

$$(ii) \quad E(UPC) = \pi_1.$$

Απόδειξη. Αφού η $v(T)$ είναι ανεξάρτητη της ακολουθίας $\{X_t\}$, η μη δεσμευμένη κατανομή του UPC και η μέση τιμή της απορρέουν απευθείας από το Θεώρημα 4.1. \square

Στην πράξη, υποθέτουμε συχνά ότι το πλήθος των επισκέψεων $v(T)$ ακολουθεί μια περικομμένη κατανομή Poisson με παράμετρο λ , έτσι ώστε $\mu(n) = \frac{\lambda^n e^{-\lambda}}{n!(1 - e^{-\lambda})}$, $n = 1, 2, \mathbf{K}$.

Όταν υπάρχουν επαρκείς πληροφορίες, το λ μπορεί να εκτιμηθεί από τα δεδομένα και να χρησιμοποιηθεί για τη συνοπτική παρουσίαση των επισκέψεων των ασθενών του υπό μελέτη πληθυσμού.

4.3 Παράδειγμα

Έστω μια ακολουθία $\{X_t\}$ επισκέψεων ασθενών σε μια κλινική με 3 διαθέσιμους παροχείς ($m=3$) και ας υποθέσουμε ότι η $\{X_t\}$ είναι μια μη υποβιβάσιμη, μη περιοδική

ομογενής μαρκοβιανή αλυσίδα με πίνακα πιθανοτήτων μεταπηδήσεως $A = (p_{ij}), i, j = 1, 2, 3$.

Έστω ότι ο αριθμός “1” δηλώνει τον συνήθη παροχέα του ασθενούς.

Έστω $v(T)=2$ και ότι η αρχική κατανομή του X_0 είναι $P(X_0)$. Τότε μπορούμε να βρούμε την κατανομή του UPC από το Θεώρημα 4.1. με τις ακόλουθες συνιστώσες:

(1) την εμφυτευμένη μαρκοβιανή αλυσίδα $\{Y_t = (N_t, X_t); t = 1, 2\}$

(2) τον χώρο καταστάσεων $\Omega = \{(0,1), (0,2), (0,3), (1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\}$.

Για να γίνει απολύτως κατανοητή η έννοια του χώρου καταστάσεων Ω της σύνθετης αλυσίδας $Y_t = (N_t, X_t)$, αρκεί να εξηγήσουμε τι σημαίνει κάποια από τις 9 καταστάσεις. Για παράδειγμα, η κατάσταση (1,3) σημαίνει ότι έχει χρησιμοποιηθεί 1 φορά ο συνήθης παροχέας ενώ κατά την επίσκεψη t χρησιμοποιήθηκε ο 3^{ος} παροχέας. Με την ίδια λογική εξηγούνται και οι υπόλοιπες 8 καταστάσεις.

(3) την αρχική κατανομή του Y_0 , η οποία περιγράφεται από το διάνυσμα $\xi = (P(X_0)_{1 \times 3}, 0, 0, 0, 0, 0)$

(4) τον 9×9 πίνακα πιθανοτήτων μεταπηδήσεως

$$M = \begin{matrix} (0,1) \\ (0,2) \\ (0,3) \\ (1,1) \\ (1,2) \\ (1,3) \\ (2,1) \\ (2,2) \\ (2,3) \end{matrix} \begin{bmatrix} 0 & p_{12} & p_{13} & p_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{22} & p_{23} & p_{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{32} & p_{33} & p_{31} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & p_{12} & p_{13} & p_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{22} & p_{23} & p_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{32} & p_{33} & p_{31} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} \mathbf{B} & \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \mathbf{C} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix},$$

$$\text{όπου } \mathbf{B} = \begin{bmatrix} 0 & p_{12} & p_{13} \\ 0 & p_{22} & p_{23} \\ 0 & p_{32} & p_{33} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} p_{11} & 0 & 0 \\ p_{21} & 0 & 0 \\ p_{31} & 0 & 0 \end{bmatrix} \quad \text{και} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

όπως ορίστηκαν στο Θεώρημα 4.1, και τέλος:

(1) τη διαμέριση του χώρου καταστάσεων Ω στη μορφή $\Omega = C_0 \cup C_1 \cup C_2$, όπου $C_0 = \{(0,1), (0,2), (0,3)\}$, $C_1 = \{(1,1), (1,2), (1,3)\}$, $C_2 = \{(2,1), (2,2), (2,3)\}$.

Για $x = 0,1,2$, έχουμε

$$P(UPC = \frac{x}{v(T)} | v(T) = 2, \xi) = P(N_2 = x | \xi) = \xi_{1 \times 9} \mathbf{M}^2 U'(C_x).$$

Οι κατανομές του UPC για μεγαλύτερο πλήθος επισκέψεων $v(T) > 2$ μπορούν να βρεθούν με παρόμοιο τρόπο.

Για το αριθμητικό μας παράδειγμα που συνοψίζεται στον Πίνακα 4.1, θεωρούμε $v(T) = 2, 5$ και 10 , για τις ακόλουθες 4 περιπτώσεις αρχικών πιθανοτήτων και πιθανοτήτων μεταπηδήσεως:

(Ia) Οι δοκιμές X_t είναι ανεξάρτητες και ισόνομες με ίσες πιθανότητες

$$\mathbf{A} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}, \quad (4.11)$$

και η αρχική κατανομή του X_0 είναι η εργοδική κατανομή $\boldsymbol{\pi} = (1/3, 1/3, 1/3)$.

(Ib) Ο πίνακας πιθανοτήτων μεταπηδήσεως \mathbf{A} όπως ορίζεται στην (4.11), αλλά η αρχική κατανομή του X_0 είναι $\boldsymbol{\pi}^* = (1, 0, 0)$.

(IIa) Η ακολουθία $\{X_t\}$ είναι ομογενής μαρκοβιανά εξαρτημένη με πίνακα πιθανοτήτων μεταπηδήσεως

$$\mathbf{A}^* = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}, \quad (4.12)$$

και η αρχική κατανομή του X_0 είναι η ίδια με αυτήν που ορίστηκε στο (Ia).

(IIb) Ο πίνακας πιθανοτήτων μεταπηδήσεως \mathbf{A}^* δίνεται και πάλι από τη σχέση (4.12) αλλά η αρχική κατανομή όπως στο (Ib).

Στον Πίνακα 4.1 για τις περιπτώσεις (Ia) και (Ib) όπου οι δοκιμές X_t είναι ανεξάρτητες και ισόνομες με ίσες πιθανότητες, οι κατανομές του UPC είναι όμοιες. Αυτό δείχνει ότι κάτω από την υπόθεση ανεξαρτησίας και ισονομίας για τις X_t , η αρχική κατανομή του X_0 δεν επηρεάζει την κατανομή του UPC . Ωστόσο, δε συμβαίνει το ίδιο όταν οι X_t δεν είναι ανεξάρτητες και ισόνομες, όπως στις περιπτώσεις (IIa) και (IIb). Προσέχουμε ότι στις περιπτώσεις (IIa) και (IIb), όπου οι αρχικές κατανομές του X_0 είναι ίσες με την ίδια εργοδική κατανομή, οι μέσες τιμές του UPC είναι ίδιες κάτω και από τις

δύο υποθέσεις, αλλά οι κατανομές είναι πολύ διαφορετικές. Αυτός είναι ένας λόγος για τον οποίο χρειαζόμαστε την ακριβή κατανομή του UPC .

$v(T)$	x	Πιθανότητας		
		Συνάρτηση Ia=Ib	IIa	IIb
$n = 2$	0	0.4445	0.5000	0.375
	1	0.4444	0.3333	0.375
	2	0.1111	0.1667	0.250
	$E(UPC n)$	0.3333	0.3333	0.4375
$n = 5$	0	0.1317	0.2109	0.1582
	1	0.3292	0.2813	0.2637
	2	0.3292	0.2526	0.2656
	3	0.1646	0.1615	0.1875
	4	0.0412	0.0729	0.0938
	5	0.0041	0.0208	0.0312
	$E(UPC n)$	0.3333	0.3333	0.3777
$n = 10$	0	0.0173	0.0501	0.0375
	1	0.0867	0.1224	0.1042
	2	0.1951	0.1835	0.1696
	3	0.2601	0.2038	0.2010
	4	0.2276	0.1795	0.1875
	5	0.1366	0.1291	0.1422
	6	0.0569	0.0763	0.0885
	7	0.0163	0.0368	0.0449
	8	0.0030	0.0140	0.0181
	9	0.0003	0.0039	0.0054
	10	0.00002	0.0007	0.0010
$E(UPC n)$	0.3333	0.3333	0.3555	

Πίνακας 4.1 Οι αναμενόμενες τιμές και οι ακριβείς πιθανότητες του UPC δοθέντος $m = 3$ για $v(T) = 2, 5$ και 10 , κάτω από τις υποθέσεις: (I) i.i.d. με ίσες πιθανότητες και (II) μαρκοβιανής εξάρτησης, με αρχικές κατανομές για την περίπτωση (a) $\pi = (1/3, 1/3, 1/3)$ και την περίπτωση (b) $\pi^* = (1, 0, 0)$.

4.4 Παρατηρήσεις

Οι αρχικές κατανομές π και π^* του παραδείγματος της ενότητας 4.3 αντιπροσωπεύουν δύο συνήθη πραγματικά σενάρια. Με την έννοια της συνέχειας, η $\pi = (1/3, 1/3, 1/3)$ σημαίνει ότι ο κάθε παροχέας έχει τις ίδιες πιθανότητες να χρησιμοποιηθεί στην πρώτη επίσκεψη του ασθενούς στην κλινική και η $\pi^* = (1, 0, 0)$ δείχνει ότι ο συνήθης παροχέας χρησιμοποιείται πάντα πρώτος για ένα συγκεκριμένο επεισόδιο περίθαλψης (π.χ. ο παθολόγος για τη βασική περίθαλψη σε πολλούς οργανισμούς ιατρικής υποστήριξης). Ωστόσο, σε άλλα κλινικά ιδρύματα ο ασθενής μπορεί να αναζητήσει τον πρώτο διαθέσιμο παροχέα σε ορισμένες επισκέψεις και οι αρχικές κατανομές να είναι πολύ διαφορετικές από τις εργοδικές. Όπως φαίνεται στον Πίνακα 4.1, όταν το πλήθος των επισκέψεων n είναι μικρό, η κατανομή του UPC εξαρτάται ισχυρά από την αρχική κατανομή $P(X_0)$ και ο UPC δεν είναι αμερόληπτος εκτιμητής. Εντούτοις, η επίδραση του $P(X_0)$ στις πιθανότητες του UPC μειώνεται πολύ γρήγορα καθώς αυξάνεται το n και η μεροληψία γίνεται αμελητέα για μεγάλο αριθμό επισκέψεων.

Στην ενότητα 4.2, η ακριβής κατανομή του UPC εξάγεται κάτω από την πιο ρεαλιστική υπόθεση της μαρκοβιανής εξάρτησης για μια ακολουθία επισκέψεων, χρησιμοποιώντας την προσέγγιση της εμφύτευσης σε πεπερασμένη μαρκοβιανή αλυσίδα. Ο προτεινόμενος τύπος είναι κατάλληλος όχι μόνο για ανεξάρτητες, ισόνομες και ομογενείς δοκιμές πολλών καταστάσεων με μαρκοβιανή εξάρτηση, αλλά μπορεί επίσης να επεκταθεί σε μη ομογενείς περιπτώσεις με μαρκοβιανή εξάρτηση. Οι ροπές του UPC μπορούν εύκολα να υπολογιστούν από το Θεώρημα 4.1 αντικαθιστώντας απλά το διάνυσμα διαμέρισης $U'(C_x)$ με $x^r U'(C_x)$ στην εξίσωση (4.5) για να πάρουμε την αντίστοιχη ροπή r τάξεως.

Σχετικά με τον παραπάνω πίνακα, για τον υπολογισμό της κατανομής χρησιμοποιήθηκε το στατιστικό πακέτο S-plus. Σημειώνουμε ότι αφού ο πίνακας πιθανοτήτων μεταπηδήσεως M περιέχει ουσιαστικά δύο μη μηδενικούς πίνακες, δηλαδή τον B και τον C , που βρίσκονται επί της κύριας διαγωνίου και στο πάνω διαγώνιο μέρος, ο χρόνος υπολογισμού μπορεί να μειωθεί σημαντικά μέσω αποτελεσματικού προγραμματισμού, όπως με τη χρήση αναδρομικών εξισώσεων. Είναι προφανές από την εμφυτευμένη μαρκοβιανή αλυσίδα ότι το μέγεθος των δύο πινάκων καθορίζεται από τον

αριθμό των διαθέσιμων παροχών m , ο οποίος στην πράξη είναι συνήθως μικρότερος από 10. Επομένως, η εκτίμηση της κατανομής περιλαμβάνει κυρίως τον πολλαπλασιασμό πινάκων μετρίου μεγέθους και δεν απαιτεί χρονοβόρους υπολογισμούς.

Η παραπάνω ανάλυση επικεντρώθηκε στη συζήτηση για τη δεσμευμένη κατανομή του UPC δοθέντος του αριθμού των επισκέψεων $v(T) = n$. Όταν ο αριθμός των επισκέψεων $v(T)$ θεωρηθεί τυχαία μεταβλητή, η μη δεσμευμένη κατανομή μπορεί να βρεθεί σύμφωνα με το Πόρισμα 4.2. Από πρακτική άποψη, η μη δεσμευμένη κατανομή είναι χρήσιμη όταν συγκεντρώνουμε πληροφορίες από άτομα με διαφορετικό αριθμό επισκέψεων. Όταν χρειαζόμαστε τέτοιες πληροφορίες για τη σύγκριση διαφορετικών πληθυσμών, είναι απαραίτητο να εφαρμόσουμε κατάλληλες μεθόδους για το άθροισμα δεικτών συνέχειας που βασίζονται στα άτομα, λαμβάνοντας συγχρόνως υπόψη τη μεταβλητότητα στο πλήθος και των παροχών και των επισκέψεων, καθώς και την εξάρτηση μεταξύ των επισκέψεων.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΚΕΦΑΛΑΙΟ 5

Η ακριβής κατανομή του *NOP*

5.1 Εισαγωγή

Το πλήθος των διαφορετικών παροχών που χρησιμοποιούνται από ένα άτομο κατά την περίοδο του follow-up μπορούν να χρησιμεύσουν ως ένα μέτρο της έλλειψης συγκέντρωσης. Με τη χρήση κάποιας φθίνουσας συνάρτησης, ο δείκτης *NOP* μπορεί να μετατραπεί σε ένα μέτρο που έχει επιθυμητό εύρος, όπως μεταξύ του μηδενός και της μονάδας, έτσι ώστε οι υψηλές τιμές να αντιστοιχούν σε υψηλό βαθμό της συνέχειας της περίθαλψης με την έννοια της συγκέντρωσης. Ο δείκτης *NOP* έχει αποτελέσει τη βάση για άλλους πιο πολύπλοκους δείκτες της συνέχειας της περίθαλψης, όπως ο δείκτης *K* που παρουσιάσαμε στο Κεφάλαιο 2. Σημειώνουμε ότι για ένα σταθερό πλήθος διαθέσιμων παροχών, ο *NOP* δεν είναι ευαίσθητος στις αλλαγές που μπορεί να γίνουν στην κατανομή των επισκέψεων, αλλά μετράει τη διασπορά. Επομένως, αυτό το εξαιρετικά απλό μέτρο χρησιμοποιείται συχνά σε συνδυασμό με άλλα μέτρα συνέχειας για να αντισταθμίσει αυτή την αδυναμία του.

Η μαθηματική περιγραφή του *NOP* παρουσιάζεται στην Ενότητα 5.2 μαζί με τα αποτελέσματα για το απλό τυχαίο μοντέλο. Το κύριο αποτέλεσμα για την πιο γενική περίπτωση της μαρκοβιανής εξάρτησης, που βασίζεται στην προσέγγιση της εμφύτευσης της πεπερασμένης μαρκοβιανής αλυσίδας των Fu and Koutras (1994), παρουσιάζεται στην Ενότητα 5.3, μαζί με κάποια αριθμητικά παραδείγματα στην Ενότητα 5.4. Στην τελευταία ενότητα δίνονται και ορισμένες παρατηρήσεις για τη μέθοδο εύρεσης της ακριβούς κατανομής του δείκτη.

5.2 Προκαταρκτικές έννοιες

Έστω m το πλήθος των διαθέσιμων παροχών (π.χ. γιατροί που εργάζονται στην ίδια κλινική, συνήθως $2 \leq m \leq 8$), και έστω n ο συνολικός αριθμός επισκέψεων που

πραγματοποιήθηκαν από έναν ασθενή κατά τη διάρκεια της χρονικής περιόδου $[0, T]$. Η ακολουθία των παροχών που χρησιμοποιήθηκαν από τον ασθενή κατά τη χρονική αυτή περίοδο συμβολίζεται με $\{Z_t\}_{t=1}^n$ και μπορεί να θεωρηθεί ως μία ακολουθία τυχαίων μεταβλητών που ορίζονται από δοκιμές πολλών καταστάσεων με έναν πεπερασμένο χώρο καταστάσεων $S = \{1, \mathbf{K}, m\}$. Τότε το μέτρο της συνέχειας της περίθαλψης NOP μπορεί να οριστεί ως

$$X_n = \sum_{t=1}^n I(Z_1, \mathbf{K}, Z_t) \quad (5.1)$$

όπου οι δείκτριες συναρτήσεις ορίζονται ως εξής για $t = 2, \mathbf{K}, n$,

$$I(Z_1, \mathbf{K}, Z_t) = \begin{cases} 1 & \text{αν } Z_t \neq Z_k, \text{ για όλα τα } 1 \leq k < t \\ 0 & \text{διαφορετικά,} \end{cases}$$

και $I(Z_1) \equiv 1$.

Για την περίπτωση του απλού συνεχούς μοντέλου, όπου οι Z_t είναι ανεξάρτητες και ισόνομες, θα έχουμε:

$$P(Z_t = j) = 1/m, \text{ για όλα τα } j = 1, \mathbf{K}, m, \text{ και } t = 1, 2, \mathbf{K}, n.$$

Εύκολα διαπιστώνουμε ότι η ακολουθία $\{X_t\}_{t=1}^n$ που ορίστηκε στην εξίσωση (5.1) υπό το τυχαίο μοντέλο, σχηματίζει μια ομογενή μαρκοβιανή αλυσίδα με πίνακα πιθανοτήτων μεταπηδήσεως

$$M = \begin{matrix} & \begin{matrix} 1 \\ 2 \\ \mathbf{M} \\ k \\ \mathbf{M} \\ m \end{matrix} & \begin{bmatrix} D_1 & H_1 & 0 & \mathbf{K} & \mathbf{K} & 0 \\ 0 & D_2 & H_2 & & \mathbf{K} & 0 \\ & \mathbf{O} & \mathbf{O} & \mathbf{O} & & \\ & & \mathbf{O} & D_k & H_k & \\ & 0 & 0 & & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ 0 & 0 & \mathbf{K} & & 0 & D_m \end{bmatrix} & \end{matrix},$$

όπου $D_k = k/m$ και $H_k = (m-k)/m$, $k = 1, \dots, m$. Επομένως η πιθανότητα του $X_n = k$ μπορεί να βρεθεί από την εξίσωση

$$P(X_n = k) = \xi_1 M^{n-1} U'_k, \quad k = 1, \mathbf{K}, m, \quad (5.2)$$

όπου $\xi_1 = (1, 0, \mathbf{K}, 0)_{1 \times m}$ είναι η αρχική πιθανότητα και $U'_k = (0, \mathbf{K}, 0, 1, 0, \mathbf{K}, 0)'$ είναι ένα $m \times 1$ μοναδιαίο διάνυσμα με μονάδα στην k συντεταγμένη και μηδενικά οπουδήποτε αλλού. Η κατανομή του NOP ικανοποιεί επίσης από την παρακάτω αναδρομική σχέση:

$$P(X_n = k) = H_{k-1}P(X_{n-1} = k-1) + D_k P(X_{n-1} = k). \quad (5.3)$$

Μπορεί να δειχθεί από την εξίσωση (5.3), ύστερα από μερικές αλγεβρικές πράξεις, ότι η αναμενόμενη τιμή και η διακύμανση του X_n ικανοποιούν τις παρακάτω αναδρομικές σχέσεις:

$$E(X_n) = 1 + \left(1 - \frac{1}{m}\right)E(X_{n-1}), \quad (5.4)$$

$$\text{Var}(X_n) = \left(1 - \frac{2}{m}\right)\text{Var}(X_{n-1}) + \frac{1}{m}E(X_{n-1}) - \left(\frac{1}{m}E(X_{n-1})\right)^2. \quad (5.5)$$

Χρησιμοποιώντας τις αρχικές συνθήκες $E(X_1) = 1$ και $\text{Var}(X_1) = 0$, η αναμενόμενη τιμή και η διακύμανση του X_n βρίσκεται ίση με

$$E(X_n) = m \left[1 - \left(1 - \frac{1}{m}\right)^n\right], \quad (5.6)$$

$$\begin{aligned} \text{Var}(X_n) = & \left(1 - \frac{1}{m}\right)^{n-1} \left\{ (m-1) \left[1 - \left(1 - \frac{1}{m-1}\right)^{n-1}\right] \right\} \\ & - \left[\left(1 - \frac{1}{m}\right)^{n-1} \right]^2 \left\{ (m-1)^2 \left[1 - \left(1 - \frac{1}{(m-1)^2}\right)^{n-1}\right] \right\}. \end{aligned} \quad (5.7)$$

Έπεται από τις εξισώσεις (5.6) και (5.7) ότι $E(X_n) \rightarrow m$ και $\text{Var}(X_n) \rightarrow 0$ καθώς $n \rightarrow \infty$. Συνεπώς, με βάση την ανισότητα Markov προκύπτει άμεσα ότι $P(X_n = m) \rightarrow 1$. Ο ρυθμός συγκλίσεως στην τελευταία είναι αρκετά γρήγορος, όπως μπορούμε να δούμε στα αριθμητικά παραδείγματα που παρουσιάζονται στην Ενότητα 5.4.

Σημειώνουμε ότι για ανεξάρτητες και ισόνομες δοκιμές, ορισμένα παρόμοια αποτελέσματα (π.χ. αναδρομικές εξισώσεις για το $P(X_n = k)$) έχουν ήδη βρεθεί μέσω συνδυαστικών μεθόδων. Η χρήση της προσέγγισης με μαρκοβιανές αλυσίδες για τα παραπάνω αποτελέσματα απετέλεσε ένα προοίμιο και μια ειδική περίπτωση για τους παρακάτω γενικούς τύπους σχετικά με τις δοκιμές που εμφανίζουν μαρκοβιανή εξάρτηση.

5.3 Η ακριβής κατανομή

Ας θεωρήσουμε στο εξής ότι η ακολουθία $\{Z_t\}_{t=1}^n$ ορίζεται από μια ακολουθία μη υποβιβάσιμων και μη περιοδικών μη ομογενών δοκιμών m καταστάσεων με μαρκοβιανή εξάρτηση και πίνακες πιθανοτήτων μεταπηδήσεως

$$P_z = [p_{ij}(t)]_{m \times m} \quad t = 2, \mathbf{K}, n, \quad (5.8)$$

και αρχικές πιθανότητες $\pi_1 = (p_1, \mathbf{K}, p_m)$.

Μπορεί να δειχθεί ότι η ακολουθία $\{X_t\}$ που ορίστηκε στην εξίσωση (5.1) για το στατιστικό NOP , παρόλο που παράγεται από τη μαρκοβιανή αλυσίδα $\{Z_t\}$, δεν είναι απαραίτητα μαρκοβιανή αλυσίδα. Επομένως, οι συνήθεις τεχνικές για μαρκοβιανές αλυσίδες δεν μπορούν να εφαρμοστούν άμεσα. Ωστόσο, όπως φαίνεται παρακάτω, ο NOP είναι μια τ.μ. που μπορεί να εμφυτευτεί (imbeddable) σε μια πεπερασμένη μαρκοβιανή αλυσίδα σύμφωνα με την τεχνική των Fu and Koutras (1994). Ο παρακάτω ορισμός είναι μια παραλλαγή του Ορισμού 3.1:

Ορισμός 5.1 Μια πεπερασμένη μη αρνητική ακέραια τυχαία μεταβλητή X_n είναι μια πεπερασμένη μαρκοβιανή αλυσίδα που μπορεί να εμφυτευτεί αν

(α) υπάρχει μια πεπερασμένη μαρκοβιανή αλυσίδα $\{Y_t : t = 1, \mathbf{K}, n\}$ που ορίζεται σε ένα πεπερασμένο χώρο καταστάσεων Ω και έχει πίνακες πιθανοτήτων μεταπηδήσεως $M_t, t = 2, \mathbf{K}, n$, και αρχικές πιθανότητες ξ_1 ,

(β) υπάρχει μια διαμέριση $\{C_x : x = 0, 1, \mathbf{K}, l\}$ πάνω στον χώρο καταστάσεων Ω , και

(γ) για κάθε $x = 0, 1, \mathbf{K}, l$,

$$P(X_n = x) = P(Y_n \in C_x | \xi_1).$$

Για ένα δεδομένο σύνολο m ακεραίων $S = \{1, 2, \mathbf{K}, m\}$ ορίζουμε ως

$$\Phi_m = \{[1], [2], \mathbf{K}, [m], [1, 2], \mathbf{K}, [m-1, m], \mathbf{K}, [1, 2, 3], \mathbf{K}, [m-2, m-1, m], \mathbf{K}, [1, 2, \mathbf{K}, m]\}$$

το σύνολο όλων των πιθανών υποσυνόλων των ακεραίων του S , εκτός του κενού συνόλου. Ορίζουμε επίσης τον χώρο καταστάσεων

$$\Omega(m) = \{(\omega, i) : \omega \in \Phi_m, \text{ και } i \in \omega\}. \quad (5.9)$$

Για παράδειγμα, όταν $m = 3$, ο χώρος καταστάσεων είναι

$$\Omega(3) = \{([1],1), ([2],2), ([3],3), ([1,2],1), ([1,2],2), ([1,3],1), ([1,3],3), ([2,3],2), ([2,3],3), ([1,2,3],1), ([1,2,3],2), ([1,2,3],3)\}.$$

Γενικά, το μέγεθος του συνόλου Φ_m καθώς και του χώρου καταστάσεων $\Omega(m)$ είναι ίσα με

$$\text{Card}(\Phi_m) = \sum_{i=1}^m \binom{m}{i} = 2^m - 1$$

και

$$\text{Card}(\Omega(m)) = \sum_{i=1}^m i \binom{m}{i} = m2^{m-1}$$

αντίστοιχα.

Για να δείξουμε ότι ο *NOP* μπορεί να εμφυτευτεί σε μαρκοβιανή αλυσίδα, κατασκευάζουμε αρχικά μια πεπερασμένη μαρκοβιανή αλυσίδα $\{Y_t\}$ που να ικανοποιεί τις 3 συνθήκες του Ορισμού 5.1 και ορίζουμε αντίστοιχα τους πίνακες πιθανοτήτων μεταπηδήσεως και τις διαμερίσεις της.

Έστω $\{Y_t\}$ μία μαρκοβιανή αλυσίδα που ορίζεται επί του χώρου καταστάσεων $\Omega(m)$ ως εξής:

$$Y_t = (W_t, Z_t), \quad (5.10)$$

όπου Z_t είναι το αποτέλεσμα της t δοκιμής, και τα W_t είναι διακριτοί ακέραιοι και προκύπτουν αν διατάξουμε σε αύξουσα σειρά τα αποτελέσματα των δοκιμών $\{Z_1, Z_2, \mathbf{K}, Z_t\}$.

Για παράδειγμα, δοθέντος $t = 5$ και των αποτελεσμάτων $Z_1 = 3, Z_2 = 7, Z_3 = 3, Z_4 = 1,$ και $Z_5 = 1$, έχουμε $W_5 = [1,3,7]$ και $Y_5 = ([1,3,7],1)$.

Οι πίνακες πιθανοτήτων μεταπηδήσεως για την παραπάνω εμφυτευμένη μαρκοβιανή αλυσίδα $\{Y_t\}$ είναι ίσοι με:

$$\mathbf{M}_t = (\beta_{u,v}(t)), \quad (5.11)$$

όπου κάθε ζευγάρι $u = (\omega, i)$ και $v = (\omega^*, j)$ είναι από τον χώρο καταστάσεων $\Omega(m)$ και οι πιθανότητες μεταπηδήσεως $\beta_{u,v}(t)$, για $t = 2, \mathbf{K}, n$, καθορίζονται από την έκφραση

$$\beta_{u,v}(t) = P(Y_t = u | Y_{t-1} = v) = \begin{cases} p_{ij}(t) & \text{αν } \omega = \omega^* \quad \text{ή } \omega^* = \omega \cup j \\ 0 & \text{διαφορετικά.} \end{cases} \quad (5.12)$$

Αφού η συμπεριφορά μιας μαρκοβιανής αλυσίδας χαρακτηρίζεται αποκλειστικά από τους πίνακες πιθανοτήτων μεταπηδήσεώς της, η εξίσωση (5.12) είναι το κλειδί για την επιτυχία

της προσέγγισης αυτής. Για να κάνουμε την εξίσωση αυτή πιο κατανοητή, ας υποθέσουμε ότι το Y_5 είναι στην κατάσταση $u = ([1,3,7],1)$. Αν το αποτέλεσμα της 6^{ης} δοκιμής δεν είναι νέο, δηλαδή $Z_6 \in [1,3,7]$ ή πιο απλά $\omega = \omega^*$, τότε, με πιθανότητα $p_{13}(6)$, το Y_6 θα είναι στην κατάσταση $v = ([1,3,7],3)$. Παρομοίως, πηγαίνουμε από την κατάσταση u στην κατάσταση $v = ([1,3,7],1)$ με πιθανότητα $p_{11}(6)$ και στην $v = ([1,3,7],7)$ με πιθανότητα $p_{17}(6)$. Αν το 6^ο αποτέλεσμα είναι νέο, ας πούμε $Z_6 = 2$, τότε $\omega^* = [1,2,3,7] = \omega \cup 2$. Επομένως, πηγαίνουμε από την κατάσταση u στην $v = ([1,2,3,7],2)$ με πιθανότητα $p_{12}(6)$. Αφού ο αριθμός των ακεραίων στο W_t μπορεί είτε να είναι ο ίδιος είτε ένας παραπάνω σε σχέση με το W_{t-1} , από την κατάσταση $u = (\omega, i)$ μπορούμε να μεταβούμε μόνο σε καταστάσεις με $\omega = \omega^*$ ή $\omega^* = \omega \cup j$, ενώ έχουμε πιθανότητα μηδέν να μεταβούμε σε όλες τις υπόλοιπες καταστάσεις.

Μια διαμέριση επί του χώρου καταστάσεων $\Omega(m)$ μπορεί να οριστεί θεωρώντας τα σύνολα

$$C_x = \{(\omega, i) : (\omega, i) \in \Omega(m), \text{ και } \text{Card}(\omega) = x\} \quad (5.13)$$

Δοθείσης της αρχικής κατανομής πιθανότητας του Y_1

$$\xi_1 = (p_1, \mathbf{K}, p_m, 0, \mathbf{K}, 0)_{1 \times m 2^{m-1}} = (\boldsymbol{\pi}_1, \mathbf{0}, \mathbf{K}, \mathbf{0}),$$

εύκολα διαπιστώνουμε ότι

$$P(X_n = x) = P(Y_n \in C_x | \xi_1). \quad (5.14)$$

Σημειωτέον ότι αν $n = 1$, τότε $P(X_n = 1) \equiv 1$. Στη συνέχεια θα υποθέτουμε ότι $n \geq 2$, αφού η ειδική περίπτωση $n = 1$ δεν παρουσιάζει πλέον ενδιαφέρον για να μελετηθεί.

Οι παραπάνω ορισμοί οδηγούν στο ακόλουθο θεώρημα, το οποίο δόθηκε από την W.Y.W. Lou (2000):

Θεώρημα 5.1 *Ας θεωρήσουμε μια ακολουθία μη υποβιβάσιμων και μη περιοδικών μη ομογενών δοκιμών m καταστάσεων $\{Z_t\}$ μαρκοβιανά εξαρτημένων με πίνακες πιθανοτήτων μεταπηδήσεως P_z και αρχική πιθανότητα $\boldsymbol{\pi}_1$. Για δεδομένο n και ξ_1 , η κατανομή του NOP δίνεται από τον τύπο:*

$$P(X_n = x) = \xi_1 \left(\prod_{t=2}^n M_t \right) U'(C_x), \quad x = 1, 2, \mathbf{K}, m, \quad (5.15)$$

όπου $\xi_1 = (\pi_1, \mathbf{0}, \mathbf{K}, \mathbf{0})_{1 \times m 2^{m-1}}$ είναι το διάνυσμα αρχικών πιθανοτήτων, $U'(C_x) = (0, \mathbf{K}, 0, 1, \mathbf{K}, 1, 0, \mathbf{K}0)'$ είναι ένα $m 2^{m-1} \times 1$ διάνυσμα στήλη με μονάδες στις συντεταγμένες που σχετίζονται με το στοιχείο της διαμέρισης C_x που ορίστηκε στην εξίσωση (5.13) και 0 οπουδήποτε αλλού. Επιπλέον, οι πίνακες πιθανοτήτων μεταπηδήσεως M_t , $t = 2, \mathbf{K}, n$ έχουν τη μορφή:

$$M_t = \begin{bmatrix} A_1(t) & B_1(t) & & & & \mathbf{0} \\ & A_2(t) & B_2(t) & & & \\ & & \mathbf{O} & \mathbf{O} & & \\ & & & \mathbf{O} & B_{m-1}(t) & \\ \mathbf{0} & & & & & A_m(t) \end{bmatrix}, \quad (5.16)$$

όπου το $A_i(t)$ και $B_i(t)$ ορίζονται στην (5.12).

Απόδειξη. Από τις δομές και τους ορισμούς των εξισώσεων (5.9)-(5.14) προκύπτει ότι η μαρκοβιανή αλυσίδα Y_t καθιστά το μέτρο NOP μια τ.μ. εμφυτεύσιμη σε πεπερασμένη μαρκοβιανή αλυσίδα. Τις πιθανότητες μεταπηδήσεως που ορίζονται στην εξίσωση (5.13) για ένα δεδομένο x , μπορούμε να τις δούμε ως εξής: Από κάθε κατάσταση $u = (\omega, i) \in C_x$ μεταβαίνουμε, με πιθανότητα $p_{ij}(t)$, είτε στην κατάσταση $v = (\omega, j) \in C_x$ είτε στην κατάσταση $v = (\omega^*, j) \in C_{x+1}$ ($\omega^* = \omega \cup j$, και $Card(\omega^*) \equiv Card(\omega) + 1$) μόνο. Επομένως, οι πίνακες πιθανοτήτων μεταπηδήσεως M_t έχουν τη μορφή που δίνεται στην εξίσωση (5.16). Η εξίσωση (5.15) είναι άμεση συνέπεια των αναδρομικών τύπων των Chapman-Kolmogorov.

5.4 Παράδειγμα

Για την καλύτερη κατανόηση της χρήσης των παραπάνω θεωρητικών αποτελεσμάτων για την εύρεση της κατανομής του NOP , παρατίθεται παρακάτω ένα λεπτομερές αριθμητικό παράδειγμα

Παράδειγμα. Έστω $\{Z_t\}_{t=1}^n$ μια μη ομογενής μαρκοβιανή αλυσίδα τριών καταστάσεων με πίνακες πιθανοτήτων μεταπηδήσεως

$$P_z(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & p_{13}(t) \\ p_{21}(t) & p_{22}(t) & p_{23}(t) \\ p_{31}(t) & p_{32}(t) & p_{33}(t) \end{bmatrix}, \quad t = 2, \mathbf{K}, n$$

και διάνυσμα αρχικής πιθανότητας $\pi_1 = (1/3, 1/3, 1/3)$.

Σύμφωνα με την εξίσωση (5.9), ο χώρος καταστάσεων ορίζεται ως

$$\Omega(3) = \{([1],1), ([2],2), ([3],3), ([1,2],1), ([1,2],2), ([1,3],1), ([1,3],3), ([2,3],2), ([2,3],3), ([1,2,3],1), ([1,2,3],2), ([1,2,3],3)\}$$

και από τους τύπους (5.13) προκύπτει ότι η διαμέριση $\{C_x\}$ επί του χώρου καταστάσεων $\Omega(3)$

είναι

$$C_1 = \{([1],1), ([2],2), ([3],3)\},$$

$$C_2 = \{([1,2],1), ([1,2],2), ([1,3],1), ([1,3],3), ([2,3],2), ([2,3],3)\},$$

$$C_3 = \{([1,2,3],1), ([1,2,3],2), ([1,2,3],3)\}.$$

Οι πίνακες πιθανοτήτων μεταπηδήσεως για την εμφυτευμένη μαρκοβιανή αλυσίδα μπορούν να εκφραστούν μέσω της εξίσωσης (5.12) ως εξής:

$$M_t = \begin{bmatrix} p_{11}(t) & 0 & 0 & 0 & p_{12}(t) & 0 & p_{13}(t) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{22}(t) & 0 & p_{21}(t) & 0 & 0 & 0 & 0 & p_{23}(t) & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{33}(t) & 0 & 0 & p_{31}(t) & 0 & p_{32}(t) & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & p_{11}(t) & p_{12}(t) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{13}(t) \\ 0 & 0 & 0 & p_{21}(t) & p_{22}(t) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{23}(t) \\ 0 & 0 & 0 & 0 & 0 & p_{11}(t) & p_{13}(t) & 0 & 0 & 0 & p_{12}(t) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{31}(t) & p_{33}(t) & 0 & 0 & 0 & p_{32}(t) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{22}(t) & p_{23}(t) & p_{21}(t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{32}(t) & p_{33}(t) & p_{31}(t) & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{11}(t) & p_{12}(t) & p_{13}(t) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{21}(t) & p_{22}(t) & p_{23}(t) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{31}(t) & p_{32}(t) & p_{33}(t) & 0 \end{bmatrix}$$

για $t = 2, \mathbf{K}, n$. Οι καταστάσεις που σχετίζονται με το M_t , έχουν ταξινομηθεί με την ίδια σειρά όπως εκείνες στο $\Omega(3)$, δηλαδή με τη λεξικογραφική διάταξη.

Εστω ότι το διάνυσμα αρχικής πιθανότητας ξ_1 ισούται με

$$\xi_1 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \right).$$

Με την έννοια της συνέχειας, αυτό σημαίνει ότι ο κάθε γιατρός έχει την ίδια πιθανότητα να χρησιμοποιηθεί στην πρώτη επίσκεψη του ασθενούς στην κλινική. Οι κατανομές του NOP , κάτω από την υπόθεση της ανεξαρτησίας και της ισονομίας, καθώς και κάτω από την υπόθεση της ομογενούς μαρκοβιανής εξάρτησης για το $\{Z_t\}$ συνοψίζονται στον Πίνακα 5.1:

- (i) (IIDEQ): Οι $\{Z_t\}$ είναι ανεξάρτητες και ισόνομες με ίσες πιθανότητες

$$P_z = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \quad t = 2, \mathbf{K}, n$$

- (ii) (IIDUE): Οι $\{Z_t\}$ είναι ανεξάρτητες και ισόνομες με άνισες πιθανότητες

$$P_z = \begin{bmatrix} \frac{8}{10} & \frac{1}{10} & \frac{1}{10} \\ \frac{8}{10} & \frac{1}{10} & \frac{1}{10} \\ \frac{8}{10} & \frac{1}{10} & \frac{1}{10} \end{bmatrix}, \quad t = 2, \mathbf{K}, n$$

- (iii) (MARKOV): Οι $\{Z_t\}$ είναι μαρκοβιανά εξαρτημένες με πίνακα πιθανοτήτων μεταπηδήσεως

$$P_z = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}, \quad t = 2, \mathbf{K}, n$$

Βλέπουμε στον Πίνακα 5.1 ότι καθώς αυξάνεται το n , η πιθανότητα στο $x = 1$ μειώνεται, η πιθανότητα στο $x = 3$ προσεγγίζει τη μονάδα και η αναμενόμενη τιμή $E(X_n)$ συγκλίνει γρήγορα στο 3. Ο ρυθμός σύγκλισης εξαρτάται ελαφρά από τον πίνακα πιθανοτήτων μεταπηδήσεως του $\{Z_t\}_{t=1}^n$.

Μοντέλο	n	Συνάρτηση Πιθανότητας			Μέσος
		x = 1	x = 2	x = 3	
IIDEQ	3	0.111111	0.666667	0.222222	2.111111
	4	0.037037	0.518519	0.444444	2.407407
	5	0.012346	0.370370	0.617284	2.604938
	7	0.001372	0.172840	0.825789	2.824417
	10	5.08e-05	0.051923	0.948026	2.947975
	20	8.60e-10	0.000902	0.999098	2.999098
	30	1.46e-14	1.56e-05	0.999984	2.999984
	50	4.18e-24	4.70e-09	1	3
IIDUE	3	0.220000	0.666667	0.113333	1.893333
	4	0.171333	0.634667	0.194000	2.022667
	5	0.136600	0.602667	0.260733	2.124133
	7	0.087382	0.533867	0.378751	2.291369
	10	0.044739	0.427083	0.528178	2.483439
	20	0.004804	0.170506	0.824690	2.819886
	30	0.000516	0.061770	0.937714	2.937198
	50	5.95e-06	0.007623	0.992371	2.992365
MARKOV	3	0.250000	0.625000	0.125000	1.875000
	4	0.125000	0.593750	0.281250	2.156250
	5	0.062500	0.507813	0.429688	2.367188
	7	0.015625	0.324707	0.659668	2.644043
	10	0.001953	0.146263	0.851784	2.849831
	20	1.91e-06	0.008453	0.991545	2.991543
	30	1.86e-09	0.000473	0.999524	2.999524
	50	1.78e-15	1.51e-06	0.999999	2.999998

Πίνακας 5.1. Οι αναμενόμενες τιμές και οι ακριβείς πιθανότητες του NOP δοθέντος $m = 3$ και $\pi_1 = (1/3, 1/3, 1/3)$ κάτω από την υπόθεση *i.i.d.* με ίσες και άνισες πιθανότητες καθώς και κάτω από την υπόθεση της μαρκοβιανής εξάρτησης για ορισμένα επιλεγμένα n .

5.5 Παρατηρήσεις

Η ακριβής κατανομή του NOP βρέθηκε με χρήση της μεθόδου εμφύτευσης σε πεπερασμένη μαρκοβιανή αλυσίδα. Ένα σημαντικό πλεονέκτημα της διατύπωσης αυτής είναι η πολλαπλή της χρησιμότητα, δηλαδή ο ίδιος τύπος μπορεί να χρησιμοποιηθεί για τη μελέτη της ακριβούς κατανομής όταν οι δοκιμές πολλών καταστάσεων είναι είτε *i.i.d.* είτε εξαρτημένες σύμφωνα με μια ομογενή/μη ομογενή μαρκοβιανή αλυσίδα.

Τα αριθμητικά αποτελέσματα που παρουσιάστηκαν στην Ενότητα 5.4 υπολογίστηκαν με τη χρήση της γλώσσας προγραμματισμού S-plus. Όταν το n είναι μεγάλο, ο χρόνος υπολογισμού δεν αποτελεί γενικά πρόβλημα, αφού η εκτίμηση της κατανομής του NOP απαιτεί μόνο πολλαπλασιασμούς των πινάκων πιθανοτήτων μεταπηδήσεως M_i . Είναι προφανές από το Θεώρημα 5.1 ότι η δομή των πινάκων πιθανοτήτων μεταπηδήσεως για την εμφυτευμένη μαρκοβιανή αλυσίδα εξαρτάται μόνο από το πλήθος των διαθέσιμων παροχών m (συνήθως το m είναι μικρότερο του 10) και από τους τετραγωνικούς διαγώνιους πίνακες $A_i(t)$ και $B_i(t)$ όπως αυτοί ορίστηκαν στην εξίσωση (5.16).

Η αναμενόμενη τιμή του NOP τείνει στο m καθώς το πλήθος των επισκέψεων n γίνεται πολύ μεγάλο, εφόσον η μαρκοβιανή αλυσίδα $\{Z_t\}$ είναι μη υποβιβάσιμη και μη περιοδική. Για πιο γενικές περιπτώσεις, η αναμενόμενη τιμή και οι υπόλοιπες ροπές του στατιστικού NOP μπορούν εύκολα να υπολογιστούν αντικαθιστώντας το $U'(C_x)$ με το $x^r U'(C_x)$ στο Θεώρημα 5.1 για την αντίστοιχη ροπή r τάξεως.

Κατά τη διάρκεια της ανάλυσης μας, το πλήθος των επισκέψεων n θεωρήθηκε δεδομένο. Σε ορισμένες πραγματικές καταστάσεις, ωστόσο, το πλήθος των επισκέψεων δεν μπορεί να καθοριστεί από πριν, και επομένως το n πρέπει να θεωρηθεί τυχαία μεταβλητή, άρα χρειαζόμαστε την από κοινού κατανομή των n και NOP .

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κ Ε Φ Α Λ Α Ι Ο 6

Case study σε πραγματικά δεδομένα

6.1 Εισαγωγή

Προκειμένου να μελετηθούν περαιτέρω οι δείκτες συνέχειας της περίθαλψης και να διερευνηθεί η εφαρμογή τους σε πραγματικές καταστάσεις, συγκεντρώθηκαν στοιχεία από πραγματικούς ασθενείς. Τα στοιχεία συγκεντρώθηκαν από υποκατάστημα του Ιδρύματος Κοινωνικών Ασφαλίσεων (ΙΚΑ) επαρχιακής πόλης της Ελλάδας και αναφέρονται σε επισκέψεις ασθενών σε 5 ιατρούς παθολόγους του Ιδρύματος από τον Ιανουάριο έως και το Σεπτέμβριο του 2004. Η συλλογή των στοιχείων έγινε από τα ημερήσια φύλλα των ραντεβού των 5 γιατρών, αλλά και από μη προγραμματισμένες επισκέψεις, δηλ. επισκέψεις εμβόλιμες χωρίς να έχει προηγηθεί ραντεβού μέσω της τηλεφωνικής υπηρεσίας των ραντεβού (184).

Αρχικά, τα δεδομένα καταχωρήθηκαν σε ένα αρχείο το ένα κάτω από το άλλο στην εξής μορφή:

ΕΠΩΝΥΜΟ	ΟΝΟΜΑ	ΗΜΕΡΟΜΗΝΙΑ	ΓΙΑΤΡΟΣ
Αντωνίου	Στέλλα	17/3/2004	B

κοκ.

Το πλήθος των δεδομένων αυτών ήταν 9658. Στο πλήθος αυτό συμπεριλαμβάνονται βέβαια ασθενείς που είχαν επισκεφθεί παθολόγο του ΙΚΑ παραπάνω από μία φορά, δεν επρόκειτο δηλαδή για 9658 ασθενείς.

Στη συνέχεια, και αφού τοποθετήθηκαν τα δεδομένα σε αλφαβητική σειρά, εξετάστηκε ποιοι ασθενείς είχαν παραπάνω από 4 επισκέψεις σε έναν ή περισσότερους γιατρούς, έτσι ώστε να είναι δυνατός ο υπολογισμός των δεικτών συνέχειας της περίθαλψης.

Από την ταξινόμηση αυτή καταλήξαμε σε 472 περιπτώσεις ασθενών με πλήθος επισκέψεων 4, 5, 6 και 7. Υπήρξαν και ασθενείς με 9 και 10 επισκέψεις κατά την περίοδο αναφοράς, επειδή όμως δεν ήταν πολλοί, το τελικό μέγεθος του δείγματος ήταν 472.

Η διάρθρωση των δεδομένων έχει ως εξής:

Πλήθος επισκέψεων	Πλήθος ασθενών
4	250
5	137
6	52
7	33

Στη συνέχεια θα επιχειρήσουμε αρχικά μια περιγραφική ανάλυση των δεδομένων και στη συνέχεια μια ανάλυση με βάση τις ακριβείς κατανομές των δεικτών της συνέχειας της περίθαλψης.

6.2 Περιγραφική ανάλυση

Στην ενότητα αυτή θα επιχειρήσουμε μια περιγραφική ανάλυση των δεδομένων που συγκεντρώθηκαν από τους 472 ασθενείς του Ιδρύματος Κοινωνικών Ασφαλίσεων. Αφού υπολογίσουμε τους δείκτες *UPC*, *SECON* και *NOP* για όλους τους ασθενείς, στη συνέχεια θα ταξινομήσουμε τα αποτελέσματα κατά πλήθος επισκέψεων ($n = 4, 5, 6$ και 7) και κατά δείκτη και θα προσπαθήσουμε να ανιχνεύσουμε αν υπάρχει συνάφεια μεταξύ των αποτελεσμάτων των τριών δεικτών συνέχειας. Επίσης, θα επιχειρήσουμε να ερμηνεύσουμε τα αποτελέσματα λαμβάνοντας υπόψη τα ελληνικά δεδομένα στο χώρο της υγείας, καθώς και τη σχέση των Ελλήνων ασθενών με τον παθολόγο τον οποίο επισκέπτονται.

Ας ξεκινήσουμε με τους υπολογισμούς που αφορούν τους ασθενείς με πλήθος επισκέψεων $n = 4$. Επειδή οι δείκτες *UPC* και *SECON* έχουν το ίδιο πεδίο τιμών, θα παρουσιάσουμε τα αποτελέσματα για τους δύο αυτούς δείκτες στον ίδιο πίνακα και θα ακολουθήσει ένας πίνακας για το δείκτη *NOP*. Για το δείκτη *UPC* η ερμηνεία του πίνακα είναι απλή. Στην πρώτη στήλη (Τιμή *UPC* και *SECON*) γράφονται οι πιθανές τιμές του δείκτη, στη στήλη «Απόλυτες συχνότητες ασθενών για δείκτη *UPC*» γράφεται για πόσους ασθενείς η ακολουθία των επισκέψεων τους μας δίνει τη συγκεκριμένη τιμή του δείκτη και τέλος, στη στήλη «Σχετικές συχνότητες ασθενών για το δείκτη *UPC*» γράφονται οι τιμές της προηγούμενης

στήλης εκπεφρασμένες σε ποσοστό επί του συνόλου των ασθενών. Οι ίδιες διευκρινήσεις ισχύουν και για τους πίνακες με τα αποτελέσματα για τους δείκτες *SECON* και *NOP*:

Τιμή <i>UPC</i> και <i>SECON</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>UPC</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>UPC</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>SECON</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>SECON</i>
0	0	0.00	3	1.20
1/4	0	0.00	15	6.00
2/4	14	5.60	31	12.40
3/4	35	14.00	201	80.40
1	201	80.40	0	0.00
ΣΥΝΟΛΟ	250	100.00	250	100.00

Τιμή <i>NOP</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>NOP</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>NOP</i>
1/5	201	80.40
2/5	42	16.80
3/5	7	2.80
4/5	0	0.00
1	0	0.00
ΣΥΝΟΛΟ	250	100.00

Από τους παραπάνω πίνακες βλέπουμε ότι οι 3 δείκτες συνέχειας μας δίνουν τα ίδια αποτελέσματα μόνο όταν σε όλες τις επισκέψεις έχει χρησιμοποιηθεί ο ίδιος παροχέας. Όταν η τιμή του *UPC*, του *SECON* και του *NOP* είναι 1, 3/4 και 1/5 αντίστοιχα, δηλαδή όταν και στις 4 επισκέψεις δεν έχουμε διαφοροποίηση όσον αφορά τον παροχέα, το αντίστοιχο ποσοστό των ασθενών είναι και στις τρεις περιπτώσεις το ίδιο, δηλαδή 80.40%. Σε καμία άλλη περίπτωση δεν έχουμε ταύτιση των ποσοστών, γεγονός αναμενόμενο εφόσον μπορεί μεν οι δείκτες να αποτελούν μέτρα της συνέχειας της περίθαλψης, όμως ο καθένας μας δίνει μια διαφορετική πληροφορία.

Εκτός από τη συνάφεια μεταξύ των τριών δεικτών, παρατηρούμε ότι η συντριπτική πλειοψηφία (80.40%) των ασθενών και τις 4 φορές επισκέφθηκε τον ίδιο γιατρό. Σε γενικές γραμμές είναι λογικό και αναμενόμενο, αφού ο παθολόγος είναι ο γιατρός τον οποίο θεωρούμε ως τον οικογενειακό γιατρό και γνωρίζει το ιστορικό μας. Επομένως, μπορούμε να πούμε ότι η τέλεια αυτή συνέχεια οφείλεται κυρίως σε υποκειμενικούς και συναισθηματικούς λόγους. Οι περισσότεροι άνθρωποι όταν χρειαστούν παθολόγο δεν αναζητούν τον οποιοδήποτε από τους διαθέσιμους παθολόγους, πηγαίνουν κατευθείαν σε αυτόν που έχουν συνηθίσει να πηγαίνουν.

Όταν το πλήθος των επισκέψεων είναι ίσο με $n = 5$, οι αντίστοιχοι πίνακες έχουν ως εξής:

Τιμή <i>UPC</i> και <i>SECON</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>UPC</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>UPC</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>SECON</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>SECON</i>
0	0	0.00	1	0.73
1/5	0	0.00	3	2.19
2/5	1	0.73	19	13.87
3/5	15	10.95	15	10.95
4/5	22	16.06	99	72.26
1	99	72.26	0	0.00
ΣΥΝΟΛΟ	137	100.00	137	100.00

Τιμή <i>NOP</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>NOP</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>NOP</i>
1/5	99	72.26
2/5	34	24.82
3/5	4	2.92
4/5	0	0.00
1	0	0.00
ΣΥΝΟΛΟ	137	100.00

Τέλος, για τους 52 ασθενείς που επισκέφθηκαν 6 φορές το ΙΚΑ πήραμε τους επόμενους δύο πίνακες:

Τιμή <i>UPC</i> και <i>SECON</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>UPC</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>UPC</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>SECON</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>SECON</i>
0	0	0.00	1	1.92
1/6	0	0.00	2	3.85
2/6	1	1.92	1	1.92
3/6	3	5.77	7	13.46
4/6	5	9.62	6	11.54
5/6	8	15.38	35	67.31
1	35	67.31	0	0.00
ΣΥΝΟΛΟ	52	100.00	52	100.00

Τιμή <i>NOP</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>NOP</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>NOP</i>
1/5	35	67.31
2/5	14	26.92
3/5	1	1.92
4/5	2	3.85
1	0	0.00
ΣΥΝΟΛΟ	52	100.00

Βλέπουμε ότι, τόσο όταν $n = 5$, όσο και όταν $n = 6$, τα ποσοστά των ασθενών για τους τρεις δείκτες της συνέχειας της περίθαλψης είναι τα ίδια μόνο όταν σε όλες τις επισκέψεις έχει χρησιμοποιηθεί ο ίδιος παροχέας (72.26% και 67.31% αντίστοιχα). Αυτό συμβαίνει για τους ίδιους λόγους που αναφέρθηκαν και παραπάνω. Αυτό που παρατηρούμε αν συγκρίνουμε μεταξύ τους τα ποσοστά για τα οποία οι τρεις δείκτες, *UPC*, *SECON* και *NOP* συμπίπτουν, δηλαδή όταν $n = 4, 5$ και 6 , είναι ότι όσο μεγαλώνει το n , τόσο μικραίνει το ποσοστό εκείνων

των ασθενών που επισκέφθηκαν έναν και μόνο γιατρό (80.40% όταν $n = 4$, 72.26% όταν $n = 5$ και 67.31% όταν $n = 6$). Αυτό συμβαίνει επειδή όσο μεγαλύτερο είναι το πλήθος των επισκέψεων, τόσο μεγαλύτερη είναι και η πιθανότητα να χρησιμοποιηθούν και άλλοι γιατροί είτε διότι ο συγκεκριμένος γιατρός δεν είναι διαθέσιμος είτε διότι ο ασθενής θέλει να επισκεφθεί άλλο γιατρό.

Η διαπίστωση αυτή επιβεβαιώνεται και στην περίπτωση $n = 7$, όπου το αντίστοιχο ποσοστό είναι ακόμα μικρότερο. Τα αποτελέσματα για την περίπτωση αυτή φαίνονται στους επόμενους πίνακες:

Τιμή <i>UPC</i> και <i>SECON</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>UPC</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>UPC</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>SECON</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>SECON</i>
0	0	0.00	0	0.00
1/7	0	0.00	0	0.00
2/7	0	0.00	3	9.09
3/7	3	9.09	1	3.03
4/7	3	9.09	5	15.15
5/7	3	9.09	5	15.15
6/7	5	15.15	19	57.58
1	19	57.58	0	0.00
ΣΥΝΟΛΟ	33	100.00	33	100.00

Τιμή <i>NOP</i>	Απόλυτες Συχνότητες ασθενών για το δείκτη <i>NOP</i>	Σχετικές Συχνότητες (%) ασθενών για το δείκτη <i>NOP</i>
1/5	19	57.58
2/5	11	33.33
3/5	3	9.09
4/5	0	0.00
1	0	0.00
ΣΥΝΟΛΟ	33	100.00

6.3 Σύγκριση εμπειρικών τιμών με τις ακριβείς κατανομές

Στην ενότητα αυτή θα εξετάσουμε περιπτώσεις με διάφορους πίνακες πιθανοτήτων μεταπηδήσεως και διάφορες αρχικές κατανομές και θα συγκρίνουμε τα αποτελέσματα με τα εμπειρικά μας δεδομένα. Με τον τρόπο αυτό θα προσπαθήσουμε να δούμε ποιο μοντέλο προσαρμόζεται καλύτερα στα δεδομένα που συγκεντρώσαμε.

Ας θυμηθούμε τώρα από ποιους τύπους δίνονται οι ακριβείς κατανομές των τριών δεικτών συνέχειας, δηλαδή του *SECON*, του *UPC* και του *NOP*.

Για το δείκτη *SECON* είδαμε στο Κεφάλαιο 3 στο Θεώρημα 3.2 ότι, αν $p_{11} = \mathbf{K} = p_{mm}$, τότε η ακριβής κατανομή του δίνεται από τον παρακάτω τύπο:

$$P(S_n = s) = \binom{n}{s} \boldsymbol{\pi} \mathbf{B}^{n-s} \mathbf{D}^s \mathbf{1}'$$

όπου \mathbf{D} και \mathbf{B} είναι υποπίνακες του πίνακα πιθανοτήτων μεταπηδήσεως $\mathbf{A} = (p_{ij})_{m \times m}$ που ορίζονται ως εξής:

$$\mathbf{D}_{m \times m} = \begin{pmatrix} p_{11} & & 0 \\ & \mathbf{O} & \\ 0 & & p_{mm} \end{pmatrix} \text{ και } \mathbf{B}_{m \times m} = \mathbf{A}_{m \times m} - \mathbf{D}_{m \times m}$$

και $\mathbf{1} = (1, 1, \mathbf{K}, 1)$ είναι ένα $1 \times m$ διάνυσμα-γραμμή με μονάδες σε όλες τις συντεταγμένες.

Η δεσμευμένη κατανομή του δείκτη *UPC* για $0 \leq x \leq n$, όπως είδαμε στο Κεφάλαιο 4 δίνεται από τον τύπο:

$$P(N_n = x | v(T) = n) = \boldsymbol{\xi} \mathbf{M}^n \mathbf{U}'(C_x)$$

όπου $\boldsymbol{\xi} = (\boldsymbol{\pi}, \mathbf{0}, \mathbf{K}, \mathbf{0})_{1 \times d}$ είναι η αρχική κατανομή της μαρκοβιανής αλυσίδας $\{Y_t\}$ και το $\mathbf{U}'(C_x)$ δηλώνει τον ανάστροφο του διανύσματος γραμμής $\mathbf{U}(C_x) = (0, \mathbf{K}, 0, 1, \mathbf{K}, 1, 0, \mathbf{K}, 0)_{1 \times d}$ (το τελευταίο έχει μονάδες στις συντεταγμένες που σχετίζονται με καταστάσεις του συνόλου $C_x = \{(x, 1), \mathbf{K}, (x, m)\}$ και μηδενικά αλλού). Ο πίνακας πιθανοτήτων μεταπηδήσεως \mathbf{M} ορίζεται ως εξής:

$$\mathbf{M} = \begin{pmatrix} \mathbf{B} & \mathbf{C} & & & \mathbf{0} \\ & \mathbf{B} & \mathbf{C} & & \\ & \mathbf{0} & \mathbf{O} & \mathbf{O} & \\ & & \mathbf{O} & \mathbf{B} & \mathbf{C} \\ \mathbf{0} & & & \mathbf{0} & \mathbf{B}^* \end{pmatrix}_{d \times d}$$

όπου $d = (n+1) \times m$, οι υποπίνακες \mathbf{B} και \mathbf{C} ορίζονται ως

$$\mathbf{B} = \begin{pmatrix} 0 & p_{12} & \mathbf{L} & p_{1m} \\ 0 & p_{22} & \mathbf{L} & p_{2m} \\ \mathbf{L} & \mathbf{L} & \mathbf{L} & \mathbf{L} \\ 0 & p_{m2} & \mathbf{L} & p_{mm} \end{pmatrix}_{m \times m}, \quad \mathbf{C} = \begin{pmatrix} p_{11} & 0 & \mathbf{L} & 0 \\ p_{21} & 0 & \mathbf{L} & 0 \\ \mathbf{L} & \mathbf{L} & \mathbf{L} & \mathbf{L} \\ p_{m1} & 0 & \mathbf{L} & 0 \end{pmatrix}_{m \times m}$$

και $\mathbf{B}^* = I_{m \times m}$ είναι ο $m \times m$ μοναδιαίος πίνακας.

Τέλος, η εξίσωση που δίνει την ακριβή κατανομή του δείκτη NOP , όπως είδαμε στο Κεφάλαιο 5, είναι η εξής:

$$P(X_n = x) = \xi_1 \left(\prod_{t=2}^n M_t \right) U'(C_x), \quad x = 1, 2, \mathbf{K}, m,$$

όπου $\xi_1 = (\pi_1, \mathbf{0}, \mathbf{K}, \mathbf{0})_{1 \times m 2^{m-1}}$ είναι το διάνυσμα αρχικών πιθανοτήτων, $U'(C_x) = (0, \mathbf{K}, 0, 1, \mathbf{K}, 1, 0, \mathbf{K}, 0)'$ είναι ένα $m 2^{m-1} \times 1$ διάνυσμα-στήλη με μονάδες στις συντεταγμένες που σχετίζονται με καταστάσεις του συνόλου C_x που ορίστηκε στην εξίσωση (5.13) και 0 οπουδήποτε αλλού. Τέλος, οι πίνακες πιθανοτήτων μεταπηδήσεως M_t , $t = 2, \mathbf{K}, n$ έχουν τη μορφή:

$$M_t = \begin{bmatrix} A_1(t) & B_1(t) & & & \mathbf{0} \\ & A_2(t) & B_2(t) & & \\ & & \mathbf{O} & \mathbf{O} & \\ & & & \mathbf{O} & B_{m-1}(t) \\ \mathbf{0} & & & & A_m(t) \end{bmatrix},$$

όπου τα $A_i(t)$ και $B_i(t)$ ορίζονται στην εξίσωση (5.12) του ίδιου κεφαλαίου.

Θα εξετάσουμε περιπτώσεις με 6 διαφορετικούς πίνακες πιθανοτήτων μεταπηδήσεως οι οποίοι θα έχουν τη γενική μορφή:

$$\mathbf{A} = \begin{bmatrix} \frac{1}{k} & \frac{k-1}{4k} & \frac{k-1}{4k} & \frac{k-1}{4k} & \frac{k-1}{4k} \\ \frac{k-1}{4k} & \frac{1}{k-1} & \frac{k-1}{4k} & \frac{k-1}{4k} & \frac{k-1}{4k} \\ \frac{4k}{k-1} & \frac{k}{k-1} & \frac{1}{k-1} & \frac{4k}{k-1} & \frac{4k}{k-1} \\ \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{k}{k-1} & \frac{4k}{k-1} & \frac{4k}{k-1} \\ \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{k-1}{4k} & \frac{1}{k-1} & \frac{4k}{k-1} \\ \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{k}{k-1} & \frac{4k}{k-1} \\ \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{k-1}{4k} & \frac{1}{k-1} \\ \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{4k}{k-1} & \frac{k}{k-1} \end{bmatrix}, \quad k > 1.$$

Πιο συγκεκριμένα, θα εξετάσουμε τις περιπτώσεις πινάκων πιθανοτήτων μεταπηδήσεως με (A) $k = 5$, (B) $k = 2$, (Γ) $k = 3/2$, (Δ) $k = 4/3$, (E) $k = 10$ και (ΣΤ) $k = 20$. Στην περίπτωση (A), ο πίνακας γίνεται $A = (1/5)_{5 \times 5}$, το οποίο σημαίνει ότι ο ασθενής έχει τις ίδιες πιθανότητες να αλλάξει γιατρό από τη μία επίσκεψη στην άλλη με το να επισκεφθεί τον ίδιο γιατρό σε δύο συνεχόμενες επισκέψεις (η $\{X_t\}$ είναι ακολουθία ανεξάρτητων και ισόνομων τ.μ. που παίρνουν με την ίδια πιθανότητα όλες τις δυνατές τιμές). Στις περιπτώσεις (B), (Γ) και (Δ) αυξάνουμε σταδιακά την πιθανότητα παραμονής στον ίδιο γιατρό, μειώνοντας ταυτόχρονα την πιθανότητα αλλαγής γιατρού από τη μία επίσκεψη στην άλλη. Στις περιπτώσεις (E) και (ΣΤ) εξετάζουμε μοντέλα όπου η πιθανότητα αλλαγής γιατρού είναι μεγαλύτερη από την πιθανότητα παραμονής στον ίδιο γιατρό, με μεγαλύτερη πιθανότητα αλλαγής γιατρού στην περίπτωση (ΣΤ). Στις περιπτώσεις, (B), (Γ), (Δ), (E) και (ΣΤ) η $\{X_t\}$ είναι ομογενής μαρκοβιανή αλυσίδα με τους αντίστοιχους πίνακες πιθανοτήτων μεταπηδήσεως.

Ειδικά για το δείκτη *UPC* (δείκτης συνέχειας του συνήθους παροχέα), όπου παίζει ρόλο στην ακριβή κατανομή του η αρχική πιθανότητα, θα εξετάσουμε τις περιπτώσεις όπου η εργοδική κατανομή είναι (i) $\pi_1 = (1/5, 1/5, 1/5, 1/5, 1/5)$ και (ii) $\pi_2 = (1, 0, 0, 0, 0)$. Στην πρώτη περίπτωση ο κάθε παροχέας έχει τις ίδιες πιθανότητες να χρησιμοποιηθεί στην πρώτη επίσκεψη του ασθενούς, ενώ στη η δεύτερη ο συνήθης παροχέας χρησιμοποιείται πάντα πρώτος για ένα συγκεκριμένο επεισόδιο περίθαλψης.

Αφού βρούμε τις τιμές της ακριβούς κατανομής των τριών δεικτών για το καθένα από τα έξι μοντέλα, θα τις συγκρίνουμε με τα εμπειρικά δεδομένα που έχουμε συγκεντρώσει για να διαπιστώσουμε ποιο προσεγγίζει καλύτερα στην πραγματικότητα.

Ας εξετάσουμε πρώτα τα αποτελέσματα των ακριβών κατανομών των τριών δεικτών όταν το πλήθος των επισκέψεων είναι ίσο με 4. Στον παρακάτω πίνακα φαίνονται τα εμπειρικά δεδομένα μας μαζί με τις τιμές της ακριβούς κατανομής του *SECON*:

<i>SECON</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
0	0.012	0.4096	0.0625	0.01235	0.31641	0.6561	0.814506
1/4	0.060	0.4096	0.2500	0.09877	0.42188	0.2916	0.171475
3/4	0.124	0.1536	0.3750	0.2963	0.21094	0.0486	0.013538
3/4	0.804	0.0256	0.2500	0.39506	0.04688	0.0036	0.000475
1	0.000	0.0016	0.0625	0.19753	0.00391	0.0001	0.000006

και στον παρακάτω πίνακα βλέπουμε τα τετράγωνα των αποκλίσεων των 6 μοντέλων από τα εμπειρικά δεδομένα μας καθώς και το άθροισμά τους:

<i>SECON</i>	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
0	0.158086	0.0025503	1.197E-07	0.092663013	0.414865	0.64401588
1/4	0.12222	0.0361	0.0015027	0.130953516	0.053639	0.01242668
2/4	0.000876	0.063001	0.0296859	0.007558216	0.005685	0.01220185
3/4	0.605907	0.306916	0.1672303	0.573238266	0.64064	0.64565243
1	2.56E-06	0.0039063	0.0390185	1.52568E-05	1E-08	3.6E-11
Άθροισμα	0.887091	0.4124735	0.2374375	0.804428267	1.114829	1.31429683

Με κόκκινο χρώμα έχουμε σημειώσει το μοντέλο που έχει τις μικρότερες αποκλίσεις από τα εμπειρικά δεδομένα και με μπλε χρώμα το μοντέλο με τις αμέσως μικρότερες αποκλίσεις.

Για τον δείκτη *UPC*, όπου έχουμε τις δύο διαφορετικές εργοδικές κατανομές (περιπτώσεις (i) και (ii)), εξετάζουμε δώδεκα διαφορετικά μοντέλα, των οποίων οι ακριβείς κατανομές είναι οι παρακάτω:

<i>UPC</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές				
		A(i) και A(ii)	B(i)	B(ii)	Γ(i)	Γ(ii)
0	0.000	0.4096	0.53594	0.33496	0.6162	0.25675
1/4	0.000	0.4096	0.24063	0.29395	0.1528	0.22705
2/4	0.056	0.1536	0.13594	0.19922	0.1051	0.18287
3/4	0.140	0.0256	0.0625	0.10938	0.0667	0.1358
1	0.804	0.0016	0.025	0.0625	0.0592	0.19753

<i>UPC</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		Δ(i)	Δ(ii)	E(i)	E(ii)	ΣΤ(i)	ΣΤ(ii)
0	0.000	0.65918	0.20599	0.37239	0.41894	0.35466	0.421156
1/4	0.000	0.11133	0.18585	0.46733	0.45817	0.4962	0.485671
2/4	0.056	0.08418	0.15991	0.14839	0.11582	0.14366	0.091356
3/4	0.140	0.06094	0.13184	0.0117	0.00698	0.00546	0.001811
1	0.804	0.08438	0.31641	0.0002	0.0001	2.50E-05	0.000006

Τα τετράγωνα των αποκλίσεων από τα εμπειρικά δεδομένα δίνονται στον επόμενο πίνακα, όπου έχει και πάλι σημειωθεί με κόκκινο χρώμα το μοντέλο με τις μικρότερες αποκλίσεις και με μπλε χρώμα το μοντέλο με τις αμέσως μικρότερες αποκλίσεις.

<i>UPC</i>	$(O - A)^2$	$(O - B(i))^2$	$(O - B(ii))^2$	$(O - \Gamma(i))^2$	$(O - \Gamma(ii))^2$
0	0.167772	0.2872285	0.1121988	0.37970244	0.065921
1/4	0.167772	0.0579004	0.0864038	0.02334784	0.051552
2/4	0.009526	0.0063901	0.0205116	0.00241081	0.016096
3/4	0.013087	0.0060063	0.0009379	0.00537289	1.76E-05
1	0.643846	0.606841	0.5498223	0.55472704	0.367806
Άθροισμα	1.002003	0.9643662	0.7698744	0.96556102	0.501392

<i>UPC</i>	$(O - \Delta(i))^2$	$(O - \Delta(ii))^2$	$(O - E(i))^2$	$(O - E(ii))^2$	$(O - \Sigma T(i))^2$	$(O - \Sigma T(ii))^2$
0	0.43451827	0.04243338	0.1386725	0.17550737	0.1257823	0.1773724
1/4	0.01239392	0.03454098	0.2183927	0.20991975	0.2462115	0.2358763
2/4	0.00079411	0.01079773	0.0085355	0.00357831	0.0076839	0.00125
3/4	0.00625096	6.6652E-05	0.0164609	0.01769565	0.0181005	0.0190962
1	0.51786014	0.23774767	0.6460944	0.64625521	0.6463758	0.6464064
Άθροισμα	0.97181741	0.32558641	1.0281559	1.05295629	1.044154	1.0800013

Για να βρούμε την ακριβή κατανομή του δείκτη *NOP* θα πρέπει να σχηματίσουμε τον πίνακα πιθανοτήτων μεταπηδήσεως M , για $m = 5$. Ο πίνακας M , στην περίπτωση αυτή θα έχει διαστάσεις $m2^{m-1} \times m2^{m-1}$, δηλαδή 80×80 .

Οι τιμές της ακριβούς κατανομής του δείκτη *NOP* για τα έξι μοντέλα είναι:

<i>NOP</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
1/5	0.804	0.008	0.125000	0.296296	0.421875	0.001000	0.000125
2/5	0.168	0.224	0.476562	0.502315	0.458008	0.133313	0.094555
3/5	0.028	0.576	0.351563	0.187500	0.114258	0.592312	0.583805
4/5	0.000	0.192	0.046875	0.013899	0.005859	0.273375	0.321515
1	0.000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000

ενώ για τα τετράγωνα των αποκλίσεων παίρνουμε τις παρακάτω τιμές (με κόκκινο χρώμα έχουμε το μοντέλο του οποίου οι τιμές προσαρμόζονται καλύτερα στα δεδομένα μας και με μπλε το αμέσως καλύτερο μοντέλο):

<i>NOP</i>	$(\mathbf{O} - \mathbf{A})^2$	$(\mathbf{O} - \mathbf{B})^2$	$(\mathbf{O} - \mathbf{\Gamma})^2$	$(\mathbf{O} - \mathbf{\Delta})^2$	$(\mathbf{O} - \mathbf{E})^2$	$(\mathbf{O} - \mathbf{\Sigma T})^2$
1/5	0.633616	0.4610410	0.2577634	0.1460195156	0.644809	0.646215016
2/5	0.003136	0.0952105	0.1117665	0.0841046401	0.001203	0.005394168
3/5	0.300304	0.1046930	0.0254403	0.0074404426	0.318448	0.308919198
4/5	0.036864	0.0021973	0.0001929	0.0000343279	0.074734	0.103371895
1	0.000000	0.0000000	0.0000000	0.0000000000	0.000000	0.000000000
Άθροισμα	0.973920	0.6631418	0.3951630	0.2375989261	1.039194	1.063900277

Όπως βλέπουμε από τους πίνακες των τετραγώνων των αποκλίσεων των τιμών των ακριβών κατανομών από τα εμπειρικά δεδομένα, για τους δείκτες *NOP* και *UPC*, το μοντέλο που προσαρμόζεται καλύτερα στα δεδομένα μας είναι το $(\mathbf{\Delta})$ με $k = 4/3$ και αμέσως μετά το $(\mathbf{\Gamma})$ με $k = 3/2$. Μάλιστα, για το δείκτη *UPC* όπου παίζει ρόλο και η αρχική κατανομή, το καλύτερο μοντέλο είναι το $\Delta(ii)$ και αμέσως μετά το $\Gamma(ii)$, όπου $\boldsymbol{\pi} = (1,0,0,0,0)$, δηλαδή όταν έχουμε υποθέσει από την αρχή ότι ο συνήθης παροχέας χρησιμοποιείται πάντα πρώτος. Για τα μοντέλα $(\mathbf{\Gamma})$ και $(\mathbf{\Delta})$ έχουμε υποθέσει ότι η πιθανότητα αλλαγής γιατρού από επίσκεψη σε επίσκεψη είναι κατά πολύ μικρότερη από την πιθανότητα παραμονής στον ίδιο γιατρό. Συγκεκριμένα, στο μοντέλο $(\mathbf{\Gamma})$ η πιθανότητα μη αλλαγής γιατρού p_{ii} , $i = 1,2,3,4,5$ είναι ίση με $2/3$ και το υπόλοιπο $1/3$ μοιράζεται εξίσου στις πιθανότητες p_{ij} , $\forall i \neq j = 1,2,3,4,5$. Ομοίως, στο μοντέλο $(\mathbf{\Delta})$ έχουμε $p_{ii} = 3/4$ και $p_{ij} = 1/16$, $\forall i \neq j$. Το μοντέλο $(\mathbf{\Delta})$, το οποίο έχει μεγαλύτερο p_{ii} από το μοντέλο $(\mathbf{\Gamma})$, είναι εκείνο που προσαρμόζεται καλύτερα στα δεδομένα μας, όσον αφορά τους δείκτες *NOP* και *UPC*. Αυτό δείχνει ότι και στην πραγματικότητα είναι πιο πιθανό σε μια σειρά επισκέψεων να χρησιμοποιηθεί όλες τις φορές ή τουλάχιστον τις περισσότερες φορές ο ίδιος/συνήθης γιατρός. Αντιθέτως, τα μοντέλα με τη χειρότερη προσαρμογή είναι τα (\mathbf{E}) και $(\mathbf{\Sigma T})$ τα οποία περιγράφουν το μη ρεαλιστικό σενάριο της συνεχούς αλλαγής γιατρού σε μια ακολουθία επισκέψεων.

Βλέπουμε όμως ότι για το δείκτη *SECON*, το μοντέλο με την καλύτερη προσαρμογή είναι το $(\mathbf{\Gamma})$ και αμέσως μετά το (\mathbf{B}) . Αυτό συμβαίνει επειδή η φύση του δείκτη *SECON* είναι

διαφορετική από τη φύση των άλλων δύο δεικτών, *NOP* και *UPC*. Στον υπολογισμό του *SECON* παίζει πρωταρχικό ρόλο η σειρά των επισκέψεων, ενώ στους άλλους δύο δείκτες όχι, π.χ. για την ακολουθία επισκέψεων $\{A, B, A, B, A, B, A\}$ με $m = A, B, C, D, E$ έχουμε $NOP = 2/5$, $UPC = 4/7$ και $SECON = 0$. Ο δείκτης *NOP* υποδηλώνει μια καλή συνέχεια, ο *UPC* μια σχετικά καλή συνέχεια και ο *SECON* καμία συνέχεια. Μπορεί δηλαδή να χρησιμοποιήθηκαν μόνο 2 από τους διαθέσιμους παροχείς και μάλιστα ο ένας πάνω από το 50% των επισκέψεων, επειδή όμως σε οποιοσδήποτε δύο συνεχόμενες επισκέψεις ο παροχέας δεν είναι ποτέ ο ίδιος, η τιμή του *SECON* είναι ίση με το μηδέν. Η λογική στην οποία βασίζεται ο *SECON* είναι πιο σύνθετη από τη λογική του *NOP* και του *UPC*, και γι' αυτό για το δείκτη *SECON* το μοντέλο με την καλύτερη προσαρμογή είναι διαφορετικό απ' ό,τι για τους δείκτες *NOP* και *UPC*. Εντούτοις, στα μοντέλα (Γ) και (Β) έχουμε υποθέσει ότι $p_{ii} > p_{ij}$, $\forall i \neq j = 1, 2, 3, 4, 5$. Επομένως, μπορούμε να πούμε ότι τα εμπειρικά μας δεδομένα περιγράφονται καλύτερα από μοντέλα όπου η πιθανότητα αλλαγής γιατρού είναι μικρότερη από την πιθανότητα χρησιμοποίησης του ίδιου γιατρού από τη μία επίσκεψη στην άλλη, δηλαδή από μοντέλα που απεικονίζουν μια καλή συνέχεια της υγειονομικής περίθαλψης.

Όπως θα δούμε στη συνέχεια, τα αποτελέσματα όταν το πλήθος των επισκέψεων είναι 5, 6 και 7 είναι παρόμοια με τα αποτελέσματα όταν $n = 4$. Ας δούμε πρώτα τα αποτελέσματα για $n = 5$ για το δείκτη *SECON*. Οι εμπειρικές τιμές και οι ακριβείς τιμές του *SECON* στα 6 μοντέλα δίνονται στον παρακάτω πίνακα:

<i>SECON</i>	Εμπειρικές Τιμές (Ο)	Ακριβείς τιμές					
		Α	Β	Γ	Δ	Ε	ΣΤ
0	0.0073	0.32768	0.03125	0.004115	0.237305	0.59049	0.7737809
1/5	0.0219	0.40960	0.15625	0.041152	0.395508	0.32805	0.2036266
2/5	0.1387	0.20480	0.31250	0.164609	0.263672	0.07290	0.0214344
3/5	0.1095	0.05120	0.31250	0.329218	0.087891	0.00810	0.0011281
4/5	0.7226	0.00640	0.15625	0.329218	0.014648	0.00045	0.0000297
1	0.0000	0.00032	0.03125	0.131688	0.000976	0.00001	0.0000003

Για τα τετράγωνα των αποκλίσεων παίρνουμε τα επόμενα αποτελέσματα:

<i>SECON</i>	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
0	0.10264334	0.0005736	1.0144E-05	0.0529023	0.34011058	0.58749297
1/5	0.15031129	0.01804992	0.00037064	0.13958294	0.09372782	0.03302456
2/5	0.00436921	0.03020644	0.00067128	0.015618	0.00432964	0.01375122
3/5	0.00339889	0.041209	0.048276	0.00046695	0.01028196	0.01174447
4/5	0.51294244	0.32075232	0.1547494	0.50119603	0.52150062	0.52210784
1	1.024E-07	0.00097656	0.01734173	9.5258E-07	1E-10	9E-14
Άθροισμα	0.77366528	0.41176785	0.22141919	0.70976717	0.96995062	1.16812106

Για το δείκτη *UPC* έχουμε τις εξής εμπειρικές και ακριβείς τιμές για 12 μοντέλα:

<i>UPC</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές				
		A(i) και A(ii)	B(i)	B(ii)	Γ(i)	Γ(ii)
0	0.0000	0.32768	0.468945	0.293091	0.564853	0.235356
1/5	0.0000	0.40960	0.248828	0.281128	0.158719	0.215905
2/5	0.0073	0.20480	0.154883	0.207031	0.113928	0.179913
3/5	0.1095	0.05120	0.080469	0.125000	0.076080	0.138374
4/5	0.1606	0.00640	0.034375	0.062500	0.046914	0.098765
1	0.7226	0.00032	0.012500	0.031250	0.039506	0.131687

<i>UPC</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		Δ(i)	Δ(ii)	E(i)	E(ii)	ΣΤ(i)	ΣΤ(ii)
0	0.0000	0.617981	0.193119	0.288600	0.324675	0.270427	0.3211316
1/5	0.0000	0.115356	0.177669	0.459478	0.464546	0.483294	0.4949448
2/5	0.0073	0.089368	0.155701	0.216858	0.186694	0.222521	0.1731802
3/5	0.1095	0.066553	0.130737	0.033469	0.023175	0.023371	0.0106244
4/5	0.1606	0.047461	0.105469	0.001575	0.000900	0.000386	0.0001187
1	0.7226	0.063281	0.237305	0.000020	0.000010	0.000001	0.0000003

Τα τετράγωνα των αποκλίσεων των τιμών των ακριβών κατανομών από τα εμπειρικά δεδομένα δίνονται στους δύο ακόλουθους πίνακες όπου με κόκκινο χρώμα έχουμε σημειώσει

το μοντέλο που προσεγγίζει περισσότερο τις εμπειρικές τιμές μας και με μπλε χρώμα το αμέσως καλύτερο μοντέλο:

<i>UPC</i>	$(O - A)^2$	$(O - B(i))^2$	$(O - B(ii))^2$	$(O - \Gamma(i))^2$	$(O - \Gamma(ii))^2$
0	0.10737418	0.21990941	0.08590233	0.31905891	0.05539245
1/5	0.16777216	0.06191537	0.07903295	0.02519172	0.04661497
2/5	0.03900625	0.02178074	0.03989247	0.01136953	0.02979525
3/5	0.00339889	0.0008428	0.00024025	0.0011169	0.00083371
4/5	0.02377764	0.01593275	0.00962361	0.01292451	0.00382357
1	0.5216884	0.50424201	0.47796482	0.46661741	0.34917817
Άθροισμα	0.86301752	0.82462309	0.69265644	0.83627898	0.48563811

<i>UPC</i>	$(O - \Delta(i))^2$	$(O - \Delta(ii))^2$	$(O - E(i))^2$	$(O - E(ii))^2$	$(O - \Sigma\Gamma(i))^2$	$(O - \Sigma\Gamma(ii))^2$
0	0.38190052	0.03729495	0.08328996	0.10541386	0.07313076	0.1031255
1/5	0.01330701	0.03156627	0.21112003	0.21580299	0.23357309	0.24497036
2/5	0.00673516	0.02202286	0.04391456	0.03218221	0.04632008	0.02751624
3/5	0.00184444	0.00045101	0.00578071	0.00745201	0.0074182	0.00977638
4/5	0.01280043	0.00303943	0.02528895	0.02550409	0.02566853	0.02575425
1	0.43470154	0.23551124	0.52212186	0.52213631	0.52214931	0.52215033
Άθροισμα	0.8512891	0.32988575	0.89151607	0.90849145	0.90825998	0.93329306

Στον επόμενο πίνακα έχουμε τις εμπειρικές και τις ακριβείς τιμές για το δείκτη *NOP* για τα 6 μοντέλα:

<i>NOP</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
1/5	0.7226	0.0016	0.062500	0.197531	0.316406	0.000100	0.000006
2/5	0.2482	0.0960	0.360352	0.475502	0.477601	0.044227	0.027303
3/5	0.0292	0.4800	0.442383	0.281829	0.185852	0.415758	0.373868
4/5	0.0000	0.3840	0.128906	0.043981	0.019775	0.478406	0.522463
1	0.0000	0.0384	0.005859	0.001157	0.000366	0.061509	0.076360

Για τις αποκλίσεις παίρνουμε τον παρακάτω πίνακα (με κόκκινο χρώμα έχουμε σημειώσει το καλύτερο μοντέλο και με μπλε το δεύτερο καλύτερο):

<i>NOP</i>	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
1/5	0.519841	0.43573201	0.27569745	0.16499357	0.52200625	0.52214209
2/5	0.02316484	0.01257807	0.0516662	0.05262482	0.04160498	0.04879548
3/5	0.20322064	0.17072019	0.06382141	0.02453985	0.14942709	0.11879603
4/5	0.147456	0.01661676	0.00193433	0.00039105	0.2288723	0.27296759
1	0.00147456	3.4328E-05	1.3386E-06	1.3396E-07	0.00378336	0.00583085
Άθροισμα	0.89515704	0.63568136	0.39312073	0.24254942	0.94569398	0.96853204

Βλέπουμε πως ό,τι ισχύει όταν το πλήθος επισκέψεων είναι ίσο με 4, ισχύει και όταν το πλήθος των επισκέψεων είναι ίσο με 5. Δηλαδή, για τους δείκτες *NOP* και *UPC* το καλύτερο μοντέλο είναι το (Δ) και (Δii) αντίστοιχα, ενώ για το δείκτη *SECON* το καλύτερο μοντέλο είναι το (Γ).

Ας δούμε τώρα τι γίνεται όταν το πλήθος των επισκέψεων είναι ίσο με $n = 6$. Υπολογίζουμε τις ακριβείς τιμές του δείκτη *SECON* για τα 6 μοντέλα που προαναφέραμε και παίρνουμε τον παρακάτω πίνακα που περιέχει και τις εμπειρικές τιμές:

<i>SECON</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
0	0.0192	0.26214	0.015625	0.001372	0.177979	0.531441	0.73509189
1/6	0.0385	0.39322	0.093750	0.016461	0.355957	0.354294	0.23213428
2/6	0.0192	0.24576	0.234375	0.082305	0.296631	0.098415	0.03054398
3/6	0.1346	0.08192	0.312500	0.219479	0.131836	0.014580	0.00214344
4/6	0.1154	0.01536	0.234375	0.329218	0.032959	0.001215	0.00008461
5/6	0.6731	0.00154	0.093750	0.263374	0.004394	0.000054	0.00000178
1	0	0.00006	0.015625	0.087791	0.000244	0.000001	0.00000002

Για τις αποκλίσεις παίρνουμε τα επόμενα αποτελέσματα όπου με κόκκινο σημειώνεται το μοντέλο με τις μικρότερες αποκλίσεις και με μπλε το αμέσως καλύτερο μοντέλο:

<i>SECON</i>	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
0	0.05901984	1.2781E-05	0.00031784	0.02521077	0.26239084	0.5125012
1/6	0.12582628	0.00305256	0.00048572	0.10077895	0.09972585	0.03749423
2/6	0.05132943	0.04630028	0.00398224	0.07696796	0.00627502	0.00012869
3/6	0.00277518	0.03164841	0.00720444	7.6397E-06	0.0144048	0.01754474
4/6	0.010008	0.01415505	0.04571814	0.00679652	0.01303821	0.01329764
5/6	0.45099283	0.33564642	0.1678754	0.44716771	0.45299092	0.45306121
1	3.6E-09	0.00024414	0.00770726	5.9536E-08	1E-12	4E-16
Άθροισμα	0.69995158	0.43105965	0.23329103	0.65692961	0.84882564	1.03402771

Και πάλι βλέπουμε ότι το μοντέλο που προσαρμόζεται καλύτερα στις εμπειρικές τιμές του *SECON* είναι το (Γ) και αμέσως μετά το (B).

Ενδιαφέρον παρουσιάζει το γεγονός ότι παρόλο που αυξάνεται ο αριθμός των επισκέψεων, το μοντέλο με την καλύτερη προσαρμογή παραμένει το ίδιο, τουλάχιστον για το δείκτη *SECON*.

Ας δούμε αν ισχύει κάτι ανάλογο και για τους δείκτες *UPC* και *NOP*. Στον παρακάτω πίνακα έχουμε τις εμπειρικές τιμές και τις ακριβείς τιμές που προκύπτουν από τα 6 διαφορετικά μοντέλα που θεωρήσαμε στην αρχή της ενότητας για το δείκτη *UPC*:

<i>UPC</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές				
		A(i) και A(ii)	B(i)	B(ii)	Γ(i)	Γ(ii)
0	0.0000	0.262144	0.410327	0.256454	0.517782	0.215743
1/6	0.0000	0.393216	0.251221	0.266922	0.162609	0.205045
2/6	0.0192	0.245760	0.169702	0.211487	0.121126	0.176413
3/6	0.0577	0.081920	0.096875	0.138184	0.084388	0.140111
4/6	0.0962	0.015360	0.046875	0.076172	0.054835	0.103566
5/6	0.1538	0.001536	0.018750	0.035156	0.032922	0.071331
1	0.6731	0.000064	0.006250	0.015625	0.026338	0.087791

UPC	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		$\Delta(i)$	$\Delta(ii)$	$E(i)$	$E(ii)$	$\Sigma T(i)$	$\Sigma T(ii)$
0	0.0000	0.579357	0.181049	0.223665	0.251623	0.20620027	0.24486282
1/6	0.0000	0.118446	0.169784	0.431504	0.444858	0.4485313	0.47241875
2/6	0.0192	0.093762	0.151448	0.272428	0.248414	0.28687372	0.24786039
3/6	0.0577	0.071545	0.129402	0.066173	0.051107	0.05588264	0.03388935
4/6	0.0962	0.052515	0.106293	0.006030	0.003887	0.00248707	0.00096132
5/6	0.1538	0.036914	0.084045	0.000198	0.000110	0.00002494	0.00000735
1	0.6731	0.047461	0.177979	0.000002	0.000001	0.00000006	0.00000002

Ο πίνακας με τις αποκλίσεις απ' όπου φαίνεται το καλύτερο και το δεύτερο καλύτερο μοντέλο (με κόκκινο και μπλε χρώμα αντίστοιχα) δίνεται παρακάτω:

UPC	$(O - A)^2$	$(O - B(i))^2$	$(O - B(ii))^2$	$(O - \Gamma(i))^2$	$(O - \Gamma(ii))^2$
0	0.06871948	0.16836825	0.06576865	0.2680982	0.04654504
1/6	0.15461882	0.06311199	0.07124735	0.02644169	0.04204345
2/6	0.05132943	0.02265085	0.03697429	0.01038891	0.02471593
3/6	0.00058661	0.00153468	0.00647767	0.00071225	0.00679157
4/6	0.00653511	0.00243296	0.00040112	0.00171106	5.4258E-05
5/6	0.02318433	0.0182385	0.0140764	0.01461149	0.00680114
1	0.45297746	0.44468892	0.43227338	0.41830108	0.34258663
Άθροισμα	0.75795123	0.72102615	0.62721887	0.74026468	0.46953801

UPC	$(O - \Delta(i))^2$	$(O - \Delta(ii))^2$	$(O - E(i))^2$	$(O - E(ii))^2$	$(O - \Sigma T(i))^2$	$(O - \Sigma T(ii))^2$
0	0.33565453	0.03277874	0.05002603	0.06331413	0.04251855	0.0599578
1/6	0.01402945	0.02882661	0.1861957	0.19789864	0.20118033	0.22317948
2/6	0.00555949	0.01748953	0.06412442	0.05253906	0.07164922	0.05228557
3/6	0.00019168	0.00514118	7.1792E-05	4.3468E-05	3.3028E-06	0.00056695
4/6	0.00190838	0.00010187	0.00813063	0.00852169	0.00878211	0.00907041
5/6	0.01366234	0.00486576	0.02359357	0.02362062	0.02364677	0.02365218
1	0.39142416	0.2451448	0.45306092	0.45306226	0.45306353	0.45306358
Άθροισμα	0.76243004	0.33434849	0.78520307	0.79899987	0.80084381	0.82177597

Οι εμπειρικές και οι ακριβείς τιμές του δείκτη *NOP* στα 6 μοντέλα δίνονται παρακάτω:

<i>NOP</i>	Εμπειρικές Τιμές (Ο)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
1/5	0.6731	0.00032	0.031250	0.131687	0.237305	0.00001	0.0000003
2/5	0.2692	0.03968	0.256470	0.422470	0.467152	0.01446	0.0078556
3/5	0.0192	0.34560	0.466919	0.353733	0.252170	0.25852	0.2157341
4/5	0.0385	0.49920	0.223389	0.087288	0.041771	0.55786	0.5759651
1	0.0000	0.11520	0.021972	0.004822	0.001602	0.16915	0.2004449

Ο πίνακας με τις αποκλίσεις των ακριβών τιμών από τα εμπειρικά δεδομένα έχει ως εξής:

<i>NOP</i>	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
1/5	0.45263293	0.41197142	0.29312804	0.18991728	0.45305015	0.45306321
2/5	0.05267943	0.00016205	0.02349169	0.03918499	0.06489247	0.0683009
3/5	0.10653696	0.2004523	0.11191233	0.05427502	0.05727406	0.03862565
4/5	0.21224449	0.03418394	0.00238027	1.0699E-05	0.26973481	0.28886873
1	0.01327104	0.00048277	2.3252E-05	2.5664E-06	0.02861172	0.04017816
Άθροισμα	0.83736485	0.64725249	0.43093558	0.28339056	0.87356321	0.88903665

Για το δείκτη *UPC* το καλύτερο μοντέλο είναι το (Δ_{ii}) και αμέσως μετά το (Γ_{ii}) και για το δείκτη *NOP* το καλύτερο μοντέλο είναι το (Δ) και αμέσως μετά το (Γ), όπως ακριβώς ισχύει και για $n = 4$ και για $n = 5$. Ούτε εδώ έχουμε καμία μεταβολή ως προς το ποιο μοντέλο προσαρμόζεται καλύτερα στα δεδομένα μας, παρόλο που έχει αυξηθεί ο αριθμός των επισκέψεων.

Τέλος, ας δούμε τι συμβαίνει όταν το πλήθος των επισκέψεων είναι ίσο με 7. Πρώτα για το δείκτη *SECON*, ο πίνακας με τις εμπειρικές και τις ακριβείς τιμές για τα 6 μοντέλα είναι ο παρακάτω:

SECON	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
0	0.0000	0.20972	0.0078125	0.000457	0.133484	0.4782969	0.6983372961
1/7	0.0000	0.36700	0.0546875	0.006401	0.311462	0.3720087	0.2572821617
2/7	0.0909	0.27525	0.1640625	0.038409	0.311462	0.1240029	0.0406234992
3/7	0.0303	0.11469	0.2734375	0.128029	0.173035	0.0229635	0.0035634648
4/7	0.1515	0.02867	0.2734375	0.256059	0.057678	0.0025515	0.0001875508
5/7	0.1515	0.00430	0.1640625	0.307270	0.011536	0.0001701	0.0000059226
6/7	0.5758	0.00036	0.0546875	0.204847	0.001282	0.0000063	0.0000001039
1	0.0000	0.00001	0.0078125	0.058528	0.000061	0.0000001	0.0000000008

Για τις αποκλίσεις παίρνουμε τα ακόλουθα αποτελέσματα:

SECON	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
0	0.04398248	6.1035E-05	2.0885E-07	0.01781798	0.22876792	0.487674979
1/7	0.134689	0.00299072	4.0973E-05	0.09700858	0.13839047	0.066194111
2/7	0.03398492	0.00535275	0.00275531	0.0486476	0.0010958	0.002527727
3/7	0.00712167	0.05911584	0.00955096	0.02037328	5.3824E-05	0.000714842
4/7	0.01508721	0.01486875	0.01093258	0.00880257	0.02218566	0.022895457
5/7	0.02166784	0.00015782	0.02426429	0.01958992	0.02290074	0.022950455
6/7	0.33113119	0.27155824	0.13760613	0.33007093	0.33153838	0.33154552
1	1E-10	6.1035E-05	0.00342553	3.721E-09	1E-14	6.4E-19
Άθροισμα	0.58766432	0.3541662	0.18857598	0.54231086	0.7449328	0.934503092

Με κόκκινο χρώμα σημειώνεται το μοντέλο με την καλύτερη προσαρμογή και με μπλε χρώμα το αμέσως καλύτερο.

Για το δείκτη *UPC* έχουμε τους παρακάτω δύο πίνακες με τις ακριβείς τιμές για τα 12 μοντέλα, καθώς και τα εμπειρικά δεδομένα:

UPC	Εμπειρικές Τιμές (O)	Ακριβείς τιμές				
		A(i) και A(ii)	B(i)	B(ii)	Γ(i)	Γ(ii)
0	0.0000	0.209715	0.359036	0.224398	0.474634	0.197764
1/7	0.0000	0.367002	0.249127	0.251875	0.164749	0.194495
2/7	0.0000	0.275251	0.180789	0.213089	0.126852	0.172464
3/7	0.0909	0.114688	0.111536	0.149017	0.091649	0.141095
4/7	0.0909	0.028672	0.059278	0.088867	0.062144	0.107624
5/7	0.0909	0.004301	0.026953	0.045410	0.039369	0.076824
6/7	0.1515	0.000358	0.010156	0.019531	0.023045	0.051212
1	0.5758	0.000013	0.003125	0.007813	0.017558	0.058528

UPC	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		Δ(i)	Δ(ii)	E(i)	E(ii)	ΣΤ(i)	ΣΤ(ii)
0	0.0000	0.543147	0.169734	0.1733406	0.1950081	0.15722705	0.1867078997
1/7	0.0000	0.120699	0.162190	0.3928572	0.4105114	0.403020115	0.4326746086
2/7	0.0000	0.097429	0.147173	0.3117171	0.2950745	0.331785393	0.3054166378
3/7	0.0909	0.075955	0.127856	0.1056337	0.0877863	0.098676847	0.0707049308
4/7	0.0909	0.057137	0.106819	0.0154741	0.0110199	0.009072676	0.0044195503
5/7	0.0909	0.041363	0.086002	0.0009533	0.0005866	0.000215732	0.0000759341
6/7	0.1515	0.028674	0.066742	0.0000238	0.0000131	0.000001529	0.0000004379
1	0.5758	0.035596	0.133484	0.0000002	0.0000001	0.000000003	0.0000000008

Στους επόμενους πίνακες παρουσιάζονται οι αποκλίσεις των ακριβών τιμών από τα εμπειρικά δεδομένα. Όπως και προηγουμένως, με κόκκινο χρώμα έχει σημειωθεί το μοντέλο που προσεγγίζει περισσότερο στα εμπειρικά δεδομένα και με μπλε χρώμα το αμέσως καλύτερο μοντέλο.

UPC	$(O - A)^2$	$(O - B(i))^2$	$(O - B(ii))^2$	$(O - \Gamma(i))^2$	$(O - \Gamma(ii))^2$
0	0.04398038	0.12890685	0.05035446	0.22527743	0.0391106
1/7	0.13469047	0.06206426	0.06344102	0.02714223	0.03782831
2/7	0.07576311	0.03268466	0.04540692	0.01609143	0.02974383
3/7	0.00056587	0.00042584	0.00337759	5.61E-07	0.00251954
4/7	0.00387232	0.00099995	4.1331E-06	0.00082691	0.00027969
5/7	0.00749939	0.00408922	0.00206934	0.00265544	0.0001983
6/7	0.0228439	0.01997813	0.01741582	0.01650069	0.01005768
1	0.33153067	0.32795666	0.32260923	0.31163413	0.26757032
Άθροισμα	0.62074612	0.57710557	0.50467851	0.60012883	0.38730827

<i>UPC</i>	$(O - \Delta(i))^2$	$(O - \Delta(\ddot{i}))^2$	$(O - E(i))^2$	$(O - E(\ddot{i}))^2$	$(O - \Sigma T(i))^2$	$(O - \Sigma T(\ddot{i}))^2$
0	0.295008664	0.02880963	0.03004696	0.03802816	0.024720551	0.03485984
1/7	0.014568249	0.0263056	0.15433678	0.16851961	0.162425213	0.187207317
2/7	0.00949241	0.02165989	0.09716755	0.08706896	0.110081547	0.093279323
3/7	0.000223353	0.00136575	0.00021708	9.6951E-06	6.04793E-05	0.000407841
4/7	0.00113994	0.00025341	0.00568907	0.00638083	0.006695711	0.007478868
5/7	0.002453914	2.399E-05	0.00809041	0.00815651	0.008223636	0.008249011
6/7	0.015086226	0.00718392	0.02294504	0.02294828	0.022951787	0.022952117
1	0.291820362	0.19564344	0.33154541	0.33154552	0.331545637	0.331545639
Άθροισμα	0.629793118	0.28124563	0.6500383	0.66265757	0.666704561	0.685979956

Ο πίνακας με τις εμπειρικές και τις ακριβείς τιμές για το δείκτη *NOP* έχει ως εξής:

<i>NOP</i>	Εμπειρικές Τιμές (O)	Ακριβείς τιμές					
		A	B	Γ	Δ	E	ΣΤ
1/5	0.5758	0.000064	0.015625	0.087791	0.177978	0.000001	0.00000008
2/5	0.3333	0.016128	0.175919	0.360748	0.438887	0.004710	0.00225879
3/5	0.0909	0.231168	0.446365	0.400395	0.308240	0.151949	0.11885751
4/5	0.0000	0.537600	0.312195	0.138969	0.070682	0.548672	0.54164707
1	0.0000	0.215040	0.049896	0.012097	0.004213	0.294668	0.33723661

Για τις αποκλίσεις παίρνουμε τον επόμενο πίνακα, όπου και πάλι με κόκκινο χρώμα σημειώνεται το μοντέλο με τις μικρότερες αποκλίσεις και με μπλε χρώμα το μοντέλο με τις αμέσως μικρότερες αποκλίσεις:

<i>NOP</i>	$(O - A)^2$	$(O - B)^2$	$(O - \Gamma)^2$	$(O - \Delta)^2$	$(O - E)^2$	$(O - \Sigma T)^2$
1/5	0.45263293	0.41197142	0.29312804	0.18991728	0.45305015	0.45306321
2/5	0.05267943	0.00016205	0.02349169	0.03918499	0.06489247	0.0683009
3/5	0.10653696	0.2004523	0.11191233	0.05427502	0.05727406	0.03862565
4/5	0.21224449	0.03418394	0.00238027	1.0699E-05	0.26973481	0.28886873
1	0.01327104	0.00048277	2.3252E-05	2.5664E-06	0.02861172	0.04017816
Άθροισμα	0.83736485	0.64725249	0.43093558	0.28339056	0.87356321	0.88903665

6.4 Συμπεράσματα

Διαπιστώνουμε ότι για το δείκτη *SECON* το καλύτερο μοντέλο είναι το (Γ), για το δείκτη *UPC* το (Δii) και για το δείκτη *NOP* το μοντέλο (Δ) ανεξαρτήτως του αριθμού των επισκέψεων που έχουν πραγματοποιηθεί στους 5 παροχείς. Και τα τρία μοντέλα είναι μοντέλα που αποδίδουν μεγαλύτερη πιθανότητα παραμονής στον ίδιο παροχέα και μικρές πιθανότητες μετάβασης σε άλλους παροχείς. Με άλλα λόγια, οι ασθενείς έχουν την τάση να επιλέγουν τον ίδιο γιατρό για τα διάφορα περιστατικά υγείας που τους παρουσιάζονται και σπανίως αλλάζουν γιατρό κατά τη διάρκεια των επισκέψεων. Για τη συγκεκριμένη ειδικότητα του παθολόγου, το αποτέλεσμα αυτό είναι λογικό, αφού όλοι μας έχουμε έναν συγκεκριμένο παθολόγο τον οποίο επισκεπτόμαστε είτε για απλά περιστατικά είτε για να μας παραπέμψει σε γιατρό άλλης ειδικότητας. Αν αντί του παθολόγου χρησιμοποιούσαμε στη μελέτη μας γιατρό άλλης ειδικότητας, π.χ. καρδιολόγο ή ενδοκρινολόγο, τα αποτελέσματα ίσως να ήταν διαφορετικά επειδή τους γιατρούς αυτούς αφενός τους επισκεπτόμαστε σπανιότερα, αφετέρου πολλές φορές τους αλλάζουμε για να πάρουμε δεύτερη ή και τρίτη γνώμη από άλλους της ίδιας ειδικότητας. Γενικά όμως μπορούμε να πούμε ότι υπάρχει μια αρκετά καλή συνέχεια της υγειονομικής περίθαλψης, όσον αφορά τις επισκέψεις ασθενών στους παθολόγους του ΙΚΑ, όπως φανερώνουν και οι τρεις δείκτες συνέχειας που μελετήσαμε.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Breslau N., Reeb K.G. (1975). Continuity of Care in a University-Based Practice, *Journal of Medical Education*, **50**, 965-969.
- Ejlertsson G., Berg S. (1984). Continuity of Care Measures: An Analytic and Empirical Comparison, *Medical Care*, **22**, 231-239.
- Eriksson E.A. (1990). Continuity of Care Measures: Random Assignment of Patients to Providers and the Impact of Utilization Model, *Medical Care*, **28**, 180-190.
- Eriksson E.A., Mattsson L.G. (1983). Quantitative Measurement of Continuity of Care: Measures in Use and an Alternative Approach, *Medical Care*, **21**, 858-875.
- Fu J.C., Koutras M.V. (1994). Distribution Theory of Runs : a Markov Chain Approach, *Journal of the American Statistical Association*, **89**, 1050-1058.
- Fu J.C., Lou W.Y.W. (2000). On the exact distribution of *SECON* and its application, *Statistica Sinica*, **10**, 999-1010.
- Lou W.Y.W. (1997). The exact distribution of overlapping counting patterns associated with a sequence of homogeneous Markov-dependent multi-state trials: an application to *SECON*, *Nonlinear Analysis, Theory, Methods and Applications*, **30**, 3349-3357.
- Lou W.Y.W. (2000). On the exact distribution of the continuity of care measure *NOP*, *Statistics and Probability Letters*, **48**, 361-368.
- Lou W.Y.W. (2001). The distribution of the Usual Provider Continuity index under Markov dependence, *Statistics and Probability Letters*, **54**, 269-276.
- Rogers J., Curtis P., (1980). The Concept and Measurement of Continuity in Primary Care, *American Journal of Public Health*, **70**, 122-127.
- Zhou B., (2003). Simulation of the Herfindahl-Hirschman Index: the case of the St.Louis Banking Geographic Market, *Geography Online*, **4**, 17-32.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ