

ΜΑΡΙΑ ΜΟΥΤΤΗ

Η Χρήση των Μοντέλων Πολλαπλών
Καταστάσεων σε Ανάλυση Δεδομένων
από Ασθενείς με Λοίμωξη σε ΜΕΘ

Διπλωματική Εργασία Ειδίκευσης
στην Στατιστική και Επιχειρησιακή Έρευνα



Πανεπιστήμιο Αθηνών
Τμήμα Μαθηματικών
Αθήνα 2012

Η παρούσα Διπλωματική Εργασία εκπονήθηκε
στα πλαίσια των σπουδών για την απόκτηση του
Μεταπτυχιακού Διπλώματος Ειδίκευσης στην
Στατιστική και Επιχειρησιακή Έρευνα
που απονέμει το Τμήμα Μαθηματικών του
Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών

Εγκρίθηκε στις από Εξεταστική Επιτροπή αποτελούμενη από τους:

Ονοματεπώνυμο	Βαθμίδα	Υπογραφή
Λουκία Μελιγκοτσίδου	Λέκτορας	
Απόστολος Μπουρνέτας	Καθηγητής	
Φώτης Σιάννης(επιβλέπων)	Επίκουρος Καθηγητής	

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω πολύ τον επιβλέποντα καθηγητή κ.Φώτη Σιάννη για την πρόταση του θέματος της εργασίας, την πολύτιμη καθοδήγηση και το ενδιαφέρον του για την εκπόνηση αυτής της εργασίας.

Επιπλέον θα ήθελα να ευχαριστήσω την κ.Λουκία Μελιγκοτσίδου και τον κ.Χαράλαμπο Δαμιανού που ήταν οι λόγοι της μετακίνησης του ενδιαφέροντος μου από τα υπολογιστικά μαθηματικά στην στατιστική στο προπτυχιακό επίπεδο με συνέπεια την επιθυμία μου να εισαχθώ σε αυτό το μεταπτυχιακό πρόγραμμα. Επίσης να ευχαριστήσω τον κ.Αποστόλη Μπουρνέτα που με τίμησε με τη συμμετοχή του στην εξεταστική επιτροπή, καθώς και την κ.Ελένη Αποστολοπούλου αναπληρώτρια καθηγήτρια του τμήματος Νοσηλευτικής του Πανεπιστημίου Αθηνών για τη συνεργασία και την παροχή των δεδομένων της παρούσας εργασίας.

Ειδικές ευχαριστίες θέλω να απευθύνω σε όλους τους συμφοιτητές μου στο μεταπτυχιακό πρόγραμμα για το άριστο κλίμα μεταξύ μας, τη συνεργασία που είχαμε και στους φίλους μου για τη συμπαράσταση τους. Ακόμα ένα ευχαριστώ στην κ.Ελισσάβετ Λέκκα, την γραμματέα του μεταπτυχιακού προγράμματος του Τμήματος για την άψογη εξυπηρέτηση της οποιαδήποτε στιγμή και το χαμόγελο της.

Τέλος, αφιερώνω αυτή την εργασία στην οικογένεια μου και τους ευχαριστώ πάρα πολύ που βρίσκονται πάντα δίπλα μου με αγάπη.

Περιεχόμενα

1	Ανάλυση επιβίωσης (Survival analysis)	1
1.1	Βασικές έννοιες	2
1.1.1	Συνάρτηση επιβίωσης	2
1.1.2	Συνάρτηση κινδύνου	4
1.1.3	Πιθανότητες μετάβασης	6
1.1.4	Ομογενείς Μαρκοβιανές διαδικασίες	8
1.1.5	Μη ομογενείς Μαρκοβιανές διαδικασίες	8
1.2	Λογοκριμένα δεδομένα	9
1.3	Μη παραμετρικές μέθοδοι	12
1.3.1	Εκτιμητής Kaplan-Meier	13
1.3.2	Εκτιμητής Nelson-Aalen	15
1.3.3	Εκτιμητής Aalen-Johansen	16
1.3.4	Σχέση μεταξύ των εκτιμητών	16
1.4	Παραμετρικές μέθοδοι	17
1.4.1	Εκθετική κατανομή (exponential distribution)	17
1.4.2	Weibull κατανομή	17
1.4.3	Λογαριθμοκανονική κατανομή (Log-normal distribution)	19
1.4.4	Log-Λογιστική κατανομή (Log-Logistic κατανομή)	20
1.4.5	Pareto κατανομή	20
1.4.6	Inverse-Gaussian κατανομή	21
1.5	Εκτίμηση πιθανοφάνειας χρόνων επιβίωσης	21
1.6	Ημιπαραμετρικές μέθοδοι	23
1.6.1	Μοντέλο παλινδρόμησης ή αναλογικού κινδύνου του Cox	23
1.7	Συμπερασματολογία	28

1.7.1	Έλεγχος λόγου πιθανοφάνειας (Likelihood ratio test)	28
1.7.2	Έλεγχος Wald (Wald test)	29
2	Ανταγωνιστικοί κίνδυνοι (Competing risks)	30
2.1	Βασικές συναρτήσεις πιθανοτήτων	31
2.2	Λανθάνοντες χρόνοι αποτυχίας (Latent failure times)	33
2.3	Το προσδιοριστικό ζήτημα (The identifiability issue)	34
2.4	Μη παραμετρική εκτίμηση	35
3	Μοντέλα πολλαπλών καταστάσεων (Multi state models)	37
3.1	Επιλογή δομής καταστάσεων	41
3.1.1	Μοντέλο γονιμότητας	42
3.1.2	Μοντέλο ανικανότητας (Disability model)	44
3.1.3	Διμεταβλητά δεδομένα	46
3.1.4	Επαναλαμβανόμενα γεγονότα	49
3.2	Προοδευτικά μοντέλα	49
3.2.1	Μοντέλο ανταγωνιστικών κινδύνων	50
3.2.2	Μοντέλο ανικανότητας	50
3.3	Επεκτεταμένα Μαρκοβιανά μοντέλα	52
3.4	Εκτίμηση Μαρκοβιανής συναρτήσης πιθανοφάνειας	53
3.5	Μοντέλα παλινδρόμησης πολλαπλών καταστάσεων	54
4	Εφαρμογή: Ανάλυση δεδομένων από ασθενείς με λοίμωξη σε ΜΕΘ	56
4.1	Εισαγωγή	56
4.2	Περιγραφική ανάλυση	58
4.3	Επιλογή κατάλληλων στατιστικών μεθόδων	63
4.4	Στατιστική ανάλυση	65
4.5	Αποτελέσματα	70
4.6	Συζήτηση	78
Α'	Ονομασία και τιμές μεταβλητών	81

Β' Κώδικας R	101
B.1 MSM	101
B.2 Cox regression	106
B.3 ChangeLOS-etm	107
B'.4 Σχήματα	109
Βιβλιογραφία	112

Κεφάλαιο 1

Ανάλυση επιβίωσης (Survival analysis)

Η ανάλυση επιβίωσης (survival analysis) είναι ένας κλάδος της στατιστικής, ο οποίος δημιουργήθηκε για την ανάλυση δεδομένων τα οποία δεν μπορούν να επεξεργαστούν από τις συνηθισμένες στατιστικές κατανομές. Τα δεδομένα αυτά δίνουν τη χρονική διάρκεια μέχρι να γίνει ένα συγκεκριμένο γεγονός.

Ο χρόνος επιβίωσης ή χρόνος αποτυχίας (survival time ή failure time), αναφέρεται σε μια μεταβλητή που μετράει το χρόνο (ώρες, ημέρες, εβδομάδες, μήνες, κλπ) που μεσολαβεί από τη στιγμή της έναρξης της παρακολούθησης ενός ατόμου (άνθρωπος, αντικείμενο, φαινόμενο κτλ.), μέχρι τη στιγμή που το άτομο θα αντιμετωπίσει το καταληκτικό γεγονός.

Ιστορικά οι πρώτες περιπτώσεις εφαρμογής της ανάλυσης επιβίωσης αποδίδονται στον D.Bernoulli 1700-1782 όπου χρησιμοποιώντας δεδομένα θνησιμότητας από ευλογία έκανε επίδειξη της αποτελεσματικότητας των εμβολίων. Άλλα αρχικά παραδείγματα εμφανίζονται τον 17ο αιώνα με τις εργασίες του Graunt για την κατανόηση της διάρκειας της ανθρώπινης ζωής με την χρησιμοποίηση διαγραμμάτων ζωής. Κατά την διάρκεια του 2ου Παγκοσμίου Πολέμου, ορισμένες αρχικές αναλύσεις πινάκων ζωής οδήγησαν στην μελέτη και ανάλυση αξιοπιστίας των όπλων και των πολεμικών εξοπλισμών. Αργότερα, η ανάλυση αναφερόταν στον χρόνο μεταξύ της θεραπείας μέχρι τον θάνατο και γι αυτό πήρε αυτό το όνομα. Η ανάλυση επιβίωσης όμως μπορεί να εφαρμοστεί σε αρκετές περιπτώσεις, όπως για παράδειγμα στη μηχανολογία, για

την ανάλυση του χρόνου μέχρι την εμπλοκή ενός μηχανήματος ή τη γεωργία για την ανάλυση του χρόνου μέχρι την στιγμή να βγάλει καρπό ένα δέντρο.

Μολονότι τα δεδομένα είναι ποσοτικά, οι κατανομές των χρόνων επιβίωσης σε ελάχιστες μόνο περιπτώσεις ακολουθούν την κανονική κατανομή, τείνοντας να έχουν δεξιά ασυμμετρία. Κι αυτός είναι ένας από τους λόγους που χρησιμοποιούνται διαφορετικές στατιστικές κατανομές από τις συνηθισμένες. Επίσης, εφ'όσον ο χρόνος επιβίωσης ή χρόνος αποτυχίας αναφέρεται σε μια μεταβλητή που μετράει το χρόνο, οι κατανομές που χρησιμοποιούνται είναι κατανομές που παίρνουν θετικές τιμές, όπως η εκθετική και η *inverse Gaussian*.

Στις περισσότερες ιατρικές έρευνες, η ανάλυση επιβίωσης μελετά τον χρόνο ως την υποτροπή ή τον θάνατο ενός ασθενούς. Είναι προφανές πως σε αυτές τις περιπτώσεις είναι ίσως αδύνατο για τον ερευνητή να περιμένει όλο το χρονικό διάστημα που απαιτείται ώστε να πεθάνουν όλοι οι ασθενείς. Το γεγονός πως το αποτέλεσμα της έρευνας είναι κατά ένα μέρος διαθέσιμο συνιστά τη σημαντικότερη διαφορά της ανάλυσης επιβίωσης από τις υπόλοιπες στατιστικές μεθόδους.

Ο στατιστικός συνήθως θέλει να εκτιμήσει πόσο χρόνο μπορεί να επιβιώσει ένας ασθενής, δεδομένων κάποιων συγκεκριμένων χαρακτηριστικών που μπορεί να αφορούν τον χρόνο που έχει ήδη επιβιώσει, κάποια κλινικά και γονιδιακά χαρακτηριστικά του ασθενή. Επίσης καλείται να απαντήσει πως επηρεάζει ένα συγκεκριμένο χαρακτηριστικό τον χρόνο θανάτου ή την πιθανότητα επιβίωσης ενός ασθενούς.

Κάθε διαθέσιμη πληροφορία για τους ασθενείς πρέπει να συμπεριλαμβάνεται στο κατάλληλο στατιστικό μοντέλο για την καλύτερη περιγραφή του τρόπου με τον οποίο οι παράγοντες αλληλεπιδρούν μεταξύ τους και για ασφαλέστερες εκτιμήσεις του χρόνου επιβίωσης.

1.1 Βασικές έννοιες

1.1.1 Συνάρτηση επιβίωσης

Η συνάρτηση επιβίωσης (*survival function*) είναι από τις πλέον βασικές έννοιες στην ανάλυση επιβίωσης επειδή οι πιθανότητες επιβίωσης για διαφορετικές τιμές του t παρέχουν ζωτικής σημασίας πληροφορίες για τα δεδομένα επιβίωσης. Έστω T ο χρόνος επιβίωσης ενός ατόμου που επειδή δηλώνει χρόνο είναι μη αρνητική τυχαία

μεταβλητή. Αν θεωρήσουμε ως γεγονός τον θάνατο, τότε η συνάρτηση επιβίωσης που συμβολίζεται με S (από το Survival) δηλώνει την πιθανότητα του ασθενούς να επιζήσει πέρα από ένα χρονικό σημείο t , και ορίζεται ως

$$S(t) = P(T > t) \quad (1.1)$$

Η συνάρτηση επιβίωσης είναι μη αρνητική και μη αύξουσα συνάρτηση του t με $S(0) = 1$ και $S(\infty) = 0$ αφού αν αυξήσουμε απεριόριστα την περίοδο μελέτης τελικά κανένας δεν θα επιζήσει κι έτσι η γραφική παράσταση της συνάρτησης επιβίωσης πέφτει τελικά στο 0. Αν ο χρόνος είναι μια συνεχής τυχαία μεταβλητή τότε η συνάρτηση επιβίωσης είναι συμπληρωματική της αθροιστικής συνάρτησης κατανομής, δηλαδή

$$S(t) = 1 - F(t) \quad (1.2)$$

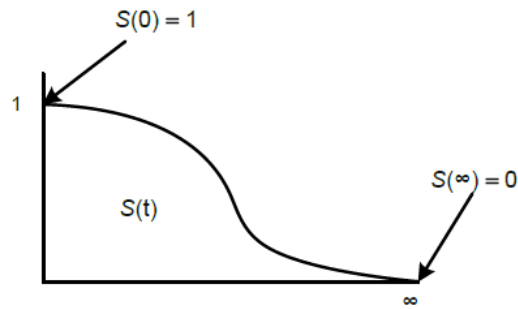
όπου $F(t) = P(T \leq t)$ δηλώνει την πιθανότητα να πεθάνει ένας ασθενής πριν τον χρόνο t . Επίσης, η συνάρτηση επιβίωσης ορίζεται ως το ολοκλήρωμα της συνάρτησης πυκνότητας πιθανότητας, $f(t)$:

$$S(t) = P(T > t) = \int_t^{\infty} f(s) ds \quad (1.3)$$

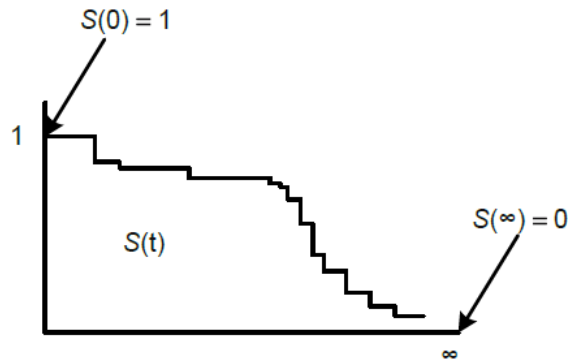
Η $f(t)$ δίνει την συνάρτηση πυκνότητας πιθανότητας να συμβεί το γεγονός την χρονική στιγμή t .

Η γραφική παράσταση της $S(t)$ συναρτήσει του t είναι γνωστή ως καμπύλη επιβίωσης (survival curve). Είναι πολύ σημαντική στην ανάλυση δεδομένων χρόνων επιβίωσης και προτάθηκε από τον Berkson το 1942. Επειδή η περίοδος μελέτης δεν είναι ποτέ άπειρη, είναι πιθανό να μην πεθάνουν (ή αποκτήσουν το γεγονός του ενδιαφέροντος) όλοι οι ασθενείς. Έτσι η εκτιμώμενη συνάρτηση επιβίωσης δεν είναι απαραίτητο ότι θα καταλήγει στο 0 στο τέλος της μελέτης. Θεωρητικά είναι μια καμπύλη που παίρνει τις τιμές από το 0 έως το ∞ όπως φαίνεται στο σχήμα (1.1).

Αλλά πρακτικά είναι μια σειρά από συναρτήσεις σκαλοπάτια λόγω του ότι οι χρόνοι επιβίωσης ποτέ δεν είναι τόσο πυκνά κατανεμημένοι στον άξονα του χρόνου. Επίσης η τιμή της $S(t)$ αλλάζει μόνο στα σημεία όπου παρατηρούνται θάνατοι ή γενικότερα αποτυχίες. (Σχήμα (1.2))



Σχήμα 1.1:



Σχήμα 1.2:

1.1.2 Συνάρτηση κινδύνου

Η συνάρτηση κινδύνου $\lambda(t)$ ορίζεται ως η πιθανότητα αποβίωσης (ή γενικότερα πραγματοποίησης του γεγονότος που εξετάζεται) την χρονική στιγμή t , δεδομένου ότι το άτομο έχει επιβιώσει μέχρι την χρονική στιγμή t . Ο όρος αυτός χρησιμοποιήθηκε για πρώτη φορά από τον Barlow το 1963. Δηλαδή,

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T \leq t + dt | T \geq t)}{dt} \quad (1.4)$$

Και από την παραπάνω σχέση μπορεί να προκύψει ότι

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T \leq t + dt, T \geq t)}{dtP(T \geq t)} = \lim_{dt \rightarrow 0} \frac{P(t \leq T \leq t + dt)}{dtP(T \geq t)} = \frac{f(t)}{S(t)} \quad (1.5)$$

Η αθροιστική συνάρτηση κινδύνου (για τη συνεχή περίπτωση) είναι:

$$\begin{aligned} \Lambda(t) &= \int_0^t \lambda(s) ds = \int_0^t \frac{f(s)}{S(s)} ds = \int_0^t \frac{f(s)}{1 - F(s)} ds = - \int_0^t \frac{d}{ds} \ln(1 - F(s)) ds \\ &= -\ln(1 - F(t)) = -\ln(S(t)) \Rightarrow \Lambda(t) = -\ln(S(t)) \end{aligned} \quad (1.6)$$

Έτσι προκύπτει και η παρακάτω σχέση που συνδέει και αυτή την συνάρτηση επιβίωσης με την συνάρτηση κινδύνου:

$$S(t) = \exp\left[-\int_0^t \lambda(u) du\right] = e^{-\Lambda(t)} \quad (1.7)$$

Αν γνωρίζουμε κάποια από τις $\lambda(t)$, $S(t)$ προκύπτει και η άλλη άμεσα.

Ως συνάρτηση το $\lambda(t)$ μπορεί να πάρει οποιαδήποτε μορφή, με τον μόνο περιορισμό ότι οι τιμές είναι μη αρνητικές για όλες τις τιμές του t .

Η συνάρτηση κινδύνου είναι γνωστή και ως στιγμιαίος λόγος αποτυχίας (instantaneous failure rate) ή ως δεσμευμένος λόγος θνησιμότητας (conditional mortality).

Κάποια σχόλια για την συνάρτηση κινδύνου είναι τα εξής:

- Δεν δίνει κάποια πιθανότητα. Η τιμή που λαμβάνεται δίνει έναν διαφορετικό αριθμό αναλόγως με τις μονάδες χρόνου, και μπορεί να δώσει ακόμη και έναν αριθμό μεγαλύτερο της μονάδας.
- Είναι ένα στιγμιαίο μέτρο ενώ η συνάρτηση επιβίωσης είναι ένα αθροιστικό μέτρο ως προς τον χρόνο.
- Μπορεί να χρησιμοποιηθεί για να προσδιορίσει ένα συγκεκριμένο μοντέλο, όπως είναι η εκθετική, μια Weibull, μια λογαριθμική καμπύλη που προσαρμόζεται στα δεδομένα ενός ατόμου.

1.1.3 Πιθανότητες μετάβασης

Οι πιθανότητες μετάβασης στην απλούστερη τους μορφή ορίζονται ως:

$$P_l(t) = P(X_t = l), \quad (1.8)$$

η πιθανότητα της διαδικασίας X να είναι στην κατάσταση l την χρονική στιγμή t .

Υπενθύμιση: Έστω ο χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, όπου Ω ο δειγματικός χώρος, $\mathcal{A}$ το σύνολο όλων των πιθανών γεγονότων (η σ-άλγεβρα) και $\mathbb{P}$ μέτρο πιθανότητας στον Ω . Η οικογένεια $X = \{X(t) : t \geq 0\}$ των τυχαίων μεταβλητών $X(t) : \Omega \rightarrow \mathbf{S}$ καλείται μια συνεχούς χρόνου διαδικασία πάνω στον χώρο καταστάσεων $\mathbf{S}$. Η στοχαστική διαδικασία X_t που περιγράφει την πορεία των μοντέλων πολλαπλών καταστάσεων ορίζεται ως l , αν η διαδικασία είναι στην κατάσταση l την χρονική στιγμή t .

Αν δεν ξεκινούν όλες οι διαδικασίες από την ίδια κατάσταση τότε η παράπάνω έκφραση θα πρέπει να εξαρτάται από την αρχική κατάσταση. Οι πιθανότητες μετάβασης εκτιμώμενες την χρονική στιγμή u ορίζονται ως:

$$P_l(s, t) = P(X_t = l | X_u, u \in [0, s]), \quad (1.9)$$

οι οποίες είναι δεσμευμένες ως προς όλη την εξέλιξη της διαδικασίας έως την χρονική στιγμή s . Προφανώς, η παραπάνω σχέση ισχύει μόνο όταν $t \geq s$. Οι πιθανότητες μετάβασης είναι αρκετά απλές για τις Μαρκοβιανές διαδικασίες, αλλά μπορούν να χρησιμοποιηθούν και σε μη μαρκοβιανές διαδικασίες. Η Μαρκοβιανή ιδιότητα μπορεί να γραφεί ως

$$P(X_t = l | X_u, u \in [0, s]) = P(X_t = l | X_s) \quad (1.10)$$

για $s \leq t$. Δηλαδή η πορεία της διαδικασίας από το 0 έως το t συνοψίζεται πλήρως από την κατάσταση στην οποία βρίσκεται την χρονική στιγμή t , ή αλλιώς η μελλοντική πιθανοθεωρητική συμπεριφορά της διαδικασίας εξαρτάται μόνο από την παρούσα κατάσταση της και όχι από την ιστορία της. Γίνεται πιο σαφές αν γραφούν οι Μαρκοβιανές πιθανότητες μετάβασης ως:

$$P_{ml}(s, t) = P(X_t = l | X_s = m), \quad (1.11)$$

όπου $s \leq t$.

Με την έννοια ιστορία (ή παρελθόν) την χρονική στιγμή t , δηλώνεται η πληροφορία που περιέχεται στην εξέλιξη της διαδικασίας μέχρι εκείνη την στιγμή, δηλαδή στο διάστημα $[0, t]$. Δηλαδή, η $X_s, 0 \leq s \leq t$, η οποία μπορεί να εξηγηθεί μαθηματικά ακριβώς μια σ-άλγεβρα που δημιουργείται αποτελούμενη από την παρατήρηση της διαδικασίας στο διάστημα $[0, t]$, όπως οι καταστάσεις που προηγουμένως επισκέφθηκε, οι χρόνοι των μεταβάσεων κτλ.

Γενικά οι πιθανότητες μετάβασης ικανοποιούν τις εξής σχέσεις:

$$P_{ml} \geq 0 \text{ και } \sum_l P_{ml} = 1. \quad (1.12)$$

Οι πιθανότητες μετάβασης πρώτης τάξης (όπως οι παραπάνω) μπορούν να χρησιμοποιηθούν για να υπολογιστεί η πιθανότητα να μεταβεί η ανάλυση από την κατάσταση m στην κατάσταση l μετά από n βήματα. Οι πιθανότητες μετάβασης n τάξης ορίζονται ως εξής:

$$P_{ml}^{(n)} = P(X_n = l | X_0 = m), n = 1, 2, \dots \quad (1.13)$$

Οι ακόλουθες εξισώσεις Chapman-Kolmogorov παρέχουν μια μέθοδο υπολογισμού των πιθανοτήτων μετάβασης n τάξης $P_{ml}^{(n)}$ μίας Μαρκοβιανής αλυσίδας.

$$P_{ml}(s, t) = \sum_r P_{mr}(s, u) P_{rl}(u, t), \quad (1.14)$$

όπου $s \leq u \leq t$. Αυτές οι εξισώσεις έχουν ισχυρή ερμηνεία αφού μπορούν να χωριστούν με βάση τους ενδιάμεσους χρόνους και τις καταστάσεις.

Οι πιθανότητες μετάβασης πρώτης τάξης συγκεντρώνονται σε ένα πίνακα που καλείται πίνακας μετάβασης πρώτης τάξης της διαδικασίας (one-step transition probability matrix). Ο πίνακας αυτός συνήθως συμβολίζεται με P και επειδή ικανοποιεί τις σχέσεις (3.6) καλείται στοχαστικός πίνακας (stochastic matrix). Το διάνυσμα-γραμμή $p^{(0)}$ με στοιχεία $p_j^{(0)} = a_j (j = 1, 2, \dots)$ ονομάζεται κατανομή αρχικών καταστάσεων και οι τετραγωνικοί πίνακες $P^{(n)}$ με στοιχεία $p_{ij}^{(n)} (i, j = 1, 2, \dots \text{ και } n = 1, 2, \dots)$ είναι οι Πίνακες Πιθανοτήτων Μετάβασης, οι οποίοι είναι μη αρνητικοί και το άθροισμα των γραμμών τους είναι ίσο με τη μονάδα.

Οι πιθανότητες μετάβασης μπορούν να προσδιορισθούν από τους κινδύνους με-

τάβασης μέσω του product integral (Grill and Johansen, 1990). Έτσι ο πίνακας πιθανοτήτων μετάβασης $P(s, t)$ με στοιχεία $\{P_{jh}(s, t)\}$ για $s < t$ είναι:

$$P(s, t) = \prod_{u \in (s, t)} (I + d\Lambda(u)) \quad (1.15)$$

όπου Λ είναι ο πίνακας των κινδύνων μετάβασης με στοιχεία $\{\Lambda_{jh}\}$.

1.1.4 Ομογενείς Μαρκοβιανές διαδικασίες

Η Μαρκοβιανή υπόθεση δηλώνει ότι το ανάπτυγμα της διαδικασίας μετά από κάποιο χρόνο t δεδομένης όλη της προϊστορίας εξαρτάται μόνο από την κατάσταση στην οποία βρίσκεται την χρονική στιγμή t . Με άλλα λόγια, για να γίνει μια πρόβλεψη την στιγμή t καμιά περαιτέρω πληροφορία για το πως εξελίχθηκε η διαδικασία μέχρι την στιγμή t δεν χρειάζεται.

Σε ένα Μαρκοβιανό μοντέλο οι συναρτήσεις κινδύνου είναι ανεξάρτητες από τον χρόνο παραμονής στην τρέχουσα κατάσταση. Όταν οι συναρτήσεις κινδύνου δεν εξαρτώνται γενικά από το χρόνο t (σταθεροί κίνδυνοι), η Μαρκοβιανή διαδικασία καλείται ομογενής.

Κάποιες φορές οι ομογενείς Μαρκοβιανές διαδικασίες ονομάζονται απλά «Μαρκοβιανές διαδικασίες», πετώντας το χαρακτηριστικό «ομογενείς». Πρόκειται για την πιο απλή περίπτωση και μια πολύ γνωστή κλάση μοντέλων. Οι πιθανότητες μετάβασης δεν εξαρτώνται από τους πραγματικούς χρόνους των γεγονότων αλλά μόνο από τη διάρκεια του χρονικού διαστήματος που χρειάστηκε για να πραγματοποιηθεί η μετάβαση. Συνεπώς οι πιθανότητες μετάβασης είναι χρονικά αναλλοίωτες. Οι χρονικά εξαρτημένες συμμεταβλητές δεν επιτρέπονται. Δηλαδή

$$P\{X_t = j | X_{t-1} = i\} = p_{ij} \quad (1.16)$$

για κάθε t .

1.1.5 Μη ομογενείς Μαρκοβιανές διαδικασίες

Σε κάποιες εφαρμογές, η υπόθεση της ομογένειας μπορεί να είναι μη ρεαλιστική και να χρειάζεται ένα μη ομογενές μοντέλο.

Οι πιθανότητες μετάβασης

$$P_{ij}(s, t) = P(X_t = j | X_s = i) = P(X_t = j | X_s = i, \text{παρελθόν}), s \leq t, \quad (1.17)$$

εξαρτώνται από το πραγματικό χρονικό διάστημα $[s, t]$ σε αντίθεση με τις ομογενείς διαδικασίες. Οι κίνδυνοι μετάβασης της Μαρκοβιανής διαδικασίας είναι συναρτήσεις του χρόνου όπως φαίνεται παρακάτω

$$\alpha_{ij}(t)dt = P(X_{(t+dt)-} = j | X_{t-} = i), i, j = 0, \dots, J, i \neq j, \quad (1.18)$$

όπου υποθέτουμε ότι τα όρια υπάρχουν και $X_{t-} = \lim_{s \nearrow t} X_s$. Η Μαρκοβιανή ιδιότητα συνεπάγεται ότι η υπόθεση $X_{t-} = i$ είναι ισοδύναμη με την υπόθεση ότι ολόκληρο το παρελθόν της διαδικασίας βρίσκεται πριν από την στιγμή t . Ορίζεται

$$\alpha_{ii}(t) = - \sum_{j=0, j \neq i}^J \alpha_{ij}(t), i = 0, \dots, J, \quad (1.19)$$

και οι αθροιστικές ποσότητες

$$A_{ij}(t) = \int_0^t \alpha_{ij}(u)du, i, j = 0, \dots, J. \quad (1.20)$$

Στην ομογενή αλυσίδα έχουμε ένα μόνο πίνακα μετάβασης, ενώ στην περίπτωση της μη ομογενούς αλυσίδας έχουμε μια ακολουθία πινάκων μετάβασης.

Οι μη ομογενείς Μαρκοβιανές διαδικασίες μπορούν επίσης να μοντελοποιηθούν μη παραμετρικά. Μέσω της γενίκευσης της προσέγγισης Kaplan-Meier για το απλό μοντέλο θνητότητας και οι Aalen και Johansen πρότειναν τη μη παραμετρική μοντελοποίηση για γενικά μοντέλα πολλαπλών καταστάσεων με πεπερασμένο αριθμό καταστάσεων.

1.2 Λογοκριμένα δεδομένα

Οι παρατηρήσεις για τις οποίες δεν είναι γνωστό το τελικό αποτέλεσμα ονομάζονται λογοκριμένες παρατηρήσεις (censored cases). Ο όρος censoring χρησιμοποιήθηκε για πρώτη φορά από τον Hald (1949). Τα δεδομένα που δεν είναι λογοκριμένα ονο-

μάζονται μη-λογοκριμένα ή πλήρη δεδομένα. Οι αιτίες που τελικά δεν παρατηρείται το τελικό γεγονός μπορεί να είναι διαφορετικές. Η βασικότερη είναι ότι μια μελέτη που ξεκινά για την αποτίμηση μιας μεθόδου έχει ένα συγκεκριμένο χρονικό ορίζοντα. Με το πέρας της μελέτης οι ασθενείς που δεν έχουν πεθάνει θεωρούνται λογοκριμένοι εφόσον θα πεθάνουν στο μέλλον αλλά δεν γνωρίζουμε την ακριβή χρονική στιγμή. Μια άλλη αιτία είναι οι περιπτώσεις ασθενών που δεν έχουμε πλέον πληροφορίες για αυτούς είτε επειδή άλλαξαν θεράποντα ιατρό είτε αρνήθηκαν την περαιτέρω συμμετοχή τους στην έρευνα είτε για κάποιον άλλο λόγο κάτι που είναι σύνηθες στις ιατρικές έρευνες.

Παρ'όλο που οι λογοκριμένες παρατηρήσεις είναι ελλιπείς μπορούν να συμβάλουν στην καλύτερη μελέτη της έρευνας. Αν ένας ασθενής καταφέρει να επιβιώσει για μερικά χρόνια χωρίς να εμφανίσει την ασθένεια, τότε ο χρόνος αυτός της επιβίωσης θα πρέπει να συμπεριληφθεί στην μελέτη αντί να αγνοηθεί. Άρα όλες οι πληροφορίες μέχρι την στιγμή που συναντάμε λογοκρισία πρέπει να συμπεριλαμβάνονται και να χρησιμοποιούνται και όχι να αγνοούνται.

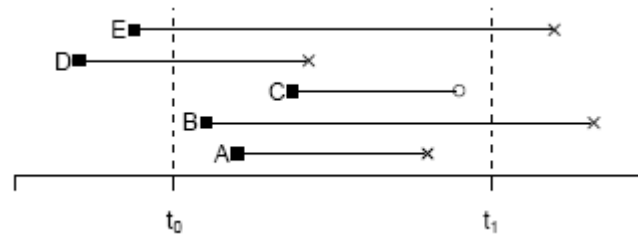
Υπάρχουν 3 είδη λογοκρισίας. Η δεξιά λογοκρισία (right censoring), η αριστερή λογοκρισία (left censoring) και η λογοκρισία διαστήματος (interval censoring).

- **Δεξιά λογοκρισία:** Ο χρόνος επιβίωσης είναι μεγαλύτερος από τον χρόνο στον οποίο σταματά η μελέτη. Είναι η πιο συνηθισμένη μορφή λογοκρισίας. Παρατηρείται σε περιπτώσεις όπου ένα άτομο χάνεται ή αποσύρεται από την παρακολούθηση, ή ακόμη όταν η μελέτη τερματίζεται σε ένα προκαθορισμένο χρόνο.
- **Αριστερή λογοκρισία:** Ο χρόνος επιβίωσης/χρόνος αποτυχίας σε αυτήν την περίπτωση συμβαίνει πριν την αρχή της παρατηρούμενης περιόδου. Έτσι ο ακριβής χρόνος επιβίωσης δεν είναι γνωστός. Το μόνο που είναι γνωστό, είναι ότι ο χρόνος επιβίωσης, T , είναι μικρότερος από ένα χρονικό διάστημα. Για παράδειγμα, η εισαγωγή των ασθενών με αρθρίτιδα σε μια ειδική κλινική. Οι περισσότεροι από τους ασθενείς εισάγονται στην κλινική μετά την έναρξη/εμφάνιση της αρθρίτιδας.
- **Λογοκρισία διαστήματος:** Παρατηρείται όταν γνωρίζουμε ότι το υπό μελέτη συμβάν έχει πραγματοποιηθεί σε ένα διάστημα (π.χ. μεταξύ των επισκέψεων) και πάλι όμως χωρίς να είναι γνωστό το ακριβές σημείο. Αυτό συμβαίνει συνήθως όταν έχουμε περιοδική παρακολούθηση του ασθενή. Για παράδειγμα, αν

παρακολουθείται μια ομάδα ατόμων για τυχόν επανεμφάνιση της ασθένειας και παρατηρηθεί ότι κάποιος ασθενής την εμφάνισε μεταξύ του 2ου και 3ου μήνα που εξετάσθηκε.

Στο σχήμα 1.3 φαίνονται οι χρόνοι μέχρι την αποτυχία πέντε ασθενών.

Ο ασθενής A αποτυγχάνει μέσα στην παρατηρούμενη χρονική περίοδο $[t_0, t_1]$ (μη



Σχήμα 1.3: Οι συμπαγείς γραμμές αντιπροσωπεύουν την περίοδο που βρίσκονται σε κίνδυνο οι ασθενείς, το "X" δηλώνει την εμφάνιση του ενδιαφέροντος γεγονότος και το "o" την εμφάνιση ενός γεγονότος διαφορετικού από το ενδιαφερόν(λογοκρισία)

λογοκριμένη-πλήρης παρατήρηση) ενώ ο ασθενής B αποτυγχάνει μετά το τέλος της παρακολούθησης t_1 και άρα λογοκρίνεται δεξιά την στιγμή t_1 . Ο ασθενής C επίσης είναι δεξιά λογοκριμένος, αλλά επειδή του συνέβη ένα διαφορετικό γεγονός από το γεγονός του ενδιαφέροντος. Το γεγονός που λογίζεται ως η αρχή της περιόδου κινδύνου του ασθενή D συνέβη πριν από την αρχή της μελέτης και έτσι θεωρείται αριστερά λογοκριμένος (ή περικομμένος-truncated). Τέλος ο ασθενής E λογοκρίνεται και αριστερά και δεξιά (λογοκρισία διαστήματος).

Συνήθως με «+» δίπλα από ένα χρόνο συμβολίζονται οι λογοκριμένες παρατηρήσεις.

Θα δούμε αργότερα ότι η ανεξαρτησία μεταξύ των χρόνων του γεγονότος και των λογοκριμένων χρόνων θεωρείται δεδομένη χωρίς ιδιαίτερες σκέψεις, αλλά αρκετές φορές δεν ισχύει.

Εάν η λογοκρισία οφείλεται στο χρονικό τέλος της έρευνας τότε γενικά μπορούμε να υποθέσουμε ότι ο μηχανισμός της λογοκρισίας είναι ανεξάρτητος με την πρόοδο του γεγονότος, π.χ. ασθένειας. Όμως εάν η λογοκρισία οφείλεται στο ότι τα άτομα χάθηκαν από την παρακολούθηση του ερευνητή τότε η λογοκρισία είναι αρνητικά

συσχετισμένη με τους χρόνους των γεγονότων και προκαλείται μεροληψία προς τα κάτω ως προς την εκτιμώμενη καμπύλη επιβίωσης. Ενώ οι λογοκριμένοι χρόνοι είναι θετικά συσχετισμένοι με τους χρόνους των γεγονότων όταν τα άτομα που βρίσκονται σε προχωρημένο στάδιο ασθένειας είναι πιθανότερο να φύγουν από τη μελέτη και αιτία γι αυτή τους την φυγή μπορεί να είναι π.χ. η επιστροφή τους στον τόπο καταγωγής τους ώστε να περάσουν το τέλος της ζωής τους εκεί. Αυτή η λογοκρισία προκαλεί μεροληψία προς τα πάνω ως προς την εκτιμώμενη καμπύλη επιβίωσης.

1.3 Μη παραμετρικές μέθοδοι

Τα περισσότερα γνωστά παραμετρικά μοντέλα που υπάρχουν στην στατιστική δεν είναι κατάλληλα για να εφαρμοσθούν στην ανάλυση επιβίωσης. Και για την αποφυγή της παραδοχής ότι κάποια παραμετρικά μοντέλα που θα δούμε αργότερα είναι σωστά, είναι σύνηθες να χρησιμοποιούνται μη παραμετρικές μέθοδοι. Ένα μεγάλο πλεονέκτημα των μη παραμετρικών μεθόδων είναι ότι παρουσιάζουν τα δεδομένα χωρίς να έχει προηγηθεί κατ'ανάγκην υπόθεση ότι ακολουθούν κάποιο μοντέλο.

Η πιο απλή μη παραμετρική εκτίμηση μιας συνάρτησης κατανομής είναι η εμπειρική κατανομή, όπου η κατανομή θεωρείται ίση με την παρατηρούμενη κατανομή. Αυτή είναι και η μη παραμετρική εκτίμηση μέγιστης πιθανοφάνειας για πλήρεις παρατηρήσεις. Έτσι, ακόμα κι αν υποθέσουμε ότι η αληθής κατανομή είναι συνεχής, αυτή εκτιμάται από μια διακριτή κατανομή.

Ένα βασικό ερώτημα είναι αν θα πρέπει να εφαρμόζεται ένα παραμετρικό ή ένα μη παραμετρικό μοντέλο. Το κύριο προσόν των μη παραμετρικών μοντέλων είναι η προσαρμογή τους αφού μπορούν να χειριστούν οποιαδήποτε κατανομή αλλά υπάρχει ένα μεγάλο κόστος γι αυτή την καλή προσαρμογή. Είναι πολύ πιο δύσκολο να εκτεθεί μια μη παραμετρική εκτίμηση εφ'όσον τα αποτελέσματα πρέπει να εμφανίζονται μέσω ενός σχήματος ή ενός πίνακα, ενώ ένα παραμετρικό μοντέλο μπορεί να περιγραφεί μέσω των τιμών λίγων παραμέτρων. Κύριο μειονέκτημα των μη παραμετρικών διαδικασιών είναι ότι η συνάρτηση κινδύνου δεν μπορεί να εκτιμηθεί, ούτε καν να οριστεί αν η κατανομή είναι διακριτή. Για ορισμένους σκοπούς, η συνάρτηση κινδύνου είναι πολύ πιο ενδιαφέρουσα και σχετική να εκτιμηθεί από τη συνάρτηση επιβίωσης και την ολοκληρωμένη συνάρτηση κινδύνου, αλλά είναι επίσης και πιο δύσκολο να εκτιμηθεί.

1.3.1 Εκτιμητής Kaplan-Meier

Όταν τα δεδομένα δεν είναι λογοκριμένα η πιο προφανής εκτίμηση θα ήταν μέσω της εμπειρικής συνάρτησης κατανομής

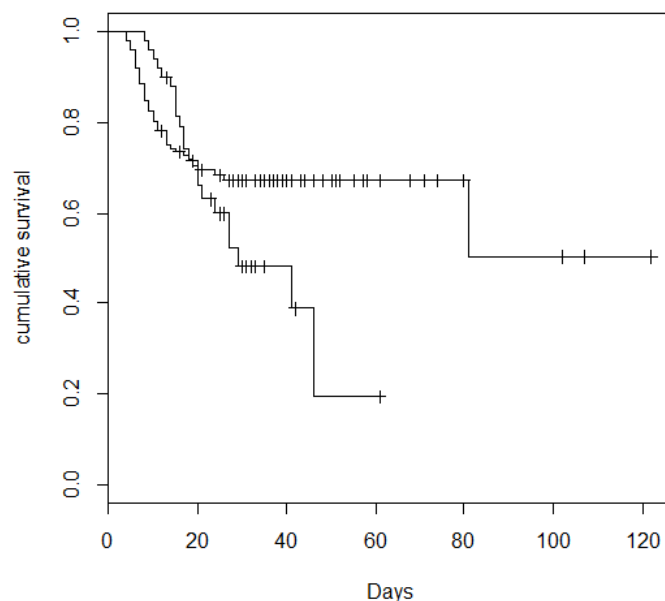
$$\hat{S}(t) = \frac{1}{n} \sum_{i=1}^n I\{t_i > t\} \quad (1.21)$$

Για τα λογοκριμένα δεδομένα όμως προτάθηκε από τους Kaplan και Meier (1958) η μέθοδος στην οποία δόθηκε το όνομα τους και ο εκτιμητής της $S(t)$ που προκύπτει, ονομάζεται εκτιμητής Kaplan-Meier αντίστοιχα. Ο εκτιμητής ονομάζεται και εκτιμητής γινομένου-ορίου (Product-Limit method (PL)) γιατί προκύπτει ως το γινόμενο πιθανοτήτων μετάβασης. Ο σκοπός αυτού του εκτιμητή είναι να εκτιμηθεί μια συνάρτηση επιβίωσης μη παραμετρικά κατά την παρουσία δεξιάς λογοκρισίας ή όταν ζητείται συγκεκριμένο αποτέλεσμα π.χ. να βρεθεί πόσο χρόνο παίρνει σε ένα ζευγάρι ώστε να συλλάβει. Συχνά χρησιμοποιείται ώστε να υπολογιστεί το κλάσμα των ασθενών που ζουν για ένα ορισμένο διάστημα του χρόνου μετά τη θεραπεία. Ένα μεγάλο πλεονέκτημα του είναι ότι δεν χρειάζεται να υποτεθεί κάποιο μοντέλο για τα δεδομένα και γι αυτό μπορεί να χρησιμοποιηθεί σχεδόν για οποιοδήποτε σύνολο δεδομένων Στην οικονομία μπορεί να χρησιμοποιηθεί για τη μέτρηση του χρόνου που οι άνθρωποι παραμένουν άνεργοι. Και στη μηχανική, για να μετρηθεί ο χρόνος μέχρι την αποτυχία κάποιων μερών της μηχανής.

Υπάρχουν τρεις παραδοχές για τη μέθοδο Kaplan-Meier.

1. Τα άτομα που λογοκρίνονται, έχουν την ίδια πιθανότητα επιβίωσης με τα άτομα που συνεχίζουν να παρακολουθούνται.
2. Οι πιθανότητες επιβίωσης είναι οι ίδιες για άτομα που εισήλθαν στην αρχή της μελέτης με των ατόμων που εισήλθαν πιο αργά στη μελέτη.
3. Το γεγονός που μελετάται (π.χ. θάνατος) συμβαίνει στον καθορισμένο χρόνο.

Η γραφική παράσταση του εκτιμητή (Σχήμα 1.4) είναι μια φθίνουσα βηματική συνάρτηση, η οποία όταν λαμβάνεται ένα αρκετά μεγάλο δείγμα, προσεγγίζει την πραγματική συνάρτηση επιβίωσης για τον πληθυσμό αυτό. Η τιμή της συνάρτησης



Σχήμα 1.4: εκτιμητής Kaplan-Meier

επιβίωσης μεταξύ διαδοχικών ξεχωριστών παρατηρήσεων του δείγματος υποτίθεται ότι είναι σταθερή.

Έστω $t_1, t_2, \dots, t_n$ οι ταξινομημένοι χρόνοι αποτυχίας και n_j ο αριθμός των ατόμων που ζουν (ή δεν έχουν λογοκριθεί ακόμα) μέχρι την στιγμή t_j , συμπεριλαμβανομένου και αυτών που πεθαίνουν την στιγμή t_j και d_j ο αριθμός των θανάτων την στιγμή t_j . Υπό αυτήν τη δομή, κάθε διάστημα ξεκινά σε έναν παρατηρούμενο χρόνο αποτυχίας, εκτός από το πρώτο το οποίο αρχίζει από την αρχή του χρόνου, και περιέχει μόνο μία αποτυχία. Μπορούν να παρατηρηθούν πάνω από μία αποτυχίες σε ένα διάστημα μόνο εάν τα άτομα παρατηρούνται να αποτυγχάνουν την ίδια χρονική στιγμή, κάτι πολύ σύνηθες σε διακριτό χρόνο.

Ο εκτιμητής Kaplan-Meier της συνάρτησης επιβίωσης την στιγμή t , $t_k \leq t \leq t_{k+1}$ είναι

$$\hat{S}(t) = \prod_{j=1}^k \hat{p}_j, \quad k = 1, 2, \dots, r \quad (1.22)$$

όπου $\hat{p}_j = \frac{n_j - d_j}{n_j}$ η εκτιμώμενη πιθανότητα το άτομο να επιβιώσει πέραν του διαστήματος j .

Η διακύμανση του εκτιμητή (γνωστή και ως τύπος Greenwood) είναι

$$\hat{Var}(\hat{S}(t)) = (\hat{S}(t))^2 \sum_{j=1}^k \frac{d_j}{n_j(n_j - d_j)} \quad (1.23)$$

Στο ίδιο πλαίσιο με τον εκτιμητή Kaplan-Meier κινείται και η μέθοδος των πινάκων επιβίωσης(life tables).

1.3.2 Εκτιμητής Nelson-Aalen

Μια απλή προσέγγιση για να εκτιμηθεί η αθροιστική συνάρτηση κινδύνου $\Lambda(t)$ είναι να ληφθεί το αντίθετο του λογαρίθμου ενός εκτιμητή της $S(t)$. Μια εναλλακτική προσέγγιση όμως είναι να εκτιμηθεί μέσω του εκτιμητή Nelson-Aalen. Είναι ένας απλός σχεδόν αμερόληπτος, μη παραμετρικός εκτιμητής της αθροιστικής συνάρτησης κινδύνου για δεξιά λογοκριμένα δεδομένα που προτάθηκε από τους Nelson και Aalen σε μία εργασία το 1972. Υπολογίζεται από τη σχέση:

$$\hat{\Lambda}(t) = \sum_{t_j \leq t} \frac{d_j}{n_j} \quad (1.24)$$

Η έκφραση αυτή εκτιμά τον αναμενόμενο αριθμό των θανάτων στο διάστημα αυτό, όπου d_j είναι ο αριθμός των θανάτων και n_j ο αριθμός των εκτιθέμενων ατόμων στον κίνδυνο την συγκεκριμένη χρονική στιγμή.

Ένας άλλος λόγος που χρησιμοποιείται ο εκτιμητής των Nelson-Aalen είναι για την αξιολόγηση διαφορετικών παραμετρικών μοντέλων. Για παράδειγμα, ένα διάγραμμα του $\hat{\Lambda}(t)$ έναντι του χρόνου t θα πρέπει να είναι κατά προσέγγιση γραμμικό αν η εκθετική κατανομή με παράμετρο θ εφαρμόζει καλά στα δεδομένα.

Η θεμελίωση του εκτιμητή της συνάρτησης κινδύνου από τον Aalen στηρίχθηκε πάνω σε counting processes και martingales. Ο Aalen ήταν ο πρώτος που προσέφερε μια αυστηρή μαθηματική θεμελίωση σε θέματα ανάλυσης επιβίωσης και όλες οι εξελίξεις πάνω στον τομέα στηρίχθηκαν στην προσέγγιση του για την απόδειξη του εκτιμητή της συνάρτησης κινδύνου.

1.3.3 Εκτιμητής Aalen-Johansen

Ο εκτιμητής Aalen-Johansen είναι ο συνήθης μη παραμετρικός εκτιμητής για τον πίνακα μεταβάσεων των μη ομογενών Μαρκοβιανών διαδικασιών και την αθροιστική συνάρτηση επίπτωσης στα μοντέλα ανταγωνιστικών κινδύνων που θα δούμε αργότερα. Επίσης γενικεύει τους εκτιμητές Nelson-Aalen/Kaplan-Meier σε δεδομένα πολλαπλών καταστάσεων.

Ο Aalen-Johansen ή γινομένου ορίου εκτιμητής είναι συνεπής και σχεδόν αμερόληπτος. Ο εκτιμητής για τον πίνακα μεταβάσεων από τον χρόνο s στον t , $P(s, t)$ δίνεται από την:

$$\hat{P}(s, t) = \prod_{i=1}^m (I + \Delta \hat{A}(T_i)). \quad (1.25)$$

όπου m ο αριθμός των συνολικών ημερών στις οποίες συνέβη τουλάχιστον μία μετάβαση, T_i το άλμα του χρόνου για κάθε μετάβαση μεταξύ των χρόνων s και t , I ο ταυτοτικός πίνακας και $\Delta \hat{A}(T_i)$ ο πίνακας με το (k, l) στοιχείο $Y_k(T_j)^{-1}$ και αντίστοιχα διαγώνια στοιχεία $-Y_k(T_j)^{-1}$, όπου η μετάβαση την στιγμή T_j είναι από το k στο l , τα υπόλοιπα στοιχεία είναι μηδενικά.

1.3.4 Σχέση μεταξύ των εκτιμητών

- Ο Breslow το 1972 πρότεινε να εκτιμάται η συνάρτηση επιβίωσης ως

$$\hat{S}(t) = \exp\{-\hat{\Lambda}(t)\} \quad (1.26)$$

όπου $\hat{\Lambda}(t)$ ο εκτιμητής Nelson-Aalen του ολοκληρωμένου κινδύνου. Οι εκτιμητές Breslow και Kaplan-Meier είναι ασυμπτωτικά ισοδύναμοι, και συνήθως είναι αρκετά κοντά μεταξύ τους και συγκεκριμένα όταν ο αριθμός των θανάτων είναι μικρός σε σχέση με τον αριθμό των ατόμων που είχαν εκτεθεί στον κίνδυνο.

- Ο αρνητικός λογάριθμος του εκτιμητή Kaplan-Meier προσεγγίζει τον εκτιμητή Nelson-Aalen όπως φαίνεται παρακάτω:

$$-\log \hat{S}_{KM}(t) = -\log \left\{ \prod_{j=1}^r \left(1 - \frac{d_j}{n_j}\right) \right\} = -\sum_{t_j \leq t} \log \left(1 - \frac{d_j}{n_j}\right) \approx -\sum_{t_j \leq t} -\frac{d_j}{n_j} = \hat{\Lambda}_{NA}(t)$$

- Ομοίως και ο εκτιμητής Kaplan-Meier μπορεί να προσεγγισθεί από τον εκθετικό εκτιμητή Nelson-Aalen: $\exp(-\hat{\Lambda}_{NA}(t))$.

1.4 Παραμετρικές μέθοδοι

Το σκεπτικό για τη χρήση μη παραμετρικών μεθόδων ήταν να μη χρειάζεται να καθοριστεί πλήρως η συνάρτηση κινδύνου. Αλλά οι παραδοχές στις παραμετρικές μεθόδους που χαρακτηρίζουν πλήρως είτε την πυκνότητα είτε την συνάρτηση κινδύνου είναι επιτρεπτές, ιδιαίτέρως σε καταστάσεις όπου προηγούμενη έρευνα προτείνει μια συγκεκριμένη μορφή για τα δεδομένα. Τα παραμετρικά μοντέλα εκτιμούνται μέσω του εκτιμητή μέγιστης πιθανοφάνειας.

1.4.1 Εκθετική κατανομή (exponential distribution)

Η εκθετική κατανομή χρησιμοποιείται για να προσδιορισθεί το χρονικό διάστημα μέχρι κάποιο συγκεκριμένο συμβάν. Η χρήση της εκθετικής κατανομής για τη μοντελοποίηση δεδομένων επιβίωσης είναι αρκετά συχνή λόγω της απλότητας της. Η συνάρτηση κατανομής της εκθετικής κατανομής είναι

$$f(t) = \theta e^{-\theta t}$$

Η συνάρτηση επιβίωσης παίρνει τη μορφή $e^{-\theta t}$. Ειδικό χαρακτηριστικό της εκθετικής κατανομής είναι ότι έχει σταθερό κίνδυνο $\lambda(t) = \frac{f(t)}{S(t)} = \theta$, ανεξάρτητο δηλαδή του χρόνου (π.χ. χρόνια ασθένεια). Η εκθετική κατανομή χαρακτηρίζεται από έλλειψη μνήμης (lack of memory), δηλαδή

$$P(T \geq t + z | T \geq t) = \frac{P(T \geq t + z)}{P(T \geq t)} = \frac{e^{-\theta(t+z)}}{e^{-\theta t}} = e^{-\theta z} \quad (1.27)$$

1.4.2 Weibull κατανομή

Η κατανομή πιθανότητας που έχει ρόλο κλειδί στην ανάλυση των δεδομένων επιβίωσης είναι η Weibull κατανομή, η οποία ορίζεται ως

$$f(t) = \theta \alpha t^{\alpha-1} e^{-\theta t^\alpha} \quad (1.28)$$

για $0 \leq t < \infty$

Όταν το μοντέλο για τους χρόνους επιβίωσης προσδιορισθεί μέσω της συνάρτησης πυκνότητας πιθανότητας, τότε όλες οι αντίστοιχες συναρτήσεις μπορούν να παρθούν μέσω των εξισώσεων:

$$S(t) = 1 - \int_0^t f(u)du = e^{-\theta t^\alpha} \quad (1.29)$$

και

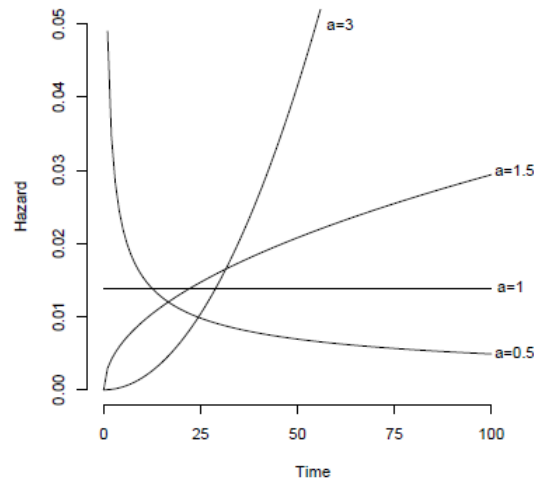
$$\lambda(t) = \frac{f(t)}{S(t)} = \theta \alpha t^{\alpha-1} \quad (1.30)$$

όπου α ονομάζεται η παράμετρος σχήματος (shape parameter) αφού το σχήμα της συνάρτησης κινδύνου εξαρτάται από αυτή την τιμή και θ η παράμετρος κλίμακας (scale parameter). Αν $\alpha = 1$ τότε

$$f(t) = \theta e^{-\theta t} \quad (1.31)$$

που είναι η συνάρτηση πυκνότητας πιθανότητας της εκθετικής κατανομής, η οποία είναι ειδική περίπτωση της Weibull κατανομής. Η Weibull κατανομή είναι πιο ευέλικτη από την εκθετική λόγω της δεύτερης παραμέτρου που διαθέτει. Για $\alpha = 2$ προκύπτει η κατανομή Rayleigh, η οποία είναι γνωστή για τη χρήση της στη μοντελοποίηση του ύψους κύματος σχετικά με τους θαλάσσιους κυματισμούς. Αν $\alpha \in (0, 1)$ η συνάρτηση κινδύνου είναι εκθετικά φθίνουσα και η συνάρτηση επιβίωσης μειώνεται πιο αργά προς το τέλος (π.χ. μετά από κάποια εγχείρηση μεταμόσχευσης, χρόνος ζωής ηλεκτρικών συσκευών), ενώ αν $\alpha > 1$ η συνάρτηση κινδύνου είναι εκθετικά αύξουσα (π.χ. για ανθρώπους μεγάλης ηλικίας).

Όμως η Weibull κατανομή δεν παρέχει πάντα το πιο ικανοποιητικό μοντέλο για τους χρόνους επιβίωσης. Ένας περιορισμός είναι ότι η Weibull συνάρτηση κινδύνου είναι μια μονότονη συνάρτηση του χρόνου ενώ υπάρχουν περιπτώσεις όπου η συνάρτηση κινδύνου μπορεί να αλλάξει διεύθυνση, π.χ. οι ασθενείς μετά από μια μεταμόσχευση καρδιάς τις πρώτες δέκα μέρες βιώνουν μεγαλύτερο κίνδυνο θανάτου ενώ αργότερα μειώνεται ο κίνδυνος με τον χρόνο.



Σχήμα 1.5: Η γενική μορφή της συνάρτησης κινδύνου Weibull για διάφορες τιμές του α

1.4.3 Λογαριθμοκανονική κατανομή (Log-normal distribution)

Σε πολλές περιπτώσεις η τυχαία μεταβλητή που εξετάζεται δεν ακολουθεί κανονική κατανομή, αλλά με έναν απλό μετασχηματισμό συχνά προκύπτει κανονική τυχαία μεταβλητή. Ένας τέτοιος μετασχηματισμός που μπορεί να προσαρμόσει τα δεδομένα σε κανονική κατανομή είναι ο λογάριθμος. Η λογαριθμοκανονική κατανομή είναι μια συνεχής κατανομή πιθανότητας μιας τυχαίας μεταβλητής της οποίας ο λογάριθμος είναι κανονικά κατανομημένος. Έτσι η κατανομή ορίζεται για τυχαίες μεταβλητές, οι οποίες παίρνουν θετικές τιμές. Η συνάρτηση πυκνότητας έχει τη μορφή

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \exp\left\{-\frac{(\log t - \mu)^2}{2\sigma^2}\right\} \quad (1.32)$$

Οι συναρτήσεις κινδύνου και επιβίωσης μπορούν να εκφραστούν μόνο μέσω ολοκληρωμάτων.

Η λογαριθμοκανονική κατανομή χρησιμοποιείται εκτενώς σε εφαρμογές αξιοπιστίας για μοντελοποίηση χρόνων αποτυχίας. Η λογαριθμοκανονική και η Weibull κατανομή είναι πιθανώς οι πιο συχνά χρησιμοποιούμενες κατανομές σε εφαρμογές α-

ξιοπιστίας.

1.4.4 Log-Λογιστική κατανομή (Log-Logistic κατανομή)

Η Log-Logistic κατανομή χρησιμοποιείται για μη αρνητικές τυχαίες μεταβλητές και για γεγονότα των οποίων ο ρυθμός κινδύνου αρχικά αυξάνεται και μετέπειτα μειώνεται, για παράδειγμα η θνησιμότητα από καρκίνο μετά από διάγνωση ή θεραπεία. Η συνάρτηση επιβίωσης έχει τη μορφή:

$$S(t) = \frac{1}{1 + at^b} \quad (1.33)$$

Η συνάρτηση κινδύνου είναι:

$$\lambda(t) = \frac{abx^{b-1}}{1 + ax^b}. \quad (1.34)$$

Η κατανομή ονομάζεται έτσι επειδή είναι η κατανομή πιθανότητας μιας τυχαίας μεταβλητής της οποίας ο λογάριθμος ακολουθεί λογιστική κατανομή. Εάν $b \leq 1$ ο κίνδυνος μειώνεται και είναι σχηματικά όμοια με τη λογαριθμοκανονική κατανομή αλλά έχει βαρύτερες ουρές, ενώ αν $b > 1$ ο κίνδυνος είναι μονοκόρυφος και φθίνει μονότονα. Η αθροιστική συνάρτηση κατανομής της μπορεί να γραφεί σε κλειστή μορφή, σε αντίθεση με της λογαριθμοκανονικής.

1.4.5 Pareto κατανομή

Η συνάρτηση πιθανότητας της κατανομής Pareto με παραμέτρο σχήματος $c > 0$ και παράμετρο κλίμακας $\alpha > 0$ είναι:

$$f(t) = \alpha \frac{c^\alpha}{t^{\alpha+1}} \quad t \geq c. \quad (1.35)$$

Έτσι η συνάρτηση επιβίωσης της είναι:

$$S(t) = \left(\frac{c}{t}\right)^\alpha \quad (1.36)$$

και η συνάρτηση κινδύνου:

$$\lambda(t) = \frac{\alpha}{t} \quad (1.37)$$

Η συνάρτηση κινδύνου είναι φθίνουσα ως προς t και τείνει στο 0 καθώς $t \rightarrow \infty$ με αργό πολυωνυμικό ρυθμό, κάτι που μας δείχνει ότι η Pareto κατανομή έχει βαριά δεξιά ουρά και όσο πιο μικρή είναι η τιμή του α τόσο βαρύτερη είναι η ουρά. Αυτή η συμπεριφορά είναι αρκετά ασυνήθιστη αλλά όχι εντελώς άγνωστη στην πράξη. Θα ήταν κατάλληλη για τα άτομα που γίνονται λιγότερο επιρρεπή στην αποτυχία όσο μεγαλώνει η ηλικία τους, π.χ. οι άνθρωποι είναι λιγότερο επιρρεπείς στα λάθη όσο περνούν τα χρόνια λόγω της εμπειρίας τους.

1.4.6 Inverse-Gaussian κατανομή

Η κατανομή αυτή πρωτοεμφανίστηκε το 1915 από τον Schrödinger ως ο χρόνος έως το πρώτο πέρασμα της κίνησης Brown. Η συνάρτηση πυκνότητας πιθανότητας της Inverse-Gaussian κατανομής (γνωστή και ως κατανομή Wald κυρίως στη Ρώσικη βιβλιογραφία) είναι μια διπαραμετρική οικογένεια

$$f(x; \mu, \lambda) = \left[\frac{\lambda}{2\pi x^3} \right]^{\frac{1}{2}} \exp \frac{-\lambda(x - \mu)^2}{2\mu^2 x} \quad (1.38)$$

για $x > 0$, όπου $\mu > 0$ είναι ο μέσος και $\lambda > 0$ είναι η παράμετρος σχήματος. Είναι δεξιά ασύμμετρη κατανομή και χρησιμοποιείται ως κατανομή πληθυσμού όταν η λογαριθμοκανονική κατανομή έχει αρκετά βαριά δεξιά ουρά. Όταν το λ τείνει στο ∞ , η Inverse-Gaussian κατανομή τείνει να μοιάζει περισσότερο στην κανονική κατανομή.

1.5 Εκτίμηση πιθανοφάνειας χρόνων επιβίωσης

Εάν παρατηρούνται όλοι οι χρόνοι επιβίωσης για όλα τα n αντικείμενα μελέτης, χωρίς να υπάρχει λογοκρισία, τους χρόνους $t_1, t_2, \dots, t_n$, η πιθανοφάνεια έχει τη μορφή:

$$l = \prod_{i=1}^n f(t_i) \quad (1.39)$$

Ενώ σε μια πιο ρεαλιστική περίπτωση όπου τα r από τα n αντικείμενα αποτυγχάνουν τους χρόνους $t_{(1)}, t_{(2)}, \dots, t_{(r)}$, τα υπόλοιπα $n-r$ λογοκρίνονται τους χρόνους $t_1^+, t_2^+, \dots, t_n^+$, η πιθανοφάνεια έχει τη μορφή:

$$l = \prod_{j=1}^r f(t_{(j)}) \prod_{k=1}^{n-r} S(t_k^+). \quad (1.40)$$

όπου το πρώτο γινόμενο αφορά τις r αποτυχίες και το δεύτερο γινόμενο τις $n-r$ λογοκριμένες παρατηρήσεις.

Αναλυτικότερα, τα παρατηρούμενα δεδομένα μπορούν να περιγραφούν από το ζεύγος (t_i, δ_i) όπου t_i τελευταίος χρόνος που παρατηρήθηκε για το αντικείμενο i και δ_i η δείκτρια μεταβλητή:

$$\delta_i = \begin{cases} 0, & \text{λογοκρισία;} \\ 1, & \text{αποτυχία.} \end{cases}$$

και έτσι η συνάρτηση πιθανοφάνειας μπορεί να γραφτεί ως:

$$l = \prod_{j=1}^n f(t_j)^{\delta_j} S(t_j)^{1-\delta_j} = \prod_{j=1}^n \{\lambda(t_j) S(t_j)\}^{\delta_j} S(t_j)^{1-\delta_j} = \prod_{j=1}^n \lambda(t_j)^{\delta_j} S(t_j). \quad (1.41)$$

Ο εκτιμητής της μέγιστης πιθανοφάνειας προκύπτει από την μεγιστοποίηση του λογαρίθμου της πιθανοφάνειας (ή την ελαχιστοποίηση του αρνητικού λογαρίθμου της πιθανοφάνειας) και λύνοντας την εξίσωση πιθανοφάνειας

$$\frac{\partial \log L(\Theta|x)}{\partial \theta_i} = 0 \quad (1.42)$$

και ελέγχοντας το παρακάτω:

$$\frac{\partial^2 \log L(\Theta|x)}{\partial \theta_i^2} < 0 \quad (1.43)$$

Αλλά στην πράξη είναι συνήθως δύσκολο να πάρουμε μια αναλυτική λύση, ειδικά όταν χρειάζεται να λυθούν μη γραμμικές εξισώσεις. Υπάρχουν όμως αρκετοί αλγόριθμοι βελτιστοποίησης ώστε να μεγιστοποιούν τον λογάριθμο της πιθανοφάνειας όπως η μέθοδος Newton-Raphson.

1.6 Ημιπαραμετρικές μέθοδοι

1.6.1 Μοντέλο παλινδρόμησης ή αναλογικού κινδύνου του Cox

Το μοντέλο αναλογικού κινδύνου του Cox (Cox proportional hazard model) ή σε συντομία το PH μοντέλο του Cox, παρουσιάστηκε από τον Cox το 1972. Αποτελεί μια μέθοδο παλινδρόμησης προσαρμοσμένη ειδικά σε δεδομένα που περιέχουν λογοκριμένες παρατηρήσεις. Με βάση αυτό το μοντέλο μπορούν να εξεταστούν σχέσεις ανάμεσα στον κίνδυνο αποτυχίας/θανάτου και τις πληροφορίες που υπάρχουν για τα άτομα που εξετάζονται, αλλά και να εκτιμηθεί ο κίνδυνος αυτός δεδομένου των πληροφοριών. Το μοντέλο γράφεται ως

$$h(t|X) = h_0(t)\exp(X\beta) \quad (1.44)$$

όπου X είναι ένας $n \times p$ πίνακας που περιέχει σε στήλες τις p ανεξάρτητες μεταβλητές πάνω σε n παρατηρήσεις/άτομα, με X_{ij} να είναι η j -οστή συμμεταβλητή του i -οστού ατόμου, β είναι το διάνυσμα των συντελεστών παλινδρόμησης διάστασης $p \times 1$ ενώ $h_0(t)$ είναι μια συνάρτηση βασικού κινδύνου (baseline hazard function) η οποία δεν χρειάζεται να καθοριστεί πλήρως. Η συνάρτηση βασικού κινδύνου δηλώνει τον κίνδυνο που έχει οποιοδήποτε άτομο να υποστεί το εξεταζόμενο γεγονός σε ένα χρόνο t όταν $X=0$. Ο ρυθμός του γεγονότος δεν μπορεί να είναι αρνητικός (παρατηρούμενοι θάνατοι δεν μπορούν να μην συμβούν) και έτσι η εκθετική μορφή παίζει σημαντικό ρόλο αφού εξασφαλίζει ότι το τελικό αποτέλεσμα θα είναι μια φυσική πιθανότητα.

Η γενική μορφή ενός μοντέλου αναλογικού κινδύνου (proportional hazards model) είναι:

$$h(t, X) = h_0(t)g(X) \quad (1.45)$$

όπου $g(X)$ είναι μια συνάρτηση του διανύσματος X . Στην περίπτωση του PH μοντέλου του Cox, η $g(X)$ είναι η συνάρτηση $e^{\beta_1 x_1 + \dots + \beta_p x_p}$.

Ο τύπος του μοντέλου του Cox έχει την ιδιότητα ότι αν όλες οι μεταβλητές X είναι ίσες με μηδέν ή αν δεν υπάρχει καμία μεταβλητή X στο μοντέλο, τότε ο τύπος μειώνεται στην συνάρτηση βασικού κινδύνου. Αυτή η ιδιότητα του μοντέλου Cox είναι ο λόγος που η $h_0(t)$ ονομάζεται βασική συνάρτηση.

Το μοντέλο του Cox έχει σαν προϋπόθεση την υπόθεση της αναλογικότητας, για αυτό ονομάζεται και μοντέλο αναλογικού κινδύνου. Δηλαδή, αν θεωρήσουμε δύο ασθενείς, με τιμές συμμεταβλητών X και X^* αντίστοιχα, η αναλογία των κινδύνων τους (hazard ratio) γράφεται:

$$\frac{h(t|X)}{h(t|X^*)} = \frac{h_0(t)\exp(X\beta)}{h_0(t)\exp(X^*\beta)} = \exp[(X - X^*)\beta]$$

Παρατηρούμε ότι ο λόγος των κινδύνων είναι σταθερός και δεν εξαρτάται από τον χρόνο.

Εάν $\exp[(X - X^*)\beta] > 1$, ο κίνδυνος ενός ασθενούς με τιμές συμμεταβλητών X θα είναι μεγαλύτερος από τον κίνδυνο ενός ασθενούς με τιμές συμμεταβλητών X^* ενώ η πιθανότητα επιβίωσης ενός ατόμου της πρώτης ομάδας θα είναι μικρότερη από την πιθανότητα επιβίωσης ενός ατόμου της δεύτερης ομάδας. Ακριβώς τα αντίθετα συμβαίνουν εάν $\exp[(X - X^*)\beta] < 1$.

Εάν αυξήσουμε μια επεξηγηματική μεταβλητή από x σε $x+1$, θα επέλθουν κάποιες αλλαγές. Στον λογαριθμικό κίνδυνο $\log h(t|X)$ θα προστεθεί ο παράγοντας β , ενώ ο κίνδυνος θα πολλαπλασιαστεί με τον παράγοντα e^β . Η καμπύλη επιβίωσης $S(t)$ θα γίνει $S(t)\exp(\beta)$, ενώ η λογαριθμική καμπύλη επιβίωσης θα πολλαπλασιαστεί με e^β .

Στην περίπτωση που κάποιες μεταβλητές είναι ποιοτικές, χρησιμοποιούνται ψευδο-μεταβλητές (Dummy variables) για να περιγράψουν την πληροφορία που εμπεριέχεται στις ποιοτικές αυτές μεταβλητές. Οι κατηγορικές μεταβλητές μπορεί να είναι είτε δυαδικές είτε να έχουν περισσότερες από 2 κατηγορίες. Για κάθε κατηγορική μεταβλητή με $i \geq 2$ κατηγορίες, απαιτείται η κατασκευή $i - 1$ ψευδομεταβλητών. Οι ψευδομεταβλητές παίρνουν τις τιμές $\{0, 1\}$.

Το μοντέλο του Cox είναι ημιπαραμετρικό επειδή ενώ η συνάρτηση βασικού κινδύνου μπορεί να είναι οποιασδήποτε μορφής, οι συμμεταβλητές εισάγονται στο μοντέλο γραμμικά.

Μπορεί να χειριστεί και συνεχή και διακριτά δεδομένα, όχι όμως και χρονικά εξαρτημένες μεταβλητές. Τις χρονικά εξαρτημένες μεταβλητές χειρίζεται το εκτεταμένο μοντέλο του Cox. Μια χρονικά εξαρτημένη μεταβλητή ορίζεται ως μια μεταβλητή της οποίας η τιμή για ένα συγκεκριμένο άτομο δεν αλλάζει με τον χρόνο, π.χ. το φύλο. Επίσης, μεταβλητές όπως η ηλικία και το βάρος που αλλάζουν με την πάροδο του χρόνου μπορεί να είναι σκόπιμο να θεωρηθούν ως χρονικά ανεξάρτητες

αν οι τιμές τους δεν αλλάζουν τόσο πολύ με τον χρόνο ή αν η επίδραση τους στον κίνδυνο επιβίωσης εξαρτάται ουσιαστικά από μία μόνο μέτρηση.

Οι συντελεστές παλινδρόμησης σε κάθε περίπτωση εκτιμούνται καλά χωρίς να έχουν προηγηθεί υποθέσεις για την συνάρτηση βασικού κινδύνου ή τον πληθυσμό από τον οποίο προέρχονται τα δεδομένα. Το μοντέλο αναλογικών κινδύνων του Cox είναι ένα robust μοντέλο, δηλαδή τα αποτελέσματα που προκύπτουν από αυτό προσεγγίζουν πολύ καλά τα αποτελέσματα ενός σωστού παραμετρικού μοντέλου. Το σωστό παραμετρικό μοντέλο είναι σίγουρα προτιμότερο αλλά δεν είναι πάντα εύκολο να βρεθεί.

Έτσι, η ευελιξία, η ευστάθεια και η ευρεία εφαρμογή του μοντέλου παλινδρόμησης του Cox, μαζί με την απλότητα της εφαρμογής του συνετέλεσαν στο να γίνει ο πιο δημοφιλής τρόπος μοντελοποίησης δεδομένων επιβίωσης. Συνεπώς, ερευνητές το χρησιμοποιούν συχνά χωρίς δεύτερες σκέψεις αγνοώντας το γεγονός ότι η αναλογικότητα του μοντέλου είναι μια περιοριστική υπόθεση που πρέπει να γίνει. Γι αυτό προτείνεται να γίνεται μια προκαταρκτική ανάλυση ώστε να διερευνηθεί το κατά πόσον η υπόθεση της αναλογικότητας ευσταθεί ή όχι.

Εκτίμηση των συντελεστών παλινδρόμησης

Η συνάρτηση βασικού κινδύνου $h_0(t)$ δεν καθορίζεται παραμετρικά, κι έτσι δεν μπορεί να χρησιμοποιηθεί η συνηθισμένη συνάρτηση πιθανοφάνειας για την εκτίμηση του διανύσματος β . Σκοπός μας λοιπόν είναι να εκτιμηθεί το β χωρίς να χρειαστεί να εμπλακεί η $h_0(t)$.

Η καινοτομία που εισήγαγε ο Cox ήταν η εκτίμηση των παραμέτρων του μοντέλου μέσω της συνάρτησης μερικής πιθανοφάνειας (partial likelihood function).

Έστω $t_{(1)}, t_{(2)}, \dots, t_{(D)}$ οι διατεταγμένοι χρόνοι θανάτου (δεν υπάρχουν ίδιοι χρόνοι για θέμα απλότητας) και X_i είναι η i -οστή γραμμή του πίνακα συμμεταβλητών που σχετίζεται με τον χρόνο t_i . Ορίζουμε το σύνολο των ατόμων σε κίνδυνο (risk set) στον χρόνο t_i , $R(t_i)$ ως το πλήθος όλων των ατόμων που βρίσκονται σε κίνδυνο και υπό μελέτη, σε χρόνο ακριβώς πριν το t_i . Η συνάρτηση πιθανοφάνειας γράφεται ως:

$$L(\beta) = \prod_{i=1}^D \frac{\exp(X_i \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \quad (1.46)$$

Στον αριθμητή είναι η πληροφορία για τα άτομα που παθαίνουν το εξεταζόμενο γεγονός, ενώ στον παρονομαστή είναι η πληροφορία όλων των ατόμων που δεν έχουν υποστεί ακόμα το γεγονός, ακόμα και αυτών που μπορεί να λογοκριθούν αργότερα. Αν στα δεδομένα υπάρχουν ισότιμες χρονικές παρατηρήσεις, τότε η συνάρτηση μερικής πιθανοφάνειας περιέχει παραλλαγές και μπορεί να γίνει πολύ χρονοβόρα. Στην περίπτωση αυτή χρησιμοποιούνται προσεγγίσεις της μερικής πιθανοφάνειας που προτάθηκαν από τον Breslow ή από τον Efron.

Η παραπάνω συνάρτηση δεν είναι συνάρτηση πλήρους πιθανοφάνειας επειδή μελετά τις πιθανότητες μόνο για τα άτομα που αποτυγχάνουν κι όχι για αυτά που λογοκρίνονται όπως και το ότι βασίζεται στην παρατηρούμενη σειρά των γεγονότων και όχι στην από κοινού κατανομή των γεγονότων. Η συνάρτηση μερικής πιθανοφάνειας δεν εξαρτάται από τη βασική συνάρτηση κινδύνου και μεγιστοποιείται σαν να ήταν πλήρης πιθανοφάνεια χρησιμοποιώντας τη μέθοδο Newton-Raphson.

Είναι πιθανό να πάρουμε μια άπειρη εκτίμηση του β . Για παράδειγμα, αν υπάρχουν μόνο δύο παρατηρήσεις, ένας άντρας και μια γυναίκα, τότε $\hat{\beta} = -\infty$ αν ο άντρας πεθάνει πρώτος και $\hat{\beta} = +\infty$ αν η γυναίκα πεθάνει πρώτη. Γενικότερα, άπειρες εκτιμήσεις προκύπτουν όταν υπάρχει μια γραμμική συνάρτηση των συμμεταβλητών τέτοια ώστε τα γεγονότα να συμβαίνουν σύμφωνα με αυτήν.

Οι εξισώσεις των scores βρίσκονται αν πάρουμε την πρώτη μερική παράγωγο της

$$l(\beta) = \sum_{i=1}^D X_i \beta - \sum_{i=1}^D \ln \left[\sum_{j \in R(t_i)} \exp(X_j \beta) \right] \quad (1.47)$$

ως προς τα β .

Έστω $U(\beta) = \frac{\partial l(\beta)}{\partial \beta}$ η συνάρτηση των scores τότε

$$U(\beta) = \sum_{i=1}^D X_i - \sum_{i=1}^D \frac{\sum_{j \in R(t_i)} X_j \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \quad (1.48)$$

Ο πίνακας πληροφορίας δίνεται από τον αρνητικό πίνακα της δεύτερης μερικής παρα-

γώγου της (1.47) με το kl στοιχείου του πίνακα να υπολογίζεται ως:

$$l_{kl} = \frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_l} = \sum_{i=1}^D \frac{\sum_{j \in R(t_i)} X_{jk} X_{jl} \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} - \sum_{i=1}^D \left[\frac{\sum_{j \in R(t_i)} X_{jk} \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \right] \left[\sum_{i=1}^D \frac{\sum_{j \in R(t_i)} X_{jl} \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \right] \quad (1.49)$$

Εκτίμηση της βασικής συνάρτησης κινδύνου και της συνάρτησης επιβίωσης

Η εκτίμηση της συνάρτησης επιβίωσης $S(t|X)$, απαιτεί την εκτίμηση της βασικής συνάρτησης κινδύνου. Ο Breslow προτείνει ως εκτιμητή της την συνάρτηση:

$$\hat{h}_0(t) = \frac{d_i}{\sum_{j \in R(t_i)} \exp(X_j \hat{\beta})} \quad (1.50)$$

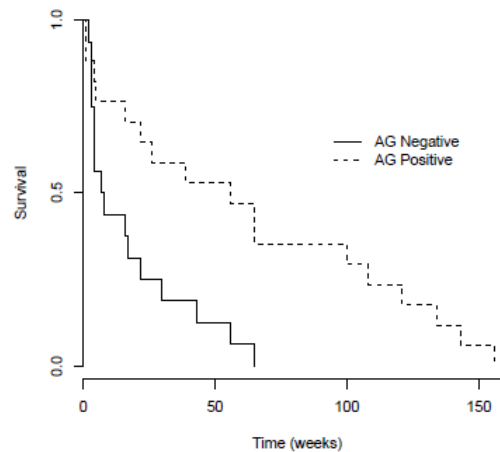
με d_i μια δείκτρια συνάρτηση που παίρνει την τιμή 1 όταν συμβεί το γεγονός και 0 αλλιώς. Άρα, η βασική αθροιστική συνάρτηση κινδύνου δίνεται από τον τύπο

$$\hat{H}_0(t) = \sum_{t_i \leq t} \hat{h}_0(t) \quad (1.51)$$

ενώ η βασική συνάρτηση επιβίωσης $S_0(t)$ εκτιμάται από την $\hat{S}_0(t) = \exp[-\hat{H}_0(t)]$. Η βασική συνάρτηση επιβίωσης δίνει τις πιθανότητες επιβίωσης για ένα άτομο όταν όλες οι τιμές των συμμεταβλητών X είναι 0. Τελικά η πιθανότητα επιβίωσης ενός ατόμου υπολογίζεται μέσω των παραπάνω ποσοτήτων ως:

$$\hat{S}(t|X_0) = \hat{S}_0(t)^{\exp(X \hat{\beta})} \quad (1.52)$$

Το μοντέλο αναλογικού κινδύνου αποτελεί μια ειδική περίπτωση ενός γενικευμένου γραμμικού μοντέλου. Συνεπώς, οι στατιστικοί έλεγχοι που χρησιμοποιούνται σε μοντέλα λογιστικής παλινδρόμησης κ.ά. και τους οποίους θα δούμε λίγο πιο μετά μπορούν να χρησιμοποιηθούν και εδώ με ανάλογο τρόπο.



Σχήμα 1.6: Γραφήμα της συνάρτησης επιβίωσης (μη παράλληλες καμπύλες) για τις δύο ομάδες θεραπείας ασθενών

Για τον έλεγχο αναλογικότητας δύο κινδύνων, δηλαδή την σύγκριση δύο θεραπειών, μπορούμε απλά να κάνουμε το γράφημα των δύο καμπύλων επιβίωσης και να παρατηρήσουμε κατά πόσον οι καμπύλες τείνουν να αποκλίνουν με τον χρόνο ή όχι. Εάν οι δύο καμπύλες είναι παράλληλες τότε αυτό σημαίνει ότι οι κίνδυνοι δεν είναι αναλογικοί και σε αυτήν την περίπτωση θα πρέπει να εξετασθούν άλλοι τρόποι μοντελοποίησης με προτίμηση στα παραμετρικά μοντέλα. Εάν κανένα μοντέλο δεν χρησιμοποιείται για να προσαρμοσθούν τα δεδομένα επιβίωσης τότε η καμπύλη επιβίωσης μπορεί να εκτιμηθεί μέσω της μεθόδου Kaplan-Meier.

1.7 Συμπερασματολογία

1.7.1 Έλεγχος λόγου πιθανοφάνειας (Likelihood ratio test)

Στην παραμετρική ανάλυση επιβίωσης όταν επιλεχθεί η κατάλληλη κατανομή για τα δεδομένα επιβίωσης το κύριο ερώτημα είναι να διαπιστωθεί ποιες συμμεταβλητές πρέπει να περιλαμβάνονται στο μοντέλο. Αυτό οδηγεί στην σύγκριση εμφωλευμένων(nested) μοντέλων. Και αυτή η σύγκριση γίνεται δυνατή μέσω του likelihood ratio test. Για παράδειγμα, έστω το Μοντέλο I που περιέχει p επεξηγηματικές μετα-

βλητές και το Μοντέλο II που περιέχει επιπλέον q επεξηγηματικές μεταβλητές.

$$Model I : h_i(t) = \exp\{\hat{\beta}_1 x_1 + \dots + \hat{\beta}_p x_p\} \hat{\theta} \hat{\alpha} t^{\hat{\alpha}-1}$$

$$Model II : h_i(t) = \exp\{\hat{\beta}_1 x_1 + \dots + \hat{\beta}_{p+q} x_{p+q}\} \hat{\theta} \hat{\alpha} t^{\hat{\alpha}-1}$$

Τότε

$$-2[\log \hat{L}_I - \log \hat{L}_{II}] \sim X_q^2 \quad (1.53)$$

Αν η διαφορά είναι μεγαλύτερη από το ποσοστιαίο σημείο της X_q^2 τότε αποφασίζεται ότι οι q συμμεταβλητές χρειάζονται στο μοντέλο.

Στο μοντέλο αναλογικών κινδύνων του Cox, μπορούν να εφαρμοσθούν τα παραπάνω μέσω της χρήσης της μερικής πιθανοφάνειας.

1.7.2 Έλεγχος Wald (Wald test)

Ο έλεγχος Wald που προτάθηκε από τους Polit(1996) και Agresti(1990) είναι ένας τρόπος (από τους πολλούς) να εξετασθεί κατά πόσον οι παράμετροι που συνδέονται με ένα σύνολο επεξηγηματικών μεταβλητών είναι μηδέν. Για κάθε επεξηγηματική μεταβλητή στο μοντέλο υπάρχει και μία παράμετρος. Η ελεγχοσυνάρτηση Wald είναι:

$$W = \frac{\hat{\beta}^2}{Var(\hat{\beta})} \sim X_1^2. \quad (1.54)$$

Είναι φανερό ότι η υπερβολική διασπορά επιδρά σημαντικά στον έλεγχο αυτό. Αν για μία συγκεκριμένη επεξηγηματική μεταβλητή, ή ένα σύνολο επεξηγηματικών μεταβλητών, ο έλεγχος Wald είναι σημαντικός, τότε προκύπτει το συμπέρασμα ότι οι παράμετροι που συνδέονται με αυτές τις μεταβλητές δεν είναι μηδέν, άρα αυτές οι μεταβλητές θα πρέπει να συμπεριληφθούν στο μοντέλο. Για μία μόνο παράμετρο ο έλεγχος Wald είναι το τετράγωνο του t ελέγχου και έτσι δίνουν ακριβώς τα ίδια αποτελέσματα (Altman(1991)).

Οι έλεγχοι Wald και λόγου πιθανοφανειών μπορεί να μη δίνουν ακριβώς τα ίδια αποτελέσματα. Έχει δειχθεί ότι από αυτές τις δύο διαδικασίες ελέγχου, ο λόγος πιθανοφάνειας έχει καλύτερες στατιστικές ιδιότητες, έτσι σε περιπτώσεις διλημάτων προτιμούμε να χρησιμοποιούμε τον έλεγχο λόγου πιθανοφάνειας.

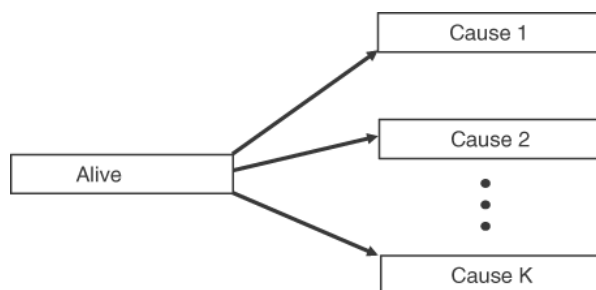
Κεφάλαιο 2

Ανταγωνιστικοί κίνδυνοι (Competing risks)

Μέχρι στιγμής, το ενδιαφέρον μας περιστράφηκε γύρω από διαδικασίες «μιας κατάστασης» αφού υποθέταμε ότι κάθε άτομο βρίσκεται σε κίνδυνο λόγω ενός γεγονότος μόνο, σε οποιαδήποτε χρονική στιγμή. Όμως, υπάρχουν περιπτώσεις στις οποίες τα άτομα είναι εκτεθειμένα σε περισσότερα από ένα είδη αποτυχίας τα οποία είναι αμοιβαίως αποκλειόμενα. Έτσι καθορίζεται μια κατάσταση ανταγωνιστικών κινδύνων. Η θεωρία των ανταγωνιστικών κινδύνων προέρχεται από τις προσπάθειες του David Bernoulli το 1760 να αποδεσμεύσει τον κίνδυνο θανάτου από ευλογία από άλλες αιτίες. Αυτό είναι ένα κλασσικό παράδειγμα ανταγωνιστικών κινδύνων όπου τα άτομα μπορεί να αποβιώσουν από πολλαπλές αιτίες. Παρόμοιες εφαρμογές εμφανίζονται στη δημογραφία και στην αναλογιστική επιστήμη. Ένας σκοπός των ανταγωνιστικών κινδύνων είναι να συγκριθούν οι ρυθμοί αποτυχίας δύο ή περισσότερων γεγονότων.

Για κάθε άτομο τα μοντέλα ανταγωνιστικών κινδύνων αναλύουν τον χρόνο μέχρι να συμβεί το πρώτο γεγονός αλλά και τον τύπο του γεγονότος εκείνη την στιγμή. Σε αντίθεση με την τυπική ανάλυση επιβίωσης που εξετάζει μόνο τον χρόνο μέχρι να συμβεί το πρώτο γεγονός. Για παράδειγμα, μοντέλα επιβίωσης εξετάζουν την διάρκεια παραμονής ασθενών σε νοσοκομείο με ένα συνδυασμένο τελικό σημείο έξοδος από το νοσοκομείο ζωντανός-θάνατος στο νοσοκομείο. Ενώ οι ανταγωνιστικοί κίνδυνοι μοντελοποιούν και τον τύπο του τελικού σημείου.

Το σχήμα (2.1) απεικονίζει ένα τυπικό μοντέλο ανταγωνιστικών κινδύνων.



Σχήμα 2.1: Τυπικό μοντέλο ανταγωνιστικών κινδύνων

2.1 Βασικές συναρτήσεις πιθανοτήτων

Κάθε άτομο συνδέεται με ένα ζεύγος (T, C) όπου $T > 0$ είναι ο χρόνος μέχρι το πρώτο γεγονός και $C \in \{1, 2, \dots, k\}$ ο τύπος της αποτυχίας. Ο χρόνος T είναι μη αρνητική και συνήθως συνεχής τυχαία μεταβλητή ενώ η αιτία αποτυχίας είναι διακριτή τυχαία μεταβλητή που παίρνει k τιμές, όσες και οι δυνατές αιτίες αποτυχίας (αν και παρατηρούμε μόνο την πρώτη αποτυχία). Σε κάθε αποτυχία αντιστοιχεί ένας χρόνος αποτυχίας $T_j, j = 1, \dots, k$, μια συνάρτηση κινδύνου λ_j και μια συνάρτηση επιβίωσης S_j .

Η απο κοινού κατανομή του ζεύγους (T, C) για ένα άτομο προσδιορίζεται από την:

$$F_j(t) = P(T \leq t, C = j); t > 0, j = 1, \dots, k.$$

η οποία ορίζεται και ως συνάρτηση υπο-κατανομής.

Κι έτσι η περιθώρια κατανομή του T δίνεται από την αθροιστική συνάρτηση κατανομής:

$$F(t) = P(T \leq t) = \sum_{j=1}^k F_j(t)$$

Ενώ η περιθώρια κατανομή του C δίνεται από την $\pi_j = P(C = j) = F_j(\infty) = F_j(0)$. Έστω $\lambda_j(t)$ ο κίνδυνος θανάτου λόγω της αιτίας $j, j = 1, \dots, k$. Η ανά αιτία συνάρ-

τηση κινδύνου(cause specific hazard function) μπορεί εναλλακτικά να δοθεί και από την παρακάτω σχέση:

$$\lambda_j = \lim_{dt} \frac{P(t < T \leq t + dt, C = j | T > t)}{dt} \quad (2.1)$$

και προφανώς προκύπτει ότι ο κίνδυνος που αντιστοιχεί στη συνολική θνησιμότητα είναι $\lambda(t) = \sum_{j=1}^k \lambda_j(t)$.

Η (2.1) εκφράζει τον στιγμιαίο ρυθμό αποτυχίας από την αιτία j παρουσία των υπόλοιπων αιτιών. Μέσω αυτής μπορεί να οριστεί η συνολική συνάρτηση επιβίωσης από όλες τις αιτίες ($S(t)$) και να ξαναοριστεί η αθροιστική επίπτωση (cumulative incidence, $F_j(t)$) για την αιτία j .

$$S(t) = \exp\left\{-\int_0^t \sum_{j=1}^k \lambda_j(u) du\right\} = \exp^{-\sum_{j=1}^k \Lambda_j(t)}. \quad (2.2)$$

$$F_j(t) = \int_0^t \lambda_j(u) S(u) du \quad (2.3)$$

όπου $\Lambda_j(t) = \int_0^t \sum_{j=1}^k \lambda_j(u) du$

Η συνολική συνάρτηση επιβίωσης (2.2) εκφράζει την πιθανότητα επιβίωσης από όλες τις αιτίες αποτυχίας τουλάχιστον για χρόνο t αλλά είναι και η συνεισφορά στην πιθανοφάνεια ενός λογοκριμένου χρόνου t , ενώ η αθροιστική επίπτωση (2.3) την πιθανότητα αποτυχίας από την αιτία j μέσα σε χρόνο t .

Η συνεισφορά ενός θανάτου από την αιτία j την στιγμή t είναι η αντίστοιχη πυκνότητα $\lambda_j(t)S(t)$. Έτσι η πιθανότητα μετάβασης για την αιτία j είναι

$$P_j(t) = \int_0^t \lambda_j(u) S(u) du = \int_0^t \lambda_j(u) \exp\left\{-\int_0^u \sum_l \lambda_l(u) du\right\}. \quad (2.4)$$

Για κάθε αιτία θανάτου, υπάρχει ένας όρος κινδύνου, εξαρτώμενος μόνο από την πραγματική αιτία, και ένας εκθετικός όρος εξαρτώμενος από τον ολοκληρωμένο κίνδυνο για όλες τις αιτίες μαζί.

Κάτω από την προϋπόθεση ότι ο χρόνος λογοκρισίας είναι ανεξάρτητος από τον

χρόνο αποτυχίας (από οποιαδήποτε αιτία) η συνάρτηση πιθανοφάνειας έχει την μορφή:

$$L = \prod_{i=1}^n \prod_{j=1}^k \lambda_j(t_i)^{d_{ij}} e^{-\Lambda_j(t_i)}. \quad (2.5)$$

Όπου ο δείκτης d_{ij} δηλώνει κατά πόσον το άτομο i πέθανε ή όχι από την αιτία j . Εάν υπάρχει λογοκρισία τότε ο χρόνος επιβίωσης θεωρείται ίσος με 0. Η ολική συνάρτηση πιθανοφάνειας είναι ένα γινόμενο από k πιθανοφάνειες, μία για κάθε αιτία αποτυχίας. Αυτό σημαίνει ότι η ανά αιτία συνάρτηση κινδύνου μπορεί να εκτιμηθεί μεγιστοποιώντας ξεχωριστές πιθανοφάνειες, εφ'όσον αυτές δεν εξαρτώνται από τις ίδιες παραμέτρους. Επίσης παρατηρούμε ότι η πιθανοφάνεια για μια συγκεκριμένη αιτία αποτυχίας είναι ακριβώς η ίδια με την πιθανοφάνεια που θα προέκυπτε αν αντιμετωπίζαμε τις υπόλοιπες αιτίες ως λογοκριμένες παρατηρήσεις. Επίσης κάτω από την προϋπόθεση των ανεξάρτητων ανταγωνιστικών κινδύνων, μπορεί να εφαρμοσθεί ο εκτιμητής Nelson-Aalen ώστε να εκτιμηθεί η αθροιστική συνάρτηση κινδύνου.

2.2 Λανθάνοντες χρόνοι αποτυχίας (Latent failure times)

Σύμφωνα με την προσέγγιση των λανθάνοντων χρόνων αποτυχίας, σε κάθε αιτία αποτυχίας j αντιστοιχεί μια υποθετική τυχαία μεταβλητή T_j που εκφράζει τον χρόνο μέχρι την εμφάνιση της j αποτυχίας. Όμως μόνο ο χρόνος μέχρι την εμφάνιση της πρώτης αιτίας μπορεί να παρατηρηθεί. Οι υπόλοιποι χρόνοι δεν παρατηρούνται προφανώς αφού οι αιτίες είναι καταληκτικά συμβάντα και έτσι αποτελούν υποθετικές ποσότητες. Δηλαδή, χειριζόμαστε τους υπόλοιπους χρόνους ως λογοκριμένους χρόνους. Για απλοποίηση, συνήθως υποθέτουμε ότι δεν γίνεται κάποιος ασθενής να αποτύχει με δύο διαφορετικούς τρόπους την ίδια ακριβώς χρονική στιγμή. Αν όμως θέλησουμε να καταγραφούν και οι υπόλοιποι χρόνοι μετά την εμφάνιση της πρώτης αιτίας αποτυχίας τότε οδηγούμαστε στα προβλήματα ημιανταγωνιστικών κινδύνων.

Έτσι ο χρόνος αποτυχίας είναι ο μικρότερος λανθάνοντας χρόνος $T = \min(T_1, T_2, \dots, T_k)$ και αιτία αποτυχίας αυτή που αντιστοιχεί στον μικρότερο αυτόν χρόνο.

Η από κοινού συνάρτηση κατανομής των λανθανόντων χρόνων αποτυχίας $T_1, \dots, T_k$

ορίζεται, μέσω της k -διάστατης συνάρτησης επιβίωσης, ως εξής:

$$\bar{K}(t_1, \dots, t_k) = P(t_1 > t_1, \dots, T_k > t_k). \quad (2.6)$$

Η περιθώρια συνάρτηση επιβίωσης (marginal survival function) του λανθάνοντος χρόνου αποτυχίας T ορίζεται ως:

$$\bar{F}(t) = \bar{K}(t, t, \dots, t) \quad (2.7)$$

Η συνάρτηση υποκατανομής μπορεί επίσης να υπολογιστεί άμεσα από την από κοινού συνάρτηση επιβίωσης ως:

$$f_j(t) = - \left(\frac{\partial \bar{K}(t_1, \dots, t_k)}{\partial t_j} \right)_{t_1=\dots=t_k=t} \quad (2.8)$$

και επομένως προκύπτει ότι η περιθώρια συνάρτηση κινδύνου του λανθάνοντος χρόνου ορίζεται ως:

$$\lambda_j(t) = \frac{f_j(t)}{\bar{F}(t)} = - \left(\frac{\partial \log \bar{K}(t_1, \dots, t_k)}{\partial t_j} \right)_{t_1=\dots=t_k=t} \quad (2.9)$$

Γενικά, η συνάρτηση υποκατανομής $f_j(t)$ και η περιθώρια συνάρτηση κινδύνου του λανθάνοντος χρόνου $\lambda_j(t)$ είναι διαφορετικές κι έχουν διαφορετικές ερμηνείες. Όμως, όταν οι χρόνοι T_j είναι ανεξάρτητοι τότε $f_j(t) = \lambda_j(t)$ για κάθε $t > 0$.

2.3 Το προσδιοριστικό ζήτημα (The identifiability issue)

Η από κοινού κατανομή καθώς και οι περιθώριες κατανομές των λανθανόντων χρόνων αποτυχίας δεν είναι εκτιμήσιμες από δεδομένα ανταγωνιστικών κινδύνων, χωρίς την αποδοχή συγκεκριμένων υποθέσεων για την από κοινού κατανομή των λανθανόντων χρόνων και την μεταξύ τους εξάρτηση. Μια προϋπόθεση που έχει χρησιμοποιηθεί αρκετά συχνά για τον προσδιορισμό των περιθώριων συναρτήσεων επιβίωσης είναι αυτή της ανεξαρτησίας μεταξύ των λανθανόντων χρόνων. Η προϋπόθεση αυτή βέβαια

είναι σωστή αν και μόνο αν ισχύει:

$$\bar{K}(t_1, \dots, t_k) = \prod_{j=1}^k \bar{G}_j^*(t_j) \quad (2.10)$$

όπου $\bar{G}_j^*(t_j) = e^{-\Lambda_j(t)}$ και $\Lambda_j(t) = \int_0^t \lambda_j(u) du$ η αθροιστική συνάρτηση υπο-κινδύνου. Κάτω από την προϋπόθεση αυτή μπορεί να γίνει προσδιορισμός των περιθωρίων συναρτήσεων επιβίωσης. Όμως, τέτοιες υποθέσεις δεν μπορούν να ελεγχθούν από δεδομένα που γνωρίζουμε, τις ποσότητες δηλαδή που μπορούμε να εκτιμήσουμε. Ο Tsiatis το 1975 απέδειξε ότι σε κάθε μοντέλο η από κοινού κατανομή με εξαρτημένους χρόνους αποτυχίας αντιστοιχεί σε ένα μοναδικό μοντέλο με ανεξάρτητους χρόνους από το οποίο προκύπτει το ίδιο ακριβώς σύνολο συναρτήσεων υποκατανομής με το σύνολο συναρτήσεων υποκατανομής που είχε η αρχική από κοινού κατανομή, άρα και την ίδια ακριβώς από κοινού κατανομή. Έτσι σύμφωνα με τις ποσότητες που μπορούν να εκτιμηθούν δεν είναι δυνατόν να προσδιορισθεί εάν το πραγματικό μοντέλο είναι αυτό που προέκυψε με τους ανεξάρτητους χρόνους ή κάποιο άλλο από τα άπειρα με τις ίδιες συναρτήσεις υποκατανομής με τους εξαρτημένους χρόνους αποτυχίας.

2.4 Μη παραμετρική εκτίμηση

Kaplan-Meier εκτιμητής

Ο εκτιμητής Kaplan-Meier μπορεί εύκολα να γενικευθεί ώστε να περιλαμβάνει και τους ανταγωνιστικούς κινδύνους. Έστω $t_{j1} < t_{j2} < \dots < t_{jk_j}$ οι k_j διακεκριμένοι χρόνοι αποτυχίας για την αιτία αποτυχίας j . Έστω n_{ji} ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο πριν την στιγμή t_{ji} και έστω d_{ji} ο αριθμός των θανάτων λόγω της αιτίας j την στιγμή t_{ji} . Τότε ο εκτιμητής Kaplan-Meier παίρνει την παρακάτω μορφή:

$$\hat{S}_j(t) = \prod_{i: t_{ji} < t} \left(1 - \frac{d_{ji}}{n_{ji}} \right). \quad (2.11)$$

Είναι ενδιαφέρον να σημειωθεί ότι η $\hat{S}_j(t)$ είναι ακριβώς η ίδια όπως στον τυπικό εκτιμητή Kaplan-Meier εάν λογοκρίνουμε όλες τις υπόλοιπες αιτίες αποτυχίας.

Εάν δεν υπάρχουν ισοπαλίες μεταξύ διαφορετικών τύπων αποτυχίας, τότε

$$\hat{S}(t) = \prod_{j=1}^m \hat{S}_j(t) \quad (2.12)$$

δηλαδή ο ολικός Kaplan-Meier εκτιμητής είναι το γινόμενο των ανά αιτία Kaplan-Meier εκτιμητών.

Κεφάλαιο 3

Μοντέλα πολλαπλών καταστάσεων (Multi state models)

Τα μοντέλα πολλαπλών καταστάσεων (multi-state models) (ή αλλιώς πολλαπλών στρωμάτων multi-strata models) των Brownie et al. (1993) και Hestbeck et al. (1991) είναι τα μοντέλα τα οποία χρησιμοποιούνται περισσότερο για την ανάπτυξη των επαναλαμβανόμενων (longitudinal) δεδομένων. (Ένα σύνολο δεδομένων λέγονται επαναλαμβανόμενα αν παρακολουθείται το ίδιο είδος πληροφοριών σχετικά με τα ίδια θέματα σε πολλαπλές χρονικές στιγμές π.χ. συγκεκριμένοι ασθενείς και τα αποτελέσματα κλινικών εξετάσεων τους σε πέντε διαδοχικά έτη). Ένα μοντέλο πολλαπλών καταστάσεων ορίζεται ως το μοντέλο για μια συνεχούς χρόνου στοχαστική διαδικασία, η οποία σε κάθε χρονικό σημείο απασχολεί ένα σύνολο από διακριτές καταστάσεις και επιτρέπει στα άτομα να μετακινούνται μεταξύ πεπερασμένου αριθμού καταστάσεων.

Στην ιατρική, οι καταστάσεις μπορούν να περιγράψουν καταστάσεις όπως υγιής, ασθενής, ασθενής με επιπλοκές και νεκρός. Όλα τα πολλαπλά δεδομένα επιβίωσης μπορούν να διατυπωθούν ως μοντέλα πολλαπλών καταστάσεων και αυτό είναι βολικό για τη μελέτη προβλέψεων. Η αλλαγή κατάστασης καλείται μετάβαση ή γεγονός. Είναι σημαντικό να αναγνωρίζεται η διαφορά μεταξύ ενός γεγονότος (όπως ο θάνατος) και μιας κατάστασης (όπως η κατάσταση νεκρός). Η δομή των καταστάσεων

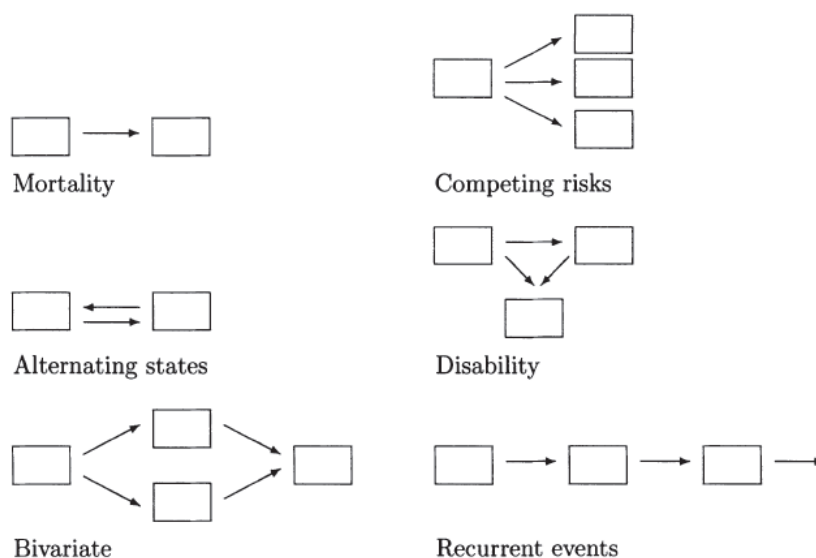
καθορίζει τις καταστάσεις και ποιες μεταβάσεις από κατάσταση σε κατάσταση είναι πιθανές. Το πλήρες στατιστικό μοντέλο καθορίζει την δομή των καταστάσεων και την μορφή της συνάρτησης κινδύνου για κάθε δυνατή μετάβαση. Η πολυπλοκότητα των μοντέλων πολλαπλών καταστάσεων εξαρτάται σημαντικά από τον αριθμό των καταστάσεων αλλά και τις πιθανές μεταβάσεις.

Είναι ένα πιο πολύπλοκο μοντέλο σε σχέση με τους ανταγωνιστικούς κινδύνους και επιτρέπει την περαιτέρω μελέτη γεγονότων που συμβαίνουν π.χ. κατά τη διάρκεια παραμονής ασθενούς στη μονάδα θεραπείας όπως το άνοιγμα ή το κλείσιμο του αναπνευστήρα, η χρήση καθετήρα ή η εμφάνιση νοσοκομειακών λοιμώξεων. Τα μοντέλα πολλαπλών καταστάσεων γενικεύουν τα μοντέλα των ανταγωνιστικών κινδύνων περιγράφοντας και αναλύοντας μεταβάσεις ενδιάμεσων γεγονότων.

Πολλά σημεία από τη μεθοδολογία των ανταγωνιστικών κινδύνων μπορούν να χρησιμοποιηθούν επίσης στην ανάλυση των μοντέλων πολλαπλών καταστάσεων και οι πιθανότητες μετάβασης τους να αναλυθούν μέσω τεχνικών των ανταγωνιστικών κινδύνων.

Η δομή των καταστάσεων δεν είναι μοναδική. Οι συνήθεις δομές καταστάσεων είναι έξι και φαίνονται παρακάτω στο σχήμα 3.1.

Το μοντέλο θνητότητας(mortality model) (με καταστάσεις ζωντανός-νεκρός και μία



Σχήμα 3.1: Τυπικά μοντέλα πολλαπλών καταστάσεων

μόνο πιθανή μετάβαση) είναι το πιο απλό πιθανό μοντέλο πολλαπλών καταστάσεων, τόσο απλό ώστε δεν υπάρχει κάποιο πλεονέκτημα να τεθεί στο πλαίσιο των μοντέλων πολλαπλών καταστάσεων. Το μοντέλο ανταγωνιστικών κινδύνων (competing risks model) (με καταστάσεις ζωντανός και νεκρός λόγω της αιτίας j , ($j = 1, \dots, k$)) επεκτείνει το απλό μοντέλο θνητότητας για δεδομένα επιβίωσης στα οποία κάθε άτομο μπορεί να πεθάνει λόγω διάφορων αιτιών. Το μοντέλο εναλλασσόμενων καταστάσεων (alternating state model) σχετίζεται με αναστρέψιμες ασθένειες όπου τα άτομα μετακινούνται μπρος-πίσω μεταξύ υγείας και ασθένειας (σε αυτήν την περίπτωση προτιμούμε ένα σύνηθες μοντέλο κινδύνου). Το μοντέλο ανικανότητας (disability model) (με καταστάσεις υγιής, άρρωστος, νεκρός) σχετίζεται με μη αναστρέψιμες ασθένειες και συγκεκριμένα όταν η ασθένεια αυξάνει το κίνδυνο θανάτου. Το διμεταβλητό μοντέλο (bivariate model) είναι το μοντέλο πολλαπλών καταστάσεων για διμεταβλητά παράλληλα δεδομένα με καταστάσεις: και οι δύο ζωντανοί, το 1ο άτομο νεκρός, το 2ο άτομο νεκρός και οι 2 νεκροί. Εναλλακτικά αυτή η περίπτωση μπορεί να μοντελοποιηθεί και με ένα σύνηθες μοντέλο κινδύνου. Και το μοντέλο επαναλαμβανόμενων γεγονότων είναι το μοντέλο το οποίο περιγράφει τους πληθυσμούς οι οποίοι επαναλαμβάνουν το ίδιο γεγονός αρκετές φορές (longitudinal data). Και αυτό μπορεί να περιγραφεί και από ένα μοντέλο πολλαπλών καταστάσεων (με τις καταστάσεις να περιγράφουν τον αθροιστικό αριθμό των γεγονότων) και από ένα σύνηθες μοντέλο κινδύνου.

Ένα επιπλέον σημαντικό στοιχείο όσον αφορά την δομή των καταστάσεων είναι ότι η γραφική διάταξη είναι σημαντική. Συγκεκριμένα, για πιο πολύπλοκα μοντέλα είναι σημαντικό να αναλογισθούμε πως απεικονίζεται το μοντέλο ώστε να αναδειχθούν με τον καλύτερο τρόπο τα σημεία-κλειδιά. Για παράδειγμα, ένας συνήθης κανόνας είναι ότι οι απλές μεταβάσεις εμφανίζονται είτε δεξιά είτε κάτω ώστε -αν υπάρχει φυσική κατεύθυνση του χρόνου- να συμβαδίζει με την κατεύθυνση του διαβάσματος.

Όταν χρειάζεται, οι καταστάσεις αριθμούνται: $1, \dots, S$. Σε μερικές περιπτώσεις, όπως τα επαναλαμβανόμενα γεγονότα, είναι πιο βολικό να δηλώσουμε τις καταστάσεις $0, \dots$. Σε κάποιες περιπτώσεις είναι πιθανός ο μη πεπερασμένος αριθμός των καταστάσεων, απαιτείται όμως να υπάρχει πεπερασμένος αριθμός πιθανών μεταβάσεων από κάθε κατάσταση. Αν όλες οι διαδικασίες ξεκινούν από την ίδια κατάσταση είναι βολικό να οριστεί η κατάσταση αυτή ως κατάσταση 1.

Μια κατάσταση μπορεί να είναι απορροφητική (absorbing), δηλαδή να μην υπάρχει

καμιά μετάβαση από την κατάσταση αυτή σε άλλη. Έτσι μετά την είσοδο στην κατάσταση αυτή η διαδικασία παραμένει σε αυτήν για πάντα. Υπάρχουν δύο περιπτώσεις απορροφητικών καταστάσεων. Υπάρχουν καταστάσεις που είναι από την φύση τους απορροφητικές, χαρακτηριστικό παράδειγμα: η κατάσταση νεκρός. Και άλλες που είναι απορροφητικές λόγω της δομής της μελέτης, για παράδειγμα αν αποφασισθεί ότι αν συμβεί κάποιο συγκεκριμένο γεγονός και δεν θέλουμε να ληφθεί υπ'όψιν στη μελέτη. Αλλιώς οι καταστάσεις λέγονται μεταβατικές.

Οι συμμεταβλητές μπορούν να ενσωματωθούν στους κινδύνους μετάβασης για να εξηγήσουν διαφορές μεταξύ των ατόμων κατά τη διάρκεια της θεραπείας. Συχνά, ενδιαμέσα γεγονότα αλλάζουν την φυσική πορεία της προόδου της ασθένειας, έτσι ώστε ο ρόλος κάποιων προγνωστικών παραγόντων να μην παραμένει ίδιος μετά το γεγονός αυτό. Επίσης, προγνωστικοί παράγοντες που συνδέονται με το ποσοστό θανάτου μπορεί να διαφέρουν σε διαφορετικές μεταβατικές καταστάσεις.

Τα μοντέλα τα οποία έχουν μελετηθεί περισσότερο από άποψη πιθανοτήτων είναι τα Μαρκοβιανά μοντέλα. Η Μαρκοβιανή υπόθεση σημαίνει ότι η παρούσα κατάσταση περιέχει όλη την πληροφορία και αυτός είναι ο λόγος για τον οποίο η γραφική προσέγγιση είναι σημαντική για τα Μαρκοβιανά μοντέλα αφού μπορεί να δει κανείς άμεσα την ιστορία. Κάτι το οποίο δεν μπορεί να επεκταθεί και σε γενικότερα μοντέλα, κι έτσι υπάρχουν μοντέλα μεταξύ των Μαρκοβιανών και των γενικότερων μοντέλων, τα επεκτεταμένα Μαρκοβιανά μοντέλα.

Η συνάρτηση κινδύνου (hazard function) μελετάει τον κίνδυνο(ένταση) της επόμενης μετάβασης. Υπάρχει μια συνάρτηση κινδύνου για κάθε μία δυνατή μετάβαση. Ο κίνδυνος της μετάβασης από την κατάσταση m στην κατάσταση l ορίζεται ως:

$$\lambda_{ml}(t|X_u, u \in [0, t)) = \lim_{\Delta t \searrow 0} \frac{P(X_{t+\Delta t} = l | F_{t-}, X_{t-} = m)}{\Delta t} \quad (3.1)$$

Ο συνολικός κίνδυνος από την κατάσταση m δίνεται από το άθροισμα όλων των άλλων καταστάσεων ($l \neq m$) και συμβολίζεται ομοίως χρησιμοποιώντας μόνο έναν δείκτη, δηλαδή $\lambda_m = \sum_{l \neq m} \lambda_{ml}$. Έτσι, είναι εύκολο να εκτιμήσουμε την πιθανότητα να μην συμβεί κανένα γεγονός κατά τη διάρκεια μιας περιόδου, αλλά η εκτίμηση του ακριβούς αριθμού και των τύπων των γεγονότων είναι αρκετά πιο περίπλοκη. Γι αυτό, πρέπει να εξεταστούν οι πιθανότητες μετάβασης.

Οι πιθανότητες μετάβασης είναι το κλειδί για να γίνουν μακροπρόθεσμες προ-

βλέψεις.

Τα πιο κοινά μοντέλα πολλαπλών καταστάσεων είναι οι μη ομογενείς χρονικά Μαρκοβιανές διαδικασίες με πεπερασμένο χώρο καταστάσεων.

Εάν όμως η ετερογένεια δεν αφορά τον χρόνο αλλά τον πληθυσμό, π.χ. υπάρχουν κάποιες γενετικές διαφορές ανάμεσα στον εξεταζόμενο πληθυσμό που δεν μπορούν πρακτικά να μετρηθούν, τότε για να υπολογιστεί αυτή η μη παρατηρούμενη ετερογένεια εισάγεται μια λανθάνουσα επίδραση (τ.μ.) Z , η οποία καλείται ευπάθεια και οδηγούμαστε στα μοντέλα ευπάθειας (δεν θα επεκταθούμε περαιτέρω σε αυτά).

3.1 Επιλογή δομής καταστάσεων

Επιλέγοντας τον καλύτερο τρόπο για να οριστούν οι καταστάσεις και οι μεταβάσεις δημιουργούνται διάφορα πλεονεκτήματα. Όπως το ότι γίνονται περισσότερο ξεκάθαρες οι υποθέσεις του μοντέλου, οι απλοποιήσεις κάποιων υπολογισμών και πιθανόν η τροπή ενός μη Μαρκοβιανού μοντέλου σε Μαρκοβιανό. Λόγω της διαισθητικής γραφικής κατανόησης ενός Μαρκοβιανού μοντέλου πολλαπλών καταστάσεων είναι μεγάλο πλεονέκτημα αν είναι δυνατή η διατύπωση του μοντέλου ως Μαρκοβιανό.

Πρέπει να εξετάζεται ποια είναι τα πιθανά γεγονότα. Στο ελάχιστο μοντέλο υπάρχει μόνο μία μετάβαση για κάθε τύπο γεγονότος. Η έννοια του τύπου ενός γεγονότος (εμφάνιση ασθένειας, επιπλοκή ή θάνατος) δεν είναι απαραίτητο να ορισθεί καλά.

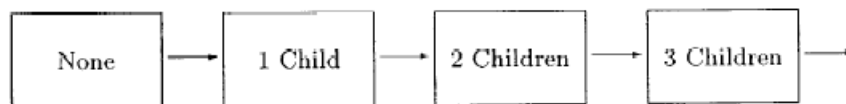
Από τη δομή των καταστάσεων μπορούμε να δούμε αν η διαδικασία είναι προοδευτική, δηλαδή αν κάθε κατάσταση έχει μόνο μία απλή πιθανή μετάβαση προς αυτήν εκτός από την αρχική κατάσταση που δεν έχει καμία είσοδο. Δηλαδή είναι πιθανό γνωρίζοντας την κατάσταση σε μια χρονική στιγμή να μπορούμε να προσδιορίσουμε ποιες καταστάσεις και με ποια σειρά επισκέφθηκε προηγουμένως. Αυτό έχει σαν πλεονέκτημα να απλοποιούνται κάποιοι υπολογισμοί. Στο μοντέλο θνητότητας, σπάζοντας την κατάσταση ζωντανός σε δύο μεταβατικές καταστάσεις λαμβάνεται το πιο απλό προοδευτικό μοντέλο τριών καταστάσεων. Και τα δύο μοντέλα είναι ειδικές περιπτώσεις του κ -προοδευτικού μοντέλου. Μεταξύ των σύννητων μοντέλων, το μοντέλο θνητότητας, το μοντέλο ανταγωνιστικών κινδύνων και το μοντέλο επαναλαμβανόμενων γεγονότων είναι προοδευτικά μοντέλα. Το μοντέλο ανικανότητας και

το διμεταβλητό μοντέλο δεν είναι προοδευτικά, αλλά μπορούν εύκολα να επεκταθούν και να γίνουν προοδευτικά, προσθέτοντας μια επιπλέον κατάσταση. Το μοντέλο εναλλασσόμενων καταστάσεων δεν είναι προοδευτικό, αφού υπάρχει είσοδος στην αρχική κατάσταση. Αυτό σημαίνει ότι αν η διαδικασία βρίσκεται στην αρχική κατάσταση δεν προκύπτει ο αριθμός των προηγούμενων μεταβάσεων, αλλά μόνο ότι ο αριθμός είναι μονός. Είναι δυνατόν να γίνει η διαδικασία προοδευτική, αλλά απαιτείται μεγάλος αριθμός καταστάσεων.

Ακολουθούν κάποια παραδείγματα συγκεκριμένων μοντέλων.

3.1.1 Μοντέλο γονιμότητας

Το μοντέλο γονιμότητας που απεικονίζεται παρακάτω είναι ένα τυπικό μοντέλο επαναλαμβανόμενων γεγονότων και περιγράφει την αναπαραγωγική ιστορία μιας γυναίκας. Η φυσική αρίθμηση των καταστάσεων είναι $0, 1, \dots$ αντίστοιχα με τον τωρινό αριθμό



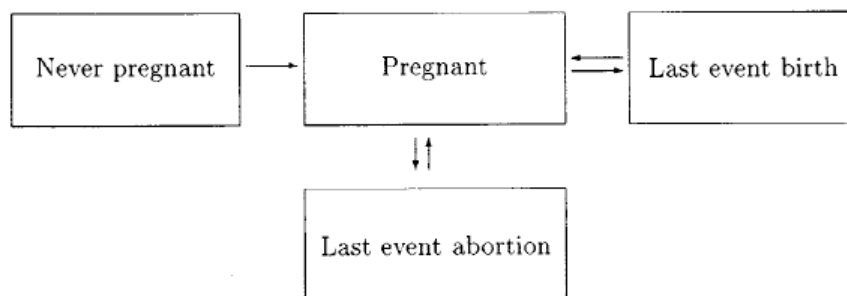
Σχήμα 3.2: Μοντέλο γονιμότητας για μια γυναίκα

των παιδιών, έτσι πιθανή λογοχρυσία συμβολίζεται με -1 . Η φυσική χρονική κλίμακα είναι η ηλικία της γυναίκας. Ένα Μαρκοβιανό μοντέλο υποστηρίζει ότι η πιθανότητα απόκτησης ενός παιδιού εξαρτάται μόνο από το πόσα παιδιά έχει ήδη η γυναίκα και την ηλικία της.

Σε αυτό το μοντέλο, η κατάσταση ορίζεται από τον προηγούμενο αριθμό των γεγονότων. Αυτό είναι ένα πλεονέκτημα των μοντέλων πολλαπλών καταστάσεων για επαναλαμβανόμενα δεδομένα, επειδή σε μερικές περιπτώσεις δεν ενδιαφερόμαστε για την πραγματική κατάσταση (π.χ. ο συνολικός αριθμός γεγονότων), αλλά μόνο για τις μεταβάσεις.

Το πιο απλό μοντέλο είναι το μοντέλο του Poisson, όπου ο «κίνδυνος» της γέννησης είναι $\lambda(t)$, κατά την ηλικία t , ανεξάρτητα από τον τωρινό αριθμό των παιδιών. Για να λειτουργήσει καλά η προσέγγιση των μοντέλων πολλαπλών καταστάσεων, πρέπει να χρησιμοποιηθεί το μοντέλο που απεικονίζεται στην εικόνα παραπάνω μαζί όμως με την υπόθεση του ίδιου «κινδύνου» σε όλες τις περιπτώσεις.

Ένα περισσότερο βιολογικό μοντέλο είναι ένα εναλλασσόμενο μοντέλο με καταστάσεις «έγκυος» και «όχι έγκυος». Αυτό το μοντέλο έχει κάποια πλεονεκτήματα, για παράδειγμα, όλοι οι τοκετοί αντιστοιχούν σε κάποια μετάβαση και είναι περισσότερο εμφανές ότι το διάστημα μεταξύ επιτυχημένων τοκετών είναι τουλάχιστον 9 μήνες. Όμως, εφαρμογή του παραπάνω απαιτεί την γνώση της κατάστασης της εγκυμοσύνης. Ένα επιπλέον πρόβλημα είναι ότι ακόμα κι αν υπάρχει η δυνατότητα να αποκτηθεί η παραπάνω γνώση, δεν αντιστοιχούν όλες οι μεταβάσεις από «έγκυος» σε «όχι έγκυος» σε γέννηση ενός υγιούς παιδιού. Η μετάβαση θα μπορούσε να αντιστοιχούσε σε μια ακούσια αποβολή. Δεν εξετάζεται το ενδεχόμενο ηθελημένης αποβολής, επειδή οι αποφάσεις αυτές ακολουθούν διαφορετικό πρότυπο. Το μοντέλο αυτό δεν μπορεί να θεωρηθεί μοντέλο πολλαπλών καταστάσεων, επειδή υπάρχουν 2 διαφορετικές μεταβάσεις από την κατάσταση «έγκυος» προς την κατάσταση «όχι έγκυος» (οι 2 μεταβάσεις είναι για την γέννα και για την αποβολή). Όμως μπορεί να μετατραπεί σε μοντέλο πολλαπλών καταστάσεων όπως απεικονίζεται παρακάτω, όπου κάθε κατάσταση αντιστοιχεί σε ένα γεγονός(αποβολή, γέννα, σύλληψη).

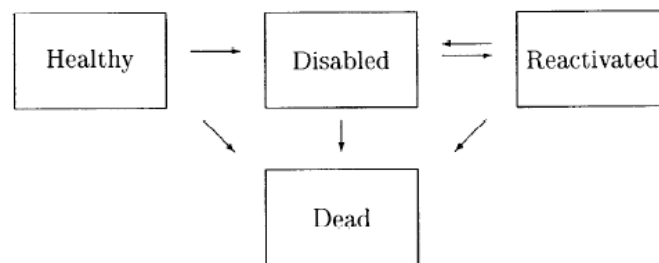


Σχήμα 3.3: Μοντέλο γονιμότητας που λογίζει το αποτέλεσμα της τελευταίας εγκυμοσύνης

Το συμπέρασμα που προκύπτει από την παραπάνω μικρή ανάλυση είναι ότι επιλέγοντας την σωστή δομή καταστάσεων, λαμβάνονται διάφορα χαρακτηριστικά, και συγκεκριμένα δημιουργείται μια διαισθητική γραφική απεικόνιση του μοντέλου οδηγώντας τις υποθέσεις να γίνουν περισσότερο φανερές. Έτσι, είναι χρήσιμο να εξεταστεί το μοντέλο λεπτομερώς πριν γίνουν οι πραγματικοί υπολογισμοί.

3.1.2 Μοντέλο ανικανότητας (Disability model)

Ονομάζεται αλλιώς και μοντέλο ασθένειας-θανάτου. Όπως είδαμε από την απεικόνιση του μοντέλου στην εισαγωγή του κεφαλαίου αυτού δεν είναι δυνατή η επανερгоποίηση, δηλαδή να επιστρέψει κανείς στην κατάσταση «υγιής» από την κατάσταση «ασθενής». Χρησιμοποιείται για να μελετηθεί η εμφάνιση της ασθένειας και ο ρυθμός του θανάτου, ενώ εκτιμάται επίσης και το κατά πόσον όσοι προηγουμένως είχαν ασθενήσει έχουν τον ίδιο κίνδυνο θανάτου με όσους ήταν υγιείς σε όλη τους τη ζωή. Αυτό μπορεί να μοντελοποιηθεί αφήνοντας τον κίνδυνο θανάτου να εξαρτάται από το αν η ιστορία της διαδικασίας περιέχει ασθένεια. Αυτό επεξηγείται καλύτερα με την παρακάτω απεικόνιση του μοντέλου.



Σχήμα 3.4: Μοντέλο ανικανότητας με την επανερгоποίηση ως ξεχωριστή κατάσταση

Σε αυτό το μοντέλο είναι εμφανές ότι ο κίνδυνος θανάτου για όσους είχαν προηγουμένως ασθενήσει είναι πιθανώς διαφορετικός από τον κίνδυνο θανάτου για αυτούς που δεν είχαν ασθενήσει ποτέ. Αυτό το χαρακτηριστικό μπορεί να ληφθεί κάτω από την υπόθεση ότι πρόκειται για ένα Μαρκοβιανό μοντέλο. Το μοντέλο αυτό μπορεί να δώσει για παράδειγμα κάποια εικόνα για την επιληψία, την είσοδο σε ψυχιατρική κλινική, τον καρκίνο, την έμμηνο ρύση. Ένα κλασσικό παράδειγμα στη δημογραφία είναι η οικογενειακή κατάσταση. Σε όλες τις περιπτώσεις χρησιμοποιείται ένα εντελώς διαφορετικό σύνολο από παραμέτρους για το πρώτο γεγονός σε σχέση με τα επόμενα γεγονότα. Για παράδειγμα, το μοντέλο για την πρώτη έμμηνο ρύση προσδιορίζεται από την ηλικία ενώ οι επόμενες έμμηνες ρύσες προσδιορίζονται από τον χρόνο του τελευταίου γεγονότος.

Για το μοντέλο ανικανότητας οι πιθανότητες μετάβασης υπολογίζονται από τις συναρ-

τήσεις κινδύνου και τις διαφορικές εξισώσεις του Kolmogorov (Cox,DR.;Miller,HD.)[9] και εκφράζονται μέσω των παρακάτω εξισώσεων:

$$p_{11}(s, t) = e^{-(\Lambda_{12}(s,t)+\Lambda_{13}(s,t))} \quad (3.2)$$

$$p_{22}(s, t) = e^{-(\Lambda_{23}(s,t))} \quad (3.3)$$

$$p_{12}(s, t) = \int_s^t p_{11}(s, u) \lambda_{12}(u) p_{22}(u, t) du \quad (3.4)$$

όπου $\Lambda_{hj}(s, t) = \int_s^t \lambda_{hj}(u) du$ είναι ο αθροιστικός κίνδυνος μεταξύ s και t .

Εάν πρόκειται για χρονικά ομογενή Μαρκοβιανή αλυσίδα οι τρεις παραπάνω εξισώσεις γίνονται:

$$p_{11}(t) = e^{-(\lambda_{12}+\lambda_{13})t} \quad (3.5)$$

$$p_{22}(t) = e^{-\lambda_{23}t} \quad (3.6)$$

$$p_{12}(t) = \frac{\lambda_{12}(e^{-\lambda_{23}t} - e^{-(\lambda_{12}+\lambda_{13})t})}{\lambda_{12} + \lambda_{13} - \lambda_{23}} \quad (3.7)$$

Επίσης οι πιθανότητες μετάβασης αυτές μπορούν να εκτιμηθούν χρησιμοποιώντας τον εκτιμητή Kaplan-Meier:

$$\hat{p}_{11}(s, t) = \prod_{s < t_{(k)} \leq t} \left(1 - \frac{d_{12k} + d_{13k}}{n_{1k}}\right) \quad (3.8)$$

$$\hat{p}_{22}(s, t) = \prod_{s < t_{(k)} \leq t} \left(1 - \frac{d_{23k}}{n_{2k}}\right) \quad (3.9)$$

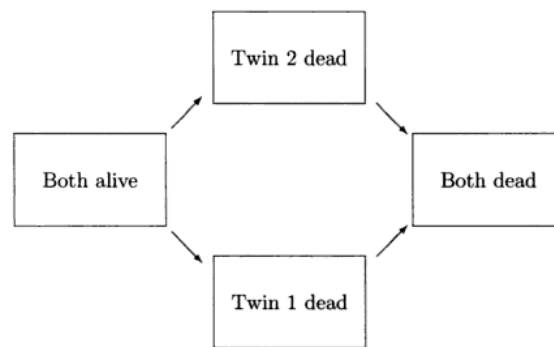
$$\hat{p}_{12}(s, t) = \sum_{s < t_{(k)} \leq t} \hat{p}_{11}(s, t_{(k-1)}) \frac{d_{13k}}{n_{1k}} \hat{p}_{22}(t_{(k)}, t) \quad (3.10)$$

όπου $t_{(1)} < t_{(2)} < \dots < t_{(d)}$ είναι οι χρόνοι των γεγονότων(είτε ασθένεια είτε θάνατος) ταξινομημένοι κατά αύξουσα σειρά, n_{1k} και n_{2k} ο αριθμός των υγιών και ασθενών ατόμων, αντίστοιχα, ακριβώς πριν το γεγονός $t_{(k)}$. Με d_{12k} δηλώνεται ο αριθμός των ατόμων που αρρώστησαν την στιγμή $t_{(k)}$, ενώ d_{13k} και d_{23k} ο αριθμός των υγιών και ασθενών ατόμων που πέθαναν την στιγμή $t_{(k)}$ αντίστοιχα.

3.1.3 Διμεταβλητά δεδομένα

Η προσέγγιση των πολλαπλών καταστάσεων καλύπτει την μοντελοποίηση πολλών ατόμων από κοινού. Κάθε συνεχής πολυμεταβλητή κατανομή μπορεί να περιγραφεί από ένα μοντέλο πολλαπλών καταστάσεων, με ένα μεγάλο περιορισμό, το μοτίβο λογοκρισίας. Αυτό εξηγείται για τα διμεταβλητά δεδομένα. Ο περιορισμός είναι ότι το μοντέλο πολλαπλών καταστάσεων απαιτεί ομογενή λογοκρισία, έτσι ώστε αν το ένα άτομο λογοκριθεί, τότε και το άλλο άτομο του ζευγαριού πρέπει να λογοκριθεί την ίδια στιγμή. Βασικά, το ζήτημα είναι αν ένα ζευγάρι πρέπει να θεωρείται ως δύο διαφορετικά άτομα (παράλληλη περίπτωση) ή ως μια μονή διαδικασία για το ζευγάρι (επαναλαμβανόμενη περίπτωση). Τα μοντέλα επαναλαμβανόμενων καταστάσεων συνθέτουν την επαναλαμβανόμενη προσέγγιση. Συγκεκριμένα, θα μελετηθεί η διάρκεια ζωής ενός ζευγαριού διδύμων.

Έστω μια συνεχής διμεταβλητή παράλληλη τυχαία μεταβλητή (W_1, W_2) , έτσι ώστε W_1 να είναι ο χρόνος επιβίωσης του δίδυμου 1 και W_2 ο χρόνος επιβίωσης του δίδυμου 2. Αυτός ο συμβολισμός κάνει εμφανές το αν χρησιμοποιείται η παράλληλη διατύπωση (W) ή η επαναλαμβανόμενη διατύπωση (T). Η διμεταβλητή συνάρτηση επιβίωσης είναι $S(w_1, w_2)$. Αυτό εφαρμόζεται σε ένα μοντέλο πολλαπλών καταστάσεων όπως φαίνεται παρακάτω:



Σχήμα 3.5: Μοντέλο επιβίωσης διδύμων

Η διαδικασία είναι στην κατάσταση 1 την χρονική στιγμή t , όταν και τα δύο δίδυμα είναι ζωντανά. Αυτό είναι το ζεύγος $\{W_1 > t, W_2 > t\}$, το οποίο συμβαίνει με πιθανότητα $S(t, t)$. Η διαδικασία είναι στην κατάσταση 2 -το δίδυμο 1 πέθανε- για

το ζεύγος $\{W_1 \leq t, W_2 > t\}$, το οποίο συμβαίνει με πιθανότητα $S(0, t) - S(t, t)$. Η διαδικασία είναι στην κατάσταση 3 -το δίδυμο 2 πέθανε- για το ζεύγος $\{W_1 > t, W_2 \leq t\}$, το οποίο συμβαίνει με πιθανότητα $S(t, 0) - S(t, t)$. Η διαδικασία είναι στην κατάσταση 4 -και τα δύο δίδυμα πέθαναν- για το ζεύγος $\{W_1 \leq t, W_2 \leq t\}$, το οποίο συμβαίνει με πιθανότητα $1 - S(0, t) - S(t, 0) + S(t, t)$. Αυτό δείχνει ότι οι πιθανότητες μετάβασης πολλαπλών καταστάσεων είναι καλά ορισμένες από την συνάρτηση επιβίωσης. Η συνάρτηση κινδύνου για το θάνατο του διδύμου 1 όταν και τα δύο δίδυμα είναι ζωντανά είναι

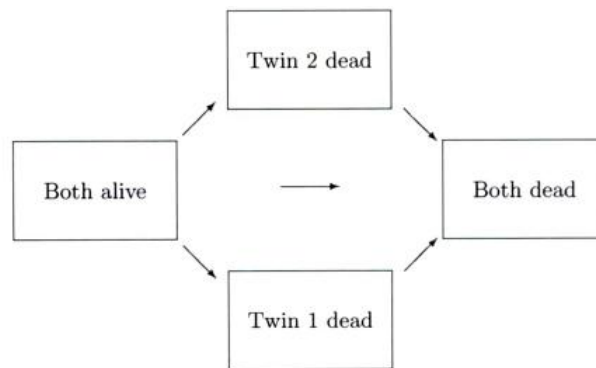
$$\lambda_1(t) = -\{D_1 \log S(t_1, t_2)\}(t, t), \quad (3.11)$$

όπου D_1 η παράγωγος ως προς την πρώτη συντεταγμένη. Ο κίνδυνος θανάτου για το δίδυμο 1, όταν το δίδυμο 2 είναι γνωστό ότι έχει πεθάνει την χρονική στιγμή t_2 είναι

$$\mu(t, t_2) = -\{D_1 \log - D_2 S(t_1, t_2)\}(t, t_2). \quad (3.12)$$

Η αντίστοιχη έκφραση για τον θάνατο του διδύμου 2 είναι ανάλογη.

Το μοντέλο μπορεί να επεκταθεί ώστε να επιτρέπει και ταυτόχρονα γεγονότα, όπως είναι ο θάνατος και των δύο διδύμων την ίδια χρονική στιγμή. Αυτό δηλώνει ότι $P(W_1 = W_2) > 0$. Σε αυτή την περίπτωση, η διμεταβλητή κατανομή δεν είναι πια συνεχής και το μοντέλο θα μοιάζει με το παρακάτω:



Σχήμα 3.6: Μοντέλο επιβίωσης διδύμων με ταυτόχρονα γεγονότα

Έτσι το παράλληλο μοντέλο το θεωρεί ως δύο γεγονότα, ενώ η στοχαστική

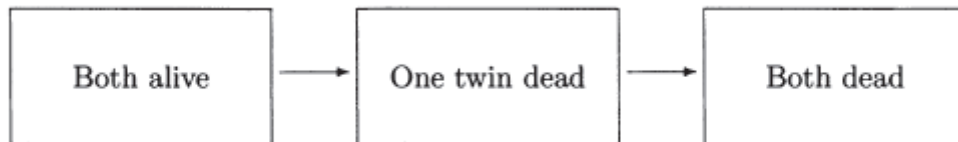
διαδικασία το θεωρεί ως μια απλή μετάβαση, αλλά μεταξύ πιο απομακρυσμένων καταστάσεων. Οι εκτιμήσεις για τα παραπάνω ζεύγη και τις συναρτήσεις επιβίωσης παραμένουν σωστές, αλλά οι συναρτήσεις κινδύνου είναι διαφορετικές. Η συνάρτηση κινδύνου του θανάτου για οποιοδήποτε από τα δίδυμα είναι $\frac{d \log S(t,t)}{dt}$. Ο συνολικός κίνδυνος θανάτου για το δίδυμο 1 περιγράφεται από την (3.15). Έτσι για να βρεθεί ο κίνδυνος θανάτου για το δίδυμο 1 μόνο, όταν και τα δύο δίδυμα είναι ζωντανά, χρησιμοποιείται ο συνολικός κίνδυνος μείον τον κίνδυνο θανάτου του δίδυμου 2, δηλαδή

$$\lambda_1(t) = -\frac{d \log S(t,t)}{dt} + \{D_2 \log S(t_1, t_2)\}(t, t).$$

Ο κίνδυνος του ταυτόχρονου γεγονότος είναι

$$\lambda_0(t) = -\frac{d \log S(t,t)}{dt} + \lambda_1(t) + \lambda_2(t).$$

Ακόμα ένας τρόπος να μελετηθεί το μοντέλο του σχήματος 3.5 είναι ο παρακάτω:



Σχήμα 3.7: Μοντέλο επιβίωσης διδύμων βασισμένο σε διατεταγμένες παρατηρήσεις

Με αυτή τη μορφή δεν καταγράφεται ποιο από τα δύο δίδυμα πεθαίνει πρώτο. Από στατιστική άποψη, αντιστοιχεί σε μείωση των δεδομένων, η οποία είναι επαρκής μόνο κάτω από την υπόθεση της συμμετρίας. Το σχήμα παραπάνω κάνει το μοντέλο να μοιάζει με πρόβλημα επαναλαμβανόμενων γεγονότων, για παράδειγμα όπως το μοντέλο γονιμότητας του σχήματος 3.2. Όμως υπάρχουν δυο κύριες διαφορές, λόγω της υπόθεσης της ανεξαρτησίας ο κίνδυνος της πρώτης μετάβασης ισούται με τον διπλάσιο κίνδυνο της δεύτερης μετάβασης. Στο πρόβλημα των επαναλαμβανόμενων γεγονότων, η υπόθεση των γεγονότων ότι συμβαίνουν ανεξάρτητα αντιστοιχεί στις δύο συναρτήσεις κινδύνου εδώ, οι οποίες για αυτό το λόγο είναι ακριβώς ίδιες. Επίσης, ο μέγιστος αριθμός των γεγονότων στο μοντέλο διδύμων είναι δύο ενώ δεν υπάρχει κάποιο όριο στα επαναλαμβανόμενα δεδομένα. Δηλαδή, υπάρχει μια απορροφητική κατάσταση στο μοντέλο διδύμων, η κατάσταση όπου και τα δύο δίδυμα είναι

νεκρά. Μια επιπλέον διαφορά είναι ότι στην περίπτωση των διδύμων είναι γνωστό ότι υπάρχουν δύο διαφορετικά αντικείμενα.

Αυτοί οι υπολογισμοί μπορούν να γίνουν σε όλα τα διμεταβλητά μοντέλα, όπως και στα μοντέλα επαναλαμβανόμενων μετρήσεων. Όμως, επειδή οι φυσικοί χρόνοι δεν είναι πάντα οι ίδιοι, δεν έχει τόσο νόημα από τη σκοπιά της πρόβλεψης. Γι αυτό το λόγο τα μοντέλα πολλαπλών καταστάσεων δεν θα πρέπει να θεωρούνται για δεδομένα επαναλαμβανόμενων μετρήσεων.

Το ίδιο μοντέλο μπορεί να γενικευθεί σε διαστάσεις μεγαλύτερες από δύο, αλλά η δομή γίνεται αρκετά πιο πολύπλοκη.

3.1.4 Επαναλαμβανόμενα γεγονότα

Αν το ενδιαφέρον εστιάζεται σε επαναλαμβανόμενα συμβάντα ενός συγκεκριμένου γεγονότος π.χ. εισαγωγή σε νοσοκομείο, γεννήσεις παιδιών, λοιμώξεις κ.ά. τότε μπορεί να θεωρηθεί το μοντέλο που περιγράφεται στο σχήμα 3.1 ως μοντέλο επαναλαμβανόμενων γεγονότων. Το γενικό μοντέλο επαναλαμβανόμενων γεγονότων έχει κίνδυνο $\lambda_1(t)$ για την πρώτη μετάβαση, $\lambda_2(t, t_1)$ για το δεύτερο γεγονός, όπου t_1 είναι ο χρόνος της πρώτης μετάβασης, $\lambda_3(t, t_1, t_2)$ για το τρίτο γεγονός, κτλ. Γενικά, θα είναι της μορφής $\lambda_j(t, t_1, \dots, t_{j-1})$ για το j -οστό γεγονός. Η συνάρτηση έχει νόημα μόνο όταν $0 < t_1 < t_2 < \dots < t_{j-1} < t$. Το μοντέλο $\lambda_1(t)$ στο οποίο ο κίνδυνος εξαρτάται μόνο από το t είναι ένα Poisson μοντέλο. Το Μαρκοβιανό μοντέλο μπορεί να προκύψει όταν ο κίνδυνος εξαρτάται μόνο από τα j και t , δηλαδή $\lambda_j(t)$. Συχνά μια ενδιαφέρουσα λειτουργία του μοντέλου αυτού είναι η εύρεση του αναμενόμενου αριθμού των συμβάντων ενός γεγονότος σε ένα διάστημα $[0, t]$.

3.2 Προοδευτικά μοντέλα

Ο απλούστερος ορισμός για ένα προοδευτικό μοντέλο είναι ότι όλες οι καταστάσεις, εκτός από την αρχική κατάσταση, έχουν μόνο μία πιθανή μετάβαση στην κατάσταση. Η αρχική κατάσταση δεν έχει καμία πιθανή μετάβαση σε αυτήν. Έχουμε δει ήδη τέτοιο μοντέλο, π.χ. το μοντέλο γονιμότητας για μια γυναίκα του σχήματος 3.2. Ενώ τα μοντέλα των σχημάτων 3.4-3.6 δεν είναι προοδευτικά αφού υπάρχουν καταστάσεις με πάνω από μία μετάβαση σε αυτές όπως και το μοντέλο εναλλασσόμενων καταστάσεων.

Αλλά με κάποιες τροποποιήσεις μπορούν και αυτά να γίνουν προοδευτικά. Σε ένα προοδευτικό μοντέλο, η τρέχουσα κατάσταση περιέχει την πληροφορία για το πόσες και ποιες καταστάσεις έχει προηγουμένως επισκεφθεί, καθώς και τη σειρά με την οποία τις επισκέφθηκε αλλά όχι το πλήθος των μεταβάσεων.

Το πλεονέκτημα του να είναι ένα μοντέλο προοδευτικό είναι ότι οι διαφορικές εξισώσεις για τις πιθανότητες μετάβασης μπορούν να απλοποιηθούν σε ολοκληρώματα, τα οποία είναι αρκετά πιο εύκολα να χειριστούν ακόμα και αν πρόκειται για πολλαπλά ολοκληρώματα. Ένα επιπλέον πλεονέκτημα της μελέτης προοδευτικών μοντέλων είναι ότι η πιθανοφάνεια των γενικών μοντέλων μπορεί να βασιστεί σε εκφράσεις για προοδευτικά μοντέλα.

3.2.1 Μοντέλο ανταγωνιστικών κινδύνων

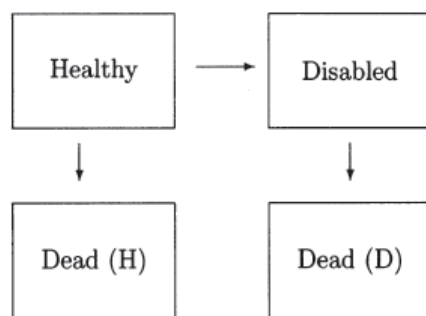
Το πιο απλό προοδευτικό μοντέλο είναι το μοντέλο ανταγωνιστικών κινδύνων όπως φαίνεται στο σχήμα 3.1. Μόνο η αρχική κατάσταση (η οποία συνήθως δηλώνει την κατάσταση «υγιής» ή «ζωντανός») δεν είναι απορροφητική. Δηλαδή, όλες οι μεταβάσεις προέρχονται από την αρχική κατάσταση και όλα τα μοντέλα είναι τετριμμένα Μαρκοβιανά μοντέλα. Μια διαδικασία μπορεί να περιγραφεί από ένα μοντέλο ανταγωνιστικών κινδύνων όταν η εμφάνιση ενός γεγονότος αλλάζει την ικανότητα να παρατηρηθεί η εκδήλωση ενός άλλου γεγονότος.

Περισσότερες λεπτομέρειες για τα μοντέλα ανταγωνιστικών κινδύνων έχουν αναφερθεί στο προηγούμενο κεφάλαιο.

3.2.2 Μοντέλο ανικανότητας

Ένα πιο πολύπλοκο μοντέλο είναι η προοδευτική εκδοχή του μοντέλου ανικανότητας, όπως φαίνεται παρακάτω στο σχήμα 3.8. Ο κίνδυνος θανάτου για ένα υγιές άτομο συμβολίζεται με $\lambda(t)$, για ένα άρρωστο άτομο συμβολίζεται με $\mu(t, t_1)$, όπου t_1 είναι ο χρόνος που αρρώστησε και ο κίνδυνος να αρρωστήσει ένα υγιές άτομο συμβολίζεται με $\psi(t)$. Η πιθανότητα του να μείνει στην υγιή κατάσταση είναι

$$P_1(t) = \exp\left\{-\int_0^t \lambda(u) + \psi(u) du\right\},$$



Σχήμα 3.8: Το μοντέλο ανικανότητας ως προοδευτικό μοντέλο

σε απόλυτη αναλογία με την (2.2). Η πιθανότητα να πεθάνει ένα άτομο χωρίς προηγουμένως να έχει αρρωστήσει, είναι

$$P_3(t) = \int_0^t \lambda(u) \exp\left[-\int_0^u \{\lambda(v) + \psi(v)\} dv\right] du,$$

μαθηματικά πανομοιότυπη με τις πιθανότητες στην περίπτωση των ανταγωνιστικών κινδύνων (2.4) και ανεξάρτητη από την έκφραση $\mu(t, t_1)$. Η πιθανότητα να βρίσκεται κανείς στην κατάσταση «άρρωστος» είναι

$$P_2(t) = \int_0^t \psi(u) \exp\left[-\int_0^u \{\lambda(v) + \psi(v)\} dv\right] \exp\left\{-\int_u^t \mu(v, u) dv\right\} du. \quad (3.13)$$

Υπάρχει ένας όρος για τη μη ύπαρξη συμβάντων την περίοδο $0 - u$, μετά ένας όρος κινδύνου την στιγμή u και ένας όρος παραμονής στην κατάσταση την περίοδο $u - t$, ο οποίος χρειάζεται να ολοκληρωθεί ως προς u . Η πιθανότητα $P_4(t)$ να είναι κάποιος στην τελική κατάσταση είναι πιο εύκολο να εκτιμηθεί ως η διαφορά της μονάδας από το άθροισμα των πιθανοτήτων όλων των άλλων καταστάσεων. Στην περίπτωση των Μαркоβιανών συναρτήσεων κινδύνου μπορεί να γραφεί ως

$$\int_0^t \int_u^t P_{11}(0, u) \psi(u) P_{22}(u, w) \mu(w) P_{44}(w, t) dw du. \quad (3.14)$$

Υπάρχουν όροι για τις συναρτήσεις κινδύνου τις στιγμές που συμβαίνουν τα γεγονότα και για τις συναρτήσεις πιθανοτήτων για μη ύπαρξη γεγονότος ενδιάμεσα. Δηλαδή,

αρχικά η διαδικασία βρίσκεται στην κατάσταση 1 από την στιγμή 0 έως u , μεταπηδά στην κατάσταση 2 (άρρωστος), παραμένει στην κατάσταση 2 μέχρι την στιγμή w , μετά μεταβαίνει στην κατάσταση 4 (νεκρός) και τέλος παραμένει στην κατάσταση 4 για το υπόλοιπο του χρόνου. Ο όρος P_{44} είναι περιττός τεχνικά αφού ισούται με 1, αλλά συμπεριλαμβάνεται ώστε να είναι πιο ξεκάθαρη η διαδικασία.

3.3 Επεκτεταμένα Μαρκοβιανά μοντέλα

Μια μικρή επέκταση των Μαρκοβιανών μοντέλων είναι όταν η συνάρτηση κινδύνου επιτρέπεται να εξαρτάται από την διάρκεια στην τρέχουσα κατάσταση. Αυτά τα μοντέλα κάποιες φορές καλούνται μοντέλα με εξάρτηση από τη διάρκεια. Οι εξισώσεις *Chapman – Kolmogorov* (1.14) δεν ισχύουν σε αυτή την περίπτωση.

Αν μελετηθούν μόνο οι δύο πρώτες μεταβάσεις σε μια διαδικασία, όλα τα μοντέλα είναι επεκτεταμένα Μαρκοβιανά μοντέλα. Έτσι το μοντέλο διμεταβλητών δεδομένων (σχήμα 3.5) είναι ένα επεκτεταμένο Μαρκοβιανό μοντέλο με οποιεσδήποτε υποθέσεις στις συναρτήσεις κινδύνου.

Έχειδειχθεί ότι τα δεδομένα γονιμότητας δεν μπορούν να περιγραφούν από ένα Μαρκοβιανό μοντέλο, αλλά φαίνεται λογικό να χρησιμοποιηθεί ένα επεκτεταμένο Μαρκοβιανό μοντέλο. Η πιο απλή πιθανότητα είναι να υποτεθεί ότι ο «κίνδυνος» μιας γέννας είναι 0 τους πρώτους εννιά μήνες μετά από μια προηγούμενη γέννα. Κάτι που είναι αρκετά απλό στο να συμπεριληφθεί στην εκτίμηση.

Σε μερικές περιπτώσεις, υπάρχει η επιλογή μεταξύ ενός μη Μαρκοβιανού μοντέλου με μια απλή δομή καταστάσεων και ενός Μαρκοβιανού μοντέλου με μια πολύπλοκη δομή καταστάσεων. Είναι επίσης πιθανό να επιδιωχθεί ώστε οι καταστάσεις να περιέχουν πληροφορία τις στιγμές των μεταβάσεων, όμως είναι αρκετά πιο πολύπλοκο. Το πλεονέκτημα του να είναι μια διαδικασία Μαρκοβιανή δεν είναι τόσο μεγάλο ώστε να καταστεί πλεονέκτημα αν εφαρμοσθούν κάποια πολύπλοκα τεχνάσματα. Έτσι, είναι καλύτερα να διατηρηθεί το αρχικό μοντέλο και να περιγραφεί ως επεκτεταμένο Μαρκοβιανό μοντέλο.

3.4 Εκτίμηση Μαρκοβιανής συναρτήσης πιθανοφάνειας

Έστω ότι υπάρχουν S καταστάσεις, αριθμημένες ως $1, \dots, S$. Αν το μοντέλο περιλαμβάνει άπειρο αριθμό καταστάσεων, μπορεί να μειωθεί το μοντέλο σε ένα μοντέλο με πεπερασμένες καταστάσεις, επειδή θεωρείται μια δεδομένη έκβαση της διαδικασίας. Όπως τα προοδευτικά μοντέλα για να μελετηθεί ένα μόνο αποτέλεσμα της διαδικασίας, είναι γνωστός ο πραγματικός αριθμός των μεταβάσεων αλλά και η σειρά τους. Ο κίνδυνος από την κατάσταση m στην κατάσταση l συμβολίζεται με $\lambda_{ml}(t)$. Σύμφωνα με την δομή καταστάσεων κάποιοι από αυτούς τους όρους θα ισούνται με 0, αλλά αυτό δεν αποτελεί πρόβλημα στην ισχύ των τύπων. Ο συνολικός κίνδυνος από την κατάσταση m είναι $\sum_{l \neq m} \lambda_{ml}(t)$ και συμβολίζεται απλά $\lambda_m(t)$. Μια μελέτη καλύπτει την χρονική περίοδο από το 0 έως το C , στην οποία υπάρχουν E γεγονότα. Η διαδικασία αρχίζει στην κατάσταση s_0 την χρονική στιγμή $T_0 = 0$. Το j -οστό γεγονός είναι μια μετάβαση στην κατάσταση $s_j, j = 1, \dots, E$. Έτσι η πιθανοφάνεια είναι

$$\left[\prod_{j=1}^E \lambda_{s_{j-1}s_j}(T_j) \exp\left\{-\int_{T_{j-1}}^{T_j} \lambda_{s_{j-1}}(u) du\right\} \right] \exp\left\{-\int_{T_E}^C \lambda_{s_E}(u) du\right\}. \quad (3.15)$$

Όμως, επειδή δεν είναι βολικό να υπάρχουν δύο εκθετικοί όροι, οι οποίοι είναι και αρκετά όμοιοι, εισάγοντας τους συμβολισμούς: $K = E + 1, T_K = C, s_K = 0$ στην περίπτωση της λογοκρισίας (και $K = E$ στην περίπτωση της ύπαρξης γεγονότος την στιγμή C), μπορεί η πιθανοφάνεια να αναδιατυπωθεί ως

$$\prod_{j=1}^K \lambda_{s_{j-1}s_j}(T_j)^{1_{\{s_j \neq 0\}}} \exp\left\{-\int_{T_{j-1}}^{T_j} \lambda_{s_{j-1}}(u) du\right\}. \quad (3.16)$$

Η πιθανοφάνεια δείχνει ότι για κάθε μετάβαση υπάρχει μια συνεισφορά που αντιστοιχεί στον κίνδυνο για εκείνο τον χρόνο και για τις περιόδους χωρίς μεταβάσεις υπάρχουν εκθετικοί όροι με το ολοκλήρωμα του κινδύνου που αντιμετωπίζει η διαδικασία, ο οποίος είναι ο κίνδυνος στην κατάσταση που βρίσκεται η διαδικασία. Εισάγουμε δείκτες $D_{mlj} = 1_{\{s_{j-1} = m, s_j = l\}}$ που είναι 1, όταν η j -οστή μετάβαση πηγαινει από την κατάσταση m στην κατάσταση l και ορίζουμε συνάρτηση κινδύνου $R_m(t)$, η

οποία είναι 1 όταν η διαδικασία είναι στην κατάσταση m την στιγμή t^- . Ο τύπος για τη συνάρτηση αυτή είναι

$$R_m(t) = \sum_{j=1}^K 1\{s_{j-1} = m, T_{j-1} < t \leq T_j\}. \quad (3.17)$$

Βάσει των παραπάνω προκύπτει ότι η πιθανοφάνεια είναι

$$\prod_m \prod_l [\{\prod_{j=1}^K \lambda_{ml}(T_j)^{D_{mlj}}\} \exp\{-\int_0^C R_m(u) \lambda_{ml}(u) du\}]. \quad (3.18)$$

Ο παραπάνω τύπος περιγράφει την πιθανοφάνεια ως ένα γινόμενο για όλες τις πιθανές μεταβάσεις και κάθε όρος εμπλέκει μόνο την συνάρτηση κινδύνου για την συγκεκριμένη μετάβαση.

3.5 Μοντέλα παλινδρόμησης πολλαπλών καταστάσεων

Σε πρακτικές εφαρμογές είναι χρήσιμο να συσχετίσουμε τα ατομικά χαρακτηριστικά με τους ρυθμούς κινδύνου μέσω ενός διάνυσματος συμμεταβλητών Z . Ένα γενικό μοντέλο παλινδρόμησης μπορεί να γραφεί ως

$$\lambda_{hji}(t) = \phi(\lambda_{jh,0}(t), \beta_{hj}^T Z_i) \quad (3.19)$$

όπου η $\lambda_{hj,0}(t)$ είναι η συνάρτηση βασικού κινδύνου μεταξύ των καταστάσεων h και j , β_{hj} το διάνυσμα των συντελεστών παλινδρόμησης και Z_i το διάνυσμα των συμμεταβλητών για το άτομο i .

Μια δημοφιλής επιλογή μοντέλου παλινδρόμησης (επιλέγεται και για χρονικά ομογενείς Μαρκοβιανές διαδικασίες) είναι το αναλογικό μοντέλο κινδύνου του Cox το οποίο αποκτάται αν επιλέξουμε $\phi(u(t), v) = u(t)e^v$ κι έτσι είναι:

$$\lambda_{hj}(t|Z) = \lambda_{hj,0}(t) \exp(\beta_{hj}^T Z) \quad (3.20)$$

Ένα εναλλακτικό μοντέλο για χρονικά εξαρτημένους συντελεστές παλινδρόμησης

προτάθηκε από τον Aalen για δεδομένα επιβίωσης και αργότερα χρησιμοποιήθηκε στα μοντέλα πολλαπλών καταστάσεων και είναι το:

$$\lambda_{hj}(t|Z) = \lambda_{hj,0}(t) + \beta_{hj}^T Z \quad (3.21)$$

Έτσι η εισαγωγή των συμμεταβλητών σε αυτά τα μοντέλα επιτρέπει την πρόβλεψη των πιθανοτήτων προσαρμοσμένες σε κάθε άτομο ξεχωριστά που αυτό ήταν ο στόχος μας.

Ας εξετάσουμε λίγο το παράδειγμα του μοντέλου γεννήσεως-θανάτου. Για τον κίνδυνο θανάτου χωρίς να αρρωστήσει το άτομο, $\lambda_{13}(t|Z)$, οι χρόνοι επιβίωσης των ασθενών που απέκτησαν την ασθένεια και αυτών που είναι ζωντανοί και δεν απέκτησαν την ασθένεια λογίζονται ως λογοκριμένες παρατηρήσεις. Το ίδιο θεωρούνται για το κίνδυνο να ασθενήσει ένα άτομο, $\lambda_{12}(t|Z)$, οι χρόνοι επιβίωσης των ατόμων που δεν ασθένησαν είτε πέθαναν τελικά είτε όχι.

Σε κάποιες περιπτώσεις υποθέτουμε ότι οι βασικοί κίνδυνοι των μεταβάσεων $1 \rightarrow 3$ και $2 \rightarrow 3$ είναι αναλογικοί. Αυτό ισοδυναμεί με το να ομαδοποιήσουμε τις δύο αυτές μεταβάσεις και να χρησιμοποιήσουμε αυτό το ενδιαμέσο γεγονός ως χρονικά εξαρτημένη συμμεταβλητή. Έτσι το μοντέλο για την μετάβαση στην κατάσταση 3 δίνεται από τις παρακάτω εξισώσεις:

$$\begin{aligned} \lambda_{13}(t|Z) &= \lambda_{3,0}(t) \exp(\beta_{13}^T Z) \\ \lambda_{23}(t|Z) &= \lambda_{3,0}(t) \exp(\beta_{23}^T Z + \delta) \end{aligned} \quad (3.22)$$

Κεφάλαιο 4

Εφαρμογή: Ανάλυση δεδομένων από ασθενείς με λοίμωξη σε ΜΕΘ

4.1 Εισαγωγή

Οι νοσοκομειακές λοιμώξεις δεν αποτελούν μόνο κίνδυνο για τη ζωή των ασθενών διακινδυνεύοντας την επιτυχία μιας εγχείρησης ή μιας θεραπείας αλλά και συντελούν στην αύξηση της παραμονής των ασθενών (prolongation of length of stay-LOS) στο νοσοκομείο και στη σχετιζόμενη υπέρβαση των ιατρικών εξόδων.

Σκοπός της μελέτης μας ήταν να εκτιμήσουμε αν επηρεάζονται και πως η θνησιμότητα και η διάρκεια παραμονής στη μονάδα εντατικής θεραπείας (ΜΕΘ) από τη πνευμονία λόγω μηχανικής αναπνοής μέσω αναπνευστήρα (ventilator associated pneumonia-VAP), τη βακτηραιμία (bloodstream infection-BSI) και την ουρολοίμωξη (urinary tract infection-UTI) που αποκτήθηκαν σε ΜΕΘ.

Υπολογίζεται ότι στην Ευρωπαϊκή Ένωση οι λοιμώξεις αυτές προσθέτουν ένα επιπλέον κόστος (>900 εκατομμύρια ευρώ) και παραγωγικές απώλειες τουλάχιστον 1.5 δισεκατομμύρια ευρώ κάθε χρόνο στις Ευρωπαϊκές χώρες. Η νοσοκομειακή πνευμονία είναι η πιο συχνή λοίμωξη που αποκτάται στα νοσοκομεία και η πιο θανατηφόρα αφού οδηγεί σε πάνω από 36.000 θανάτους ετησίως. Γι αυτό και η ανάλυση μας επικεντρώθηκε στη VAP και τις άλλες δύο λοιμώξεις τις μελετήσαμε σε συνάρτησι

με τη VAP. Τα βαριά περιστατικά ασθενών χρειάζονται παρατεταμένη μηχανική αναπνοή, η οποία είναι ο πιο σημαντικός παράγοντας κινδύνου για την ανάπτυξη της νοσοκομειακής πνευμονίας αφού ο κίνδυνος απόκτησης πνευμονίας κατά τη διάρκεια της αναπνευστικής υποστήριξης είναι ουσιαστικά μεγαλύτερος από οποιοδήποτε άλλο πληθυσμό ασθενών. Στις Ηνωμένες Πολιτείες πρόσφατα προτάθηκε η VAP ως ένας δείκτης ποιότητας της φροντίδας για τις ΜΕΘ. Προγράμματα ελέγχου λοιμώξεων είναι σημαντικά για την πρόληψη της VAP μέσω εφαρμογής στρατηγικών που τροποποιούν τους παράγοντες κινδύνου για τη VAP.

Υπάρχουν αρκετές μελέτες στη βιβλιογραφία που έχουν δείξει ότι η VAP περιπλέκει την πορεία των ασθενών στη ΜΕΘ, όμως παραμένει αμφιλεγόμενο ακόμα το ερώτημα εάν οι βαριά ή όχι ασθενείς βρίσκονται σε αυξημένο κίνδυνο θανάτου ή παρατείνεται η διάρκεια παραμονής τους στη ΜΕΘ ως αποτέλεσμα απόκτησης πνευμονίας.

Ο αυξημένος κίνδυνος θανάτου που οφείλεται στη VAP ποικίλλει σημαντικά μεταξύ των μελετών, από 0% έως και 80%. Τέσσερις κύριοι παράγοντες συντελούν σε αυτή τη μεταβλητότητα. Η διάγνωση της VAP είναι δύσκολο να αποδειχθεί με πλήρη βεβαιότητα επειδή δεν υπάρχουν παθογνωμικά ευρήματα. Άλλος λόγος είναι ότι αν αρχίσει η παροχή αντιμικροβιακών προτού παρθούν δείγματα για μικροβιολογικούς ελέγχους, τότε η ευαισθησία των διαγνωστικών τεχνικών μειώνεται δραματικά. Επίσης, ο πληθυσμός των ασθενών διαφέρει μεταξύ των μελετών αλλά και μεταξύ των χωρών λόγω των διαφορετικών συνθηκών στα συστήματα υγείας της κάθε χώρας. Επιπλέον, η επάρκεια των αντιμικροβιακών και το χρονοδιάγραμμα της αντιβιοτικής θεραπείας μπορεί να επηρεάσει τις εκτιμήσεις. Εκτός των παραπάνω, και οι στατιστικές μέθοδοι που χρησιμοποιούνται μπορούν να συντελέσουν σε αυτήν την παρατηρούμενη μεταβλητότητα στις τιμές της θνητότητας που αναλογεί στη VAP.

Ένα μοντέλο πολλαπλών καταστάσεων και συγκεκριμένα το προοδευτικό μοντέλο ανικανότητας λαμβάνει υπ'όψιν του και την χρονική εξάρτηση των παραγόντων κινδύνου (π.χ. της VAP) αλλά και την παρουσία ανταγωνιστικών κινδύνων (π.χ. θάνατος και έξοδος από τη ΜΕΘ) σε κάθε χρονικό σημείο, κάτι το οποίο δεν πετυχαίνει κάποια άλλη στατιστική μέθοδος.

4.2 Περιγραφική ανάλυση

Τα δεδομένα της μελέτης συλλέχθηκαν από την κ.Ελένη Αποστολοπούλου (αναπληρώτρια καθηγήτρια του τμήματος Νοσηλευτικής(ΕΚΠΑ)) συστηματικά από τέσσερις διαφορετικές μονάδες εντατικής θεραπείας της Αθήνας. Το συνολικό δείγμα ήταν 294 ασθενείς με μηχανική αναπνοή. Δυνητικές πηγές μεροληψίας μπορεί να υπάρξουν λόγω του ανθρώπινου παράγοντα.

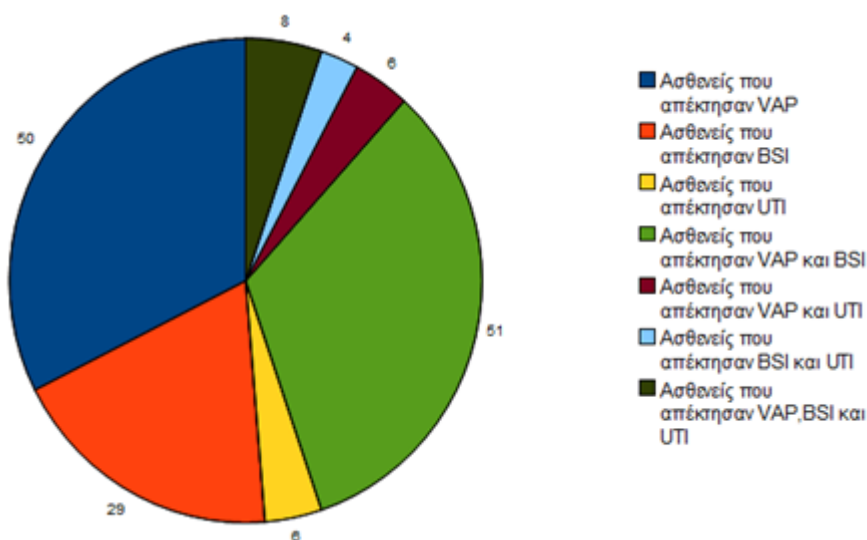
Πολλές μεταβλητές καταγράφηκαν για κάθε ασθενή, όπως το φύλο, η ηλικία, ο δείκτης βαρύτητας περιστατικού (συγκεκριμένα το chronic health evaluation II (Apache II) score το οποίο εφαρμόζεται εντός 24 ωρών από την είσοδο του ασθενή στη ΜΕΘ, είναι ένας ακέραιος αριθμός από το 0 έως το 71 και οι μεγάλες τιμές του δείκτη αυτού αντιστοιχούν σε βαρύτερες ασθένειες και μεγαλύτερο κίνδυνο θανάτου), ο τύπος του περιστατικού, η διάρκεια προηγούμενης νοσηλείας, η κύρια διάγνωση εισαγωγής, χαρακτηριστικά ασθενών κατά την εισαγωγή, ανάπτυξη λοιμώξεων και ημέρα ανάπτυξης λοιμώξεων και εξόδου από τη ΜΕΘ. (Πλήρης κατάλογος των καταγραφόμενων μεταβλητών υπάρχει στο Παράρτημα Α')

Είχαμε στη διάθεση μας και στοιχεία για τις μελετούμενες λοιμώξεις (VAP, BSI, UTI) που προκαλούνται όμως από ανθεκτικά στις Carbanemes παθογόνα. Η ανάλυση των στοιχείων αυτών δεν ήταν μέσα στις προτεραιότητες μας κι έτσι θα μπορούσε να αποτελέσει την συνέχεια της παρούσας μελέτης.

Μελέτες που διεξήχθησαν στην υπόλοιπη Ευρώπη, και όχι μόνο, αφορούσαν συνήθως χιλιάδες ασθενείς και συμπεριλαμβάνονταν στη μελέτη όσοι ασθενείς λάμβαναν μηχανική αναπνοή και παρέμεναν στη ΜΕΘ για τουλάχιστον 48 ώρες. Λόγω αυτού, λοιμώξεις που αποκτούνταν από τους ασθενείς νωρίτερα από τις πρώτες 48 ώρες παραμονής των ασθενών στη ΜΕΘ μπορεί να μη καταγράφονταν. Εμείς δεν διαθέταμε τον χρονικό περιορισμό αυτόν για το ποιοι ασθενείς θα συμπεριληφθούν στη μελέτη αλλά ούτε και την πολυτέλεια των χιλιάδων ασθενών, άρα και μεγαλύτερου πληθυσμού που θα μας προσέφερε περισσότερη ευελιξία στην ανάλυση και πιο ακριβή αποτελέσματα.

Στην ανάλυση μας, εξετάστηκαν 294 ασθενείς της Μονάδας Εντατικής Θεραπείας με μηχανική αναπνοή. Το πιο κοινό είδος λοίμωξης ήταν η VAP, η οποία αποκτήθηκε από 115 (39.1%) ασθενείς. Η δεύτερη πιο συχνή λοίμωξη παρατηρήθηκε η BSI, με 92 (31%) ασθενείς να προσβλήθηκαν από αυτήν. Και η UTI αποκτήθηκε από 24(8.2%) ασθενείς. Ενώ υπήρχαν 140(47.6%) ασθενείς οι οποίοι δεν παρουσίασαν καμία λοί-

μωξη κατά την παραμονή τους στη ΜΕΘ. Όμως μεγάλο μέρος του πληθυσμού μας, δεν απέκτησε μόνο μία λοίμωξη αλλά συνδυασμό αυτών, κάτι που παρατηρείται και από το γεγονός ότι το άθροισμα των παραπάνω αριθμών είναι μεγαλύτερο του πληθυσμού. Ο πιο πιθανός συνδυασμός λοιμώξεων ήταν η απόκτηση των VAP και BSI που συνέβηκε σε 51(17.3%) ασθενείς. Ενώ μόνο τη VAP και καμία άλλη λοίμωξη απέκτησαν 50(17%) ασθενείς, μόνο τη BSI 29(9.9%) ασθενείς και μόνο τη UTI 6(2%) ασθενείς. Αναλυτικά το πλήθος των ασθενών και το είδος των λοιμώξεων που απέκτησαν οι ασθενείς κατά την παραμονή τους στη ΜΕΘ φαίνεται στο Σχήμα 4.1.



Σχήμα 4.1: Πλήθος ασθενών συγκριτικά με τα είδη των λοιμώξεων που απέκτησαν

Συνολικά, 135 (45.9%) ασθενείς πέθαναν όσο βρίσκονταν στη ΜΕΘ. Από τους 50 ασθενείς που απέκτησαν μόνο τη VAP (χωρίς καμία προγενέστερη ή μεταγενέστερη λοίμωξη), οι 22(44%) ασθενείς απεβίωσαν στη ΜΕΘ, ενώ από τους 29 ασθενείς που προσβλήθηκαν μόνο από τη BSI, οι 15(51.7%) ασθενείς πέθαναν στη ΜΕΘ.

Ο Πίνακας 4.1 μας παρουσιάζει κάποια αναφορικά κλινικά και δημογραφικά χαρακτηριστικά για τον πληθυσμό των ασθενών. Βλέπουμε ότι η ασθένεια των πνευμόνων είναι η πιο κοινή κύρια διάγνωση εισαγωγής στη ΜΕΘ άρα κι αυτός είναι ένας λόγος που η VAP είναι η πιο συχνή λοίμωξη. Η διάρκεια μηχανικής αναπνοής των ασθενών που δεν απέκτησαν τη VAP είναι ελάχιστα παραπάνω από τη διάρκεια των ασθε-

νών που απέκτησαν τη VAP. Οι γυναίκες ασθενείς(164,55.8%) είναι περισσότερες από τους άντρες στον πληθυσμό μας και οι ασθενείς με παθολογική ή καρδιολογική νόσο-medical (192, 65.3%) περισσότεροι από τους ασθενείς με χειρουργική νόσο-surgical. Επίσης σχεδόν οι μισοί εξεταζόμενοι ασθενείς εισάγονται στη ΜΕΘ ενώ έχουν ήδη κάποια λοίμωξη, άρα είναι βεβαρυμένη η κατάσταση τους.

Πίνακας 4.1: Αναφορικά κλινικά και δημογραφικά χαρακτηριστικά για τον πληθυσμό των ασθενών

Χαρακτηριστικά	Πλήθος(%)
Ηλικία, μέσος όρος $\pm$ διακύμανση	61.56 $\pm$ 18.93
Φύλο (γυναίκες)	164(55.8)
Score ApacheII, μέσος όρος $\pm$ διακύμανση	15.82 $\pm$ 6.05
Διάρκεια παραμονής, μέσος όρος $\pm$ διακύμανση	20.42 $\pm$ 17.76
Διάρκεια προηγούμενης νοσηλείας πάνω από 2 ημέρες	100 (34)
Διάρκεια μηχανικής αναπνοής πριν τη διάγνωση της VAP για τους ασθενείς που την απέκτησαν, μέσος όρος $\pm$ διακύμανση	12.25 $\pm$ 9.05
Διάρκεια μηχανικής αναπνοής των ασθενών χωρίς VAP, μέσος όρος $\pm$ διακύμανση	12.74 $\pm$ 11.54
Τύπος ασθενή	
medical	192(65.3)
surgical	102(34.7)
Χαρακτηριστικά ασθενών κατά την είσοδο στη ΜΕΘ	
Διαβήτης	44(14.97)
Κακοήθεια	43(14.63)
Κώμα	58(19.72)
Σοκ	66(22.45)
Λοίμωξη	124(42.18)
Κύρια διάγνωση εισαγωγής	
Ασθένεια πνευμόνων	162(55.1)
Νευρολογική ασθένεια	29(9.86)
Ενδοκοιλιακή ασθένεια	35(11.9)
Καρδιαγγειακή ασθένεια	30(10.2)
Πολλαπλό τραύμα	38(12.93)

Οι ασθενείς που απέκτησαν τουλάχιστον μία λοίμωξη κατά την παραμονή τους στη ΜΕΘ είναι 154(52.4%), από τους οποίους απεβίωσαν οι 72(46.8%) ασθενείς, ενώ αξιοσημείωτο είναι το ποσοστό των ασθενών που απεβίωσαν ενώ δεν απέκτησαν καμία λοίμωξη(45%). Η συχνότητα των νοσοκομειακών λοιμώξεων και η θνητότητα για

την κάθε κατηγορία ασθενών που απέκτησαν συγκεκριμένο είδος λοίμωξης φαίνεται στον παρακάτω Πίνακα 4.2. Από τον πίνακα αυτόν παρατηρούμε ότι είτε οι ασθενείς δεν απέκτησαν καμία λοίμωξη, είτε απέκτησαν μόνο μία λοίμωξη, είτε περισσότερες, έχουν περίπου το ίδιο ποσοστό θανάτου (γύρω στο 46%). Αυτό αρχικά μας προκαλεί έκπληξη αφού το αναμενόμενο είναι ότι όσο πιο πολλές λοιμώξεις πάθει ένας ασθενής τόσο πιο επιρρεπής στο θάνατο θα είναι. Άρα ίσως να υπάρχουν άλλοι παράγοντες που επιδρούν στον κίνδυνο θανάτου. Αυτοί οι παράγοντες ίσως είναι η σοβαρότητα της κατάστασης του ασθενούς, το αν είχε αποκτήσει κάποια λοίμωξη σε προηγούμενη ΜΕΘ ή νοσοκομείο και άλλοι και θα τους μελετήσουμε αργότερα.

Πίνακας 4.2: Ποσοστά νοσοκομειακών λοιμώξεων και η θνητότητα τους

Κατηγορία ασθενών	Πλήθος(%)	Απεβίωσαν(%)
Ασθενείς που δεν απέκτησαν καμία λοίμωξη	140(47.6)	63(45)
Ασθενείς που απέκτησαν μόνο μία λοίμωξη	85(28.9)	40(47)
Ασθενείς που απέκτησαν περισσότερες από μία λοιμώξεις	69(23.5)	32(46.4)
Είδος λοίμωξης		
VAP και άλλη μία λοίμωξη	115(39.1)	53(46.1)
Μόνο VAP	50(17)	22(44)
BSI& UTI	112(38)	54(48)

Από τον συγκριτικό πίνακα (Πίνακας 4.3) κατηγοριών ασθενών με την ύπαρξη ή μη της λοίμωξης VAP και του θανάτου διακρίνουμε όπως είναι αναμενόμενο ότι οι γηραιότεροι ασθενείς έχουν σχεδόν διπλάσια ποσοστά θανάτου ως προς τους μικρότερους ασθενείς ανεξαρτήτως όμως αν προσβλήθηκαν από την VAP ή όχι. Οι ασθενείς, οι οποίοι κατά την εισαγωγή τους στη ΜΕΘ έχουν ήδη λοίμωξη είναι πιο επιρρεπείς στον θάνατο. Ο δείκτης βαρύτητας περιστατικού μας δείχνει ότι όντως όσο πιο μεγάλος είναι τόσο πιο μεγάλα είναι και τα ποσοστά θανάτου. Επιπλέον, παρατηρούμε ότι ο θάνατος στη ΜΕΘ είναι πιο πιθανόν να συμβεί στις αρχές της παραμονής των ασθενών στη ΜΕΘ και η median διάρκεια παραμονής των ασθενών που δεν απέκτησαν VAP είναι κατά πολύ μικρότερη(13.54 έναντι 31.14 ημερών) από τη διάρκεια παραμονής των ασθενών που απέκτησαν VAP (είτε φεύγουν από τη ΜΕΘ ζωντανοί είτε όχι).

Πίνακας 4.3: Συγκριτικά στοιχεία κατηγοριών ασθενών ως προς την ύπαρξη της VAP και του θανάτου

	Χωρίς VAP N=179(%)	Με VAP N=115(%)	Θάνατος χωρίς VAP N=82(45.8%)	Θάνατος με VAP N=53(46.1%)
Ηλικία>59	114(63.7)	74(64.3)	63(55.3)	41(55.4)
Ηλικία≤59	65(36.1)	41(35.7)	19(29.2)	12(29.3)
Ηλικία(median)	64	68	69	72
Φύλο:άντρες	106(59.2)	58(50.4)	51(48.1)	29(50)
γυναίκες	73(40.8)	57(49.6)	31(42.5)	24(42.1)
Λοίμωξη κατά την είσοδο	81(45.3)	43(37.4)	47(58)	23(53.5)
Χωρίς λοίμωξη κατά την είσοδο	98(54.7)	72(62.6)	35(35.7)	30(41.7)
Apache II ≤ 9	25(14)	10(8.7)	2(8)	2(20)
9<Apache II < 23	133(74.3)	90(78.3)	67(50.4)	40(44.4)
Apache II ≥ 23	21(11.7)	15(13)	13(61.9)	11(73.3)
Apache II(median)	15	16	17.50	18
Τύπος:medical	118(65.9)	74(64.3)	62(52.5)	37(50)
surgical	61(34.1)	41(35.7)	20(32.8)	16(39)
Παραμονή στη ΜΕΘ ≤21	160(89.4)	46(40)	74(46.3)	26(56.5)
21<Παραμ. στη ΜΕΘ≤31	9(5)	29(25.2)	4(44.4)	12(41.4)
Παραμονή στη ΜΕΘ >32	10(5.6)	40(34.8)	4(40)	15(37.5)
Παραμονή στη ΜΕΘ (median)	10	27	9	24

Η μέση διάρκεια παραμονής των ασθενών στη ΜΕΘ εκτιμήθηκε ως 20.42 ημέρες, ενώ ο διάμεσος της διάρκειας της παραμονής στη ΜΕΘ εκτιμήθηκε στις 15 ημέρες.

Η median διάρκεια παραμονής στη ΜΕΘ των ασθενών που βγήκαν από τη ΜΕΘ ζωντανοί είναι 16 ημέρες, ενώ για τους ασθενείς που απεβίωσαν στη ΜΕΘ είναι 13 ημέρες.

Οι ασθενείς που πέθαναν στη ΜΕΘ χωρίς να αποκτήσουν τη VAP παρέμειναν στη ΜΕΘ για 9 ημέρες (median time), ενώ οι 22 ασθενείς που πέθαναν στη ΜΕΘ έχοντας αποκτήσει μόνο τη VAP παρέμειναν στη ΜΕΘ για 17 ημέρες (median time), και για 8.5 ημέρες μετά την απόκτηση της VAP .

Όσον αφορά τους ασθενείς που αποχώρησαν από τη ΜΕΘ ζωντανοί, ο χρόνος παραμονής για αυτούς που απέκτησαν μόνο τη VAP (28 ασθενείς) είναι 24 ημέρες (median time), ενώ για αυτούς που δεν την απέκτησαν είναι 10 ημέρες (median time).

Οι ασθενείς που απέκτησαν μόνο τη VAP παρέμειναν στη ΜΕΘ για 20 ημέρες

(median time) είτε βγήκαν από αυτή ζωντανοί ή όχι. Ενώ οι ασθενείς που δεν απέκτησαν τη VAP παρέμειναν για 10 ημέρες(median time). Ο χρόνος παραμονής στη ΜΕΘ μετά την απόκτηση της λοίμωξης για τους ασθενείς που απέκτησαν μόνο τη VAP ήταν 10 ημέρες(median time). Και η διάμεσος του χρόνου μέχρι την εμφάνιση της VAP για τους ασθενείς που απέκτησαν μόνο VAP είναι 8 ημέρες (διατεταρτημοριακό εύρος: 6-11 ημέρες).

4.3 Επιλογή κατάλληλων στατιστικών μεθόδων

Μια συχνή προσέγγιση για την προσαρμογή των συνεπιδρώντων παραγόντων ήταν η παλινδρόμηση (regression) ή αλλιώς απλοϊκή προσέγγιση (naive approach) κατά τους G.Schulgen, M.Schumacher[28], που βασίζεται στα λογιστικά μοντέλα και πρωτοχρησιμοποιήθηκε στα τέλη της δεκαετίας του 1980 ώστε να εκτιμηθεί η υπερβολική θνησιμότητα που οφείλεται στη VAP. Η προσέγγιση αυτή παρέχει μια πρόβλεψη για το τι θα συνέβαινε σε ασθενείς με VAP αν δεν είχαν αποκτήσει τη λοίμωξη, χρησιμοποιώντας πληροφορία από ένα σύνολο ασθενών που δεν είχαν τη λοίμωξη και που είναι αρκετά διαφορετικοί, για παράδειγμα λιγότερο σοβαρά ως περιστατικά, από τους ασθενείς με VAP. Αυτή η προσέγγιση υποθέτει ακόμη ότι η διάρκεια παραμονής στη ΜΕΘ των ασθενών με VAP, εάν δεν είχαν προσβληθεί από τη VAP θα ήταν η ίδια με τη διάρκεια των ασθενών που δεν είχαν προσβληθεί. Λόγω των παραπάνω, η προσέγγιση αυτή δίνει μικρή σημασία στην ετερογένεια και στο χρονοδιάγραμμα, οδηγώντας σε μεροληπτικά αποτελέσματα.

Το ταίριασμα (matching) χρησιμοποιείται κυρίως αναδρομικά για να εκτιμηθεί η υπερβολική θνητότητα σε ΜΕΘ που οφείλεται στη VAP, ξεχωρίζοντας την επίδραση της από τις επιδράσεις της ηλικίας, του φύλου, της σοβαρότητας των υποκείμενων ασθενειών ή οποιουδήποτε άλλου παράγοντα που επηρεάζει την πρόγνωση για τους ασθενείς που υποστηρίζονται με μηχανική αναπνοή. Από την πληροφορία που διαθέτουμε για κάθε ασθενή κατά την έξοδο του από τη ΜΕΘ, γίνεται προσπάθεια οι ασθενείς που προσβλήθηκαν από τη VAP να ταιριάζουν με ένα ή περισσότερους ασθενείς που δεν ανέπτυξαν VAP και οι οποίοι είναι όσο πιο όμοιοι γίνεται ως προς τους ασθενείς που προσβλήθηκαν από τη VAP σε συγκεκριμένους παράγοντες κινδύνου για την διάρκεια νοσηλείας. Ένας περιορισμός για αυτή την διαδικασία είναι ότι μόνο ένα υποσύνολο ασθενών μπορεί να ταιριάζει με τους κατάλληλους ασθενείς ελέγχου,

ειδικά όταν υπάρχουν πολλά κριτήρια ταιριάσματος. Δηλαδή, αυτή η προσέγγιση αποτελεί μια ατομική σύγκριση της διάρκειας παραμονής στη ΜΕΘ των ασθενών που απέκτησαν τη λοίμωξη κάποια στιγμή κατά τη διάρκεια της παραμονής τους (cases) με τη διάρκεια των ασθενών που βγήκαν από τη ΜΕΘ χωρίς να αποκτήσουν κάποια λοίμωξη (controls). Οι δύο παραπάνω διαδικασίες μπορεί να θεωρηθούν συμπληρωματικές και να χρησιμοποιηθούν βέλτιστα σε συνδυασμό. Όμως δεν περιγράφουν το τι συμβαίνει στην πραγματικότητα και επιπλέον επειδή καμία από αυτές δεν λαμβάνει υπ'όψιν της το πραγματικό χρονικό διάγραμμα των γεγονότων, αγνοώντας την φυσική χρονική εξάρτηση των γεγονότων οδηγούμαστε σε χρονικά εξαρτημένη μεροληψία άρα και μεροληπτικές εκτιμήσεις με υπερεκτιμημένα αποτελέσματα. Σε αντίθεση με την κλασική case-control μελέτη ένας κατάλληλος ασθενής control πρέπει να πληροί τον όρο ότι δεν θα γίνει αργότερα ασθενής case.

Ο καλύτερος εκπρόσωπος των μοντέλων επιβίωσης για την εκτίμηση της υπερβολικής θνητότητας που αναλογεί στη VAP, είναι το χρονικά εξαρτημένο μοντέλο αναλογικού κινδύνου του Cox (time-dependent proportional hazards Cox model) γι αυτό και χρησιμοποιείται ευρέως.

Οι περισσότερες από τις παλιές προσεγγίσεις παραβλέπουν το γεγονός ότι υπάρχουν επιπλέον ανταγωνιστικοί κίνδυνοι. Η λογοκρισία θεωρείται ως ένας ανταγωνιστικός κίνδυνος, δεδομένου ότι η έξοδος από τη ΜΕΘ επιδρά στην πιθανότητα απόκτησης του ενδιαφερόμενου γεγονότος (π.χ. θάνατος πριν την έξοδο). Η έξοδος είναι, επίσης, ένας ανταγωνιστικός κίνδυνος που είναι χαρακτηριστικός για πολλές περιοχές της κλινικής έρευνας και χρειάζεται να μοντελοποιηθεί ξεχωριστά από τη λογοκρισία προκειμένου να αποφευχθούν λανθασμένα συμπεράσματα. Οι περισσότερες έρευνες χρησιμοποιούν τα μοντέλα ανταγωνιστικών κινδύνων, αφού εξασφαλίζουν πιο ακριβείς εκτιμήσεις συγκρινόμενα με τα λογιστικά μοντέλα και τις απλές καμπύλες επιβίωσης.

Τα μοντέλα πολλαπλών καταστάσεων μπορούν να θεωρηθούν ως μια γενίκευση του βασικού πλαισίου για τον χειρισμό δεδομένων επιβίωσης σε περίπτωση που διάφορα γεγονότα (πιθανώς και ανταγωνιστικά) συμβαίνουν επιτυχώς με την πάροδο του χρόνου, και γι αυτό επιτρέπουν τον χειρισμό και των ανταγωνιστικών κινδύνων αλλά και της λογοκρισίας. Τα μοντέλα πολλαπλών καταστάσεων αποδείχθηκαν ένα χρήσιμο και εύκολα κατανοητό πλαίσιο για τον ορισμό και την εκτίμηση της θνητότητας που αποδίδεται σε νοσοκομειακές λοιμώξεις συμπεριλαμβανομένου και της

VAP. Προφανώς αυτή η προσέγγιση πετυχαίνει διαφορετικές εκτιμήσεις συγκρινόμενη με τις παλιότερες προσεγγίσεις. Επίσης, τα μοντέλα πολλαπλών καταστάσεων συνεχώς εξελίσσονται υπολογιστικά για να υποστηρίξουν την επίλυση ολοένα και περισσότερων αλλά και διαφορετικών σχετικών ζητημάτων.

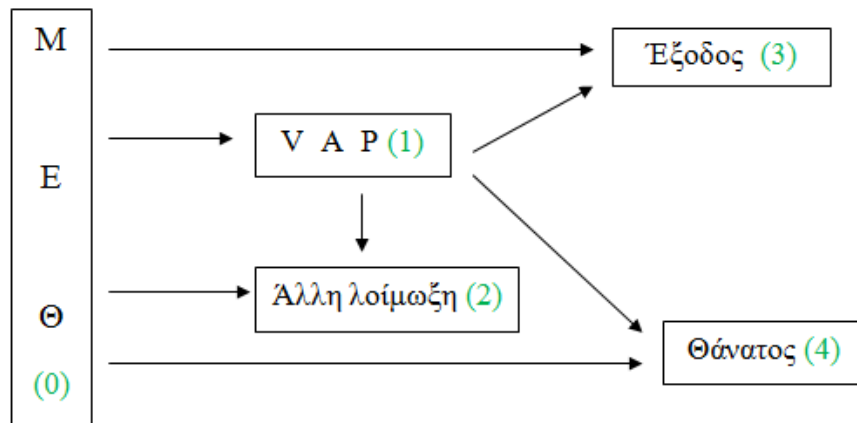
4.4 Στατιστική ανάλυση

Λόγω μειονεκτημάτων των παλιότερων μεθόδων που είδαμε στην προηγούμενη παράγραφο αποφασίσαμε να γίνει η ανάλυση των δεδομένων μας για την εκτίμηση της θνητότητας που αναλογεί στη VAP και της διάρκειας παραμονής των ασθενών με VAP στη ΜΕΘ με τα μοντέλα πολλαπλών καταστάσεων. Σε ένα μοντέλο πολλαπλών καταστάσεων για την ύπαρξη και την επίδραση των νοσοκομειακών λοιμώξεων, θεωρούμε για παράδειγμα την είσοδο στη ΜΕΘ, την νοσοκομειακή λοίμωξη, τον θάνατο και την έξοδο από τη ΜΕΘ ως πιθανές καταστάσεις ασθένειας. Η απόκτηση της VAP από έναν ασθενή κατά τη διάρκεια της παραμονής του στη ΜΕΘ θεωρείται ως μια μετάβαση στην κατάσταση “VAP” από την αρχική υγιή κατάσταση. Η εμφάνιση των λοιμώξεων, ο θάνατος και η έξοδος από τη ΜΕΘ είναι τα γεγονότα στην πορεία του ασθενούς.

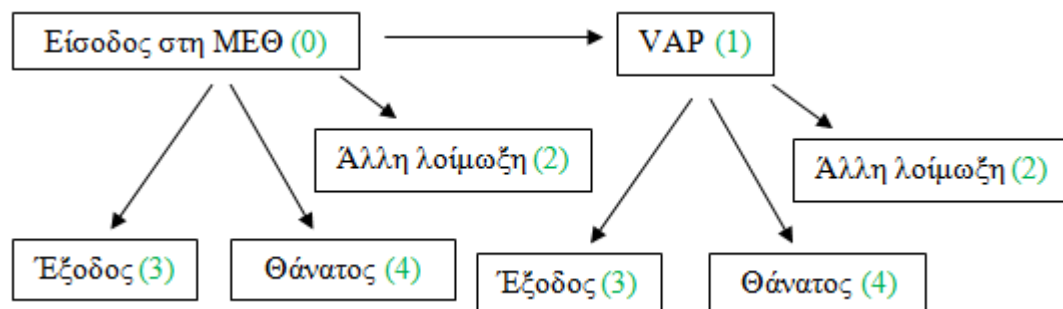
Τα μοντέλα πολλαπλών καταστάσεων υποθέτουν ότι οι μετακινήσεις ενός ασθενή μεταξύ των καταστάσεων της εισόδου στη ΜΕΘ, της νοσοκομειακής λοίμωξης, του θανάτου και της εξόδου από τη ΜΕΘ είναι Μαρκοβιανές. Δηλαδή, η μελλοντική κατάσταση μιας εισόδου στη ΜΕΘ εξαρτάται μόνο από την παρούσα κατάσταση. Το Σχήμα 4.2 δείχνει εμφανώς τις βασικές πιθανές μετακινήσεις ενός ασθενή με τη μορφή ενός μοντέλου ανικανότητας. Ενώ το Σχήμα 4.3 είναι η μετατροπή του παραπάνω μοντέλου σε προοδευτικό μοντέλο πολλαπλών καταστάσεων, δηλαδή το μοντέλο όπου κάθε κατάσταση να έχει μία μόνο μετάβαση προς αυτήν.

Οι ασθενείς και οι χρόνοι των γεγονότων τους υποτίθενται ανεξάρτητοι μεταξύ τους. Η παρατήρηση των γεγονότων γίνεται σε συνεχή χρόνο και η βασική χρονική κλίμακα είναι ο χρόνος (δηλαδή οι ημέρες) από την είσοδο των ασθενών στη ΜΕΘ.

Όπως έχουμε ήδη αναφέρει, εμείς ενδιαφερόμαστε κυρίως για την εμφάνιση της VAP που είναι το πιο συχνά εμφανίσιμο είδος νοσοκομειακής λοίμωξης. Τις BSI και UTI τις θεωρήσαμε ως δευτερεύουσες λοιμώξεις και έτσι τις εξετάσαμε ως μία λοίμωξη με το όνομα “άλλη λοίμωξη”. Στα διαθέσιμα δεδομένα μας και οι δύο κα-



Σχήμα 4.2: Το μοντέλο ανικανότητας (πολλαπλών καταστάσεων) που ακολουθήθηκε για την τελική μας μελέτη. Τα τετράγωνα δηλώνουν τις πιθανές καταστάσεις και τα βέλη τις πιθανές μεταβάσεις για έναν ασθενή κατά την διάρκεια της παραμονής του στη ΜΕΘ. Από τις πέντε καταστάσεις, οι δύο (0 και 1) είναι μεταβατικές καταστάσεις, δηλαδή οι καταστάσεις αυτές ακολουθούνται από μετάβαση σε άλλη κατάσταση. Οι άλλες τρεις καταστάσεις (2, 3 και 4) είναι απορροφητικές καταστάσεις, δηλαδή καταστάσεις από τις οποίες δεν μπορεί να υπάρξει περαιτέρω μετάβαση.



Σχήμα 4.3: Το προοδευτικό μοντέλο ανικανότητας (πολλαπλών καταστάσεων) που ακολουθήθηκε για τη μελέτη μας. Κι εδώ οι καταστάσεις 0 και 1 είναι μεταβατικές, ενώ οι υπόλοιπες απορροφητικές. Εδώ όμως έχουμε 8 καταστάσεις.

ταστάσεις ("VAP" και "άλλη λοίμωξη") είναι μη επαναλαμβανόμενες και πολλαπλές λοιμώξεις του ίδιου τύπου δεν μοντελοποιήθηκαν.

Όλοι οι ασθενείς ξεκινούν από την κοινή αρχική, μεταβατική, μη επαναλαμβανόμενη κατάσταση που δηλώνει την είσοδο τους στη ΜΕΘ την χρονική στιγμή $t = 0$. Η εμφάνιση της VAP μοντελοποιείται ως η μετάβαση στην ενδιάμεση κατάσταση και

έχουν ως ανταγωνιστικά τελικά σημεία τον θάνατο ή την έξοδο από τη ΜΕΘ.

Αρχικά είχαμε θεωρήσει διάφορα άλλα μοντέλα. Σε όλα τα μοντέλα που θα αναφερθούν παρακάτω υπήρχαν οι μεταβάσεις από οποιαδήποτε κατάσταση προς τα καταληκτικά γεγονότα θανάτου και εξόδου από τη ΜΕΘ εκτός κι αν αναφερθεί κάτι διαφορετικό. Θεωρήσαμε μοντέλα στα οποία εξετάζαμε τη VAP όταν εμφανιζόταν πρώτη ή μόνη της και λογοκρίναμε τους ασθενείς που είχαν άλλη λοίμωξη ως πρώτη ή μόνη της (αυτοί οι ασθενείς δημιουργούσαν μια ξεχωριστή κατάσταση λογοκριμένων ασθενών). Ή μοντέλα στα οποία εξετάζαμε τη VAP οπότε εμφανιζόταν χωρίς να υπάρχει λογοκρισία (οι ασθενείς που αποκτούσαν διαφορετική λοίμωξη θεωρούνταν ασθενείς χωρίς λοίμωξη). Επίσης είχαμε εξετάσει μοντέλα στα οποία έπαιζε ρόλο η σειρά με την οποία εμφανιζόντουσαν οι λοιμώξεις με σκοπό να εξετάσουμε τις επιδράσεις των άλλων λοιμώξεων όταν εμφανίζονται και πότε στη VAP. Δηλαδή ένα μοντέλο όπου οι ασθενείς αποκτούσαν πρώτα κάποια άλλη λοίμωξη, δευτερευόντως τη VAP και από εκεί κατέληγαν στην έξοδο από τη ΜΕΘ ή στο θάνατο ή άλλο μοντέλο όπου οι ασθενείς αποκτούσαν πρώτα τη VAP, μεταγενέστερα και άλλη λοίμωξη και από εκεί πήγαιναν στις απορροφητικές καταστάσεις θανάτου ή εξόδου από τη ΜΕΘ. Στα δύο παραπάνω μοντέλα οι ασθενείς που αποκτούσαν μόνο μία λοίμωξη λογοκρίνονταν. Επιπλέον θεωρήσαμε ένα μοντέλο στο οποίο οι ασθενείς μετέβαιναν στην πρώτη χρονικά αποκτηθείσα λοίμωξη, καθώς και από την κατάσταση της άλλης λοίμωξης στην κατάσταση της VAP και από εκεί στα καταληκτικά γεγονότα, με εναλλακτικό μοντέλο εκείνο όπου κατέληγαν οι ασθενείς στα καταληκτικά γεγονότα και από την κατάσταση της VAP αλλά και από την κατάσταση της άλλης λοίμωξης.

Μετά από αρκετή συζήτηση με την κ.Αποστολοπούλου καταλήξαμε στην παραπάνω μορφή του μοντέλου που είδαμε στα Σχήματα 4.2 και 4.3 για τον λόγο ότι μας ενδιαφέρουν αποκλειστικά όσοι ασθενείς παθαίνουν VAP και τίποτα άλλο συγκριτικά με τους ασθενείς που δεν απέκτησαν τη VAP κι έτσι με το που αποκτά κάποιος ασθενής άλλη λοίμωξη εκτός της VAP, λογοκρίνεται μεταβαίνοντας σε μια ξεχωριστή κατάσταση. Θεωρήσαμε την κατάσταση “άλλη λοίμωξη” ως απορροφητική μιας που δεν μας ενδιέφερε να εξετάσουμε ποια ήταν η έκβαση των ασθενών που παθαίνουν (και) άλλη λοίμωξη εκτός από τη VAP.

Προηγούμενες βιβλιογραφικές αναφορές χειρίστηκαν διαφορετικά τους ασθενείς με άλλη λοίμωξη. Σε μία περίπτωση οι ασθενείς με άλλες λοιμώξεις αναλύθηκαν ως μέρος των ασθενών που δεν απέκτησαν καμία λοίμωξη μιας που ο πληθυσμός των

ασθενών με άλλες λοιμώξεις αποτελούσε μικρό ποσοστό του συνολικού πληθυσμού που δεν είχαν αποκτήσει καμία λοίμωξη και έτσι είχε μικρή επίδραση στις εκτιμήσεις. Στη δεύτερη περίπτωση θεώρησαν τους ασθενείς λογοκριμένους σύμφωνα με το ίδιο σκεπτικό με το δικό μας αλλά δεν εξέταζαν την περίπτωση όπου κάποιος ασθενής αποκτούσε πρώτα άλλη λοίμωξη και μετά την λοίμωξη τους ενδιαφέροντος.

Ως προς το προγραμματιστικό κομμάτι της ανάλυσης μας, για τον υπολογισμό της θνητότητας που συνδέεται με τη VAP χρησιμοποιήσαμε το πακέτο της R (R Foundation for Statistical Computing, Vienna, Austria) `msm` (multi-state model written by C.Jackson) το οποίο προσαρμόζει ένα συνεχούς χρόνου Μαρκοβιανό μοντέλο πολλαπλών καταστάσεων μέσω της μέγιστης πιθανοφάνειας και μπορεί να προσαρμόσει συμμεταβλητές στις πιθανότητες μετάβασης της Μαρκοβιανής αλυσίδας. Για τον υπολογισμό (σε ημέρες) της επιπλέον διάρκειας παραμονής στη ΜΕΘ χρησιμοποιήσαμε το πακέτο της R `etm` (empirical transition matrix- written by A.Allignol) σε συνδυασμό με το πακέτο της R `changeLOS` (change in length of hospital stay- written by M.Wrangler, J.Beyersmann, M.Schumacher) όπου υπολογίζουν την αλλαγή στη διάρκεια νοσηλείας των ασθενών μέσω του εκτιμητή Aalen-Johansen για τον πίνακα πιθανοτήτων μετάβασης στα μοντέλα πολλαπλών καταστάσεων και βασίζονται στην προσέγγιση B των G.Schulgen, M.Schumacher[28], οι οποίοι θεώρησαν τη διαφορά των αναμενόμενων διαδοχικών παραμονών δεδομένης της κατάστασης της λοίμωξης την στιγμή s . Επίσης για τον υπολογισμό των αθροιστικών κινδύνων και των πιθανοτήτων μετάβασης για τα μοντέλα πολλαπλών καταστάσεων χρησιμοποιήσαμε το πακέτο της R `mvna` (multivariate Nelson-Aalen estimator, written by A.Allignol) .

Τη διάρκεια παραμονής των ασθενών στο νοσοκομείο τη μετρήσαμε με χρονικό σημείο αφετηρίας την ημέρα εισόδου ως ημέρα 1. Αυτός ο τρόπος υπολογισμού είναι ελάχιστα διαφορετικός από την τυπική ανάλυση επιβίωσης όπου θεωρούν την επόμενη ημέρα της εισόδου ως ημέρα 1. Στο δικό μας πλαίσιο, όμως, η διάρκεια παραμονής συχνά χρησιμοποιείται ως οικονομική ποσότητα και επειδή όλες οι ημέρες, ακόμα και η ημέρα της εισόδου, συνδέονται με το κόστος, μετρήσαμε την ημέρα εισόδου ως ημέρα 1. Αυτό δεν επηρεάζει τα αποτελέσματα για την αλλαγή στη διάρκεια παραμονής των ασθενών λόγω λοίμωξης, παρά μόνο τα αποτελέσματα των καμπυλών Kaplan-Meier και των αθροιστικών συναρτήσεων συχνότητας, δηλαδή ο άξονας του χρόνου είναι μετακινημένος κατά μία ημέρα δεξιά.

Επίσης θέσαμε ένα μοντέλο αναλογικών κινδύνων για τα ανταγωνιστικά τελι-

κά σημεία του θανάτου και της εξόδου από τη ΜΕΘ ώστε να εκτιμηθεί εάν η εμφάνιση της VAP επιδρά στην διάρκεια παραμονής των ασθενών στη VAP. Αυτή η μοντελοποίηση λαμβάνει υπ'όψιν της την μη απευθείας επίδραση της παραπάνω διάρκειας παραμονής στη ΜΕΘ στη θνητότητα. Είναι σημαντικό να τονίσουμε ότι χειριστήκαμε την VAP ως χρονικά εξαρτημένη συμμεταβλητή (time-dependent covariate). Δηλαδή την χρονική στιγμή (ημέρα) που ο ασθενής απέκτησε τη VAP άλλαξε η κατάσταση του από μη εκτεθειμένος σε αυτή τη λοίμωξη σε εκτεθειμένος. Από και πέρα οι ασθενείς αυτοί θεωρούνται εκτεθειμένοι μέχρι και την έξοδο τους από τη ΜΕΘ. Αυτή η προσέγγιση μπορεί να χρησιμοποιηθεί ώστε να αποφανθούμε εάν η VAP επιδρά κάπως στο συνδυασμένο κίνδυνο θανάτου ή εξόδου (π.χ. στο συνολικό χρόνο νοσηλείας στο νοσοκομείο). Αν θεωρούσαμε τη VAP ως αναφορική συμμεταβλητή (baseline covariate) τότε θα οδηγούμασταν σε υπερεκτίμηση της επίδρασης στην αναλογία κινδύνου (hazard ratio).

Στο μοντέλο μας για την στατιστική ανάλυση συμπεριλήφθηκαν αναφορικές συμμεταβλητές όπως η ηλικία, το φύλο, ο τύπος του ασθενή, το σκορ Apache II, το αν υπήρχε ήδη λοίμωξη κατά την εισαγωγή στη ΜΕΘ, η διάρκεια μηχανικής αναπνοής κ.ά. Από αυτές, μόνο το σκορ Apache2 (p-value 4.404008e-08<0.05), το αν υπήρχε ήδη λοίμωξη κατά την εισαγωγή στη ΜΕΘ (p-value 0.01623048<0.05) και η διάρκεια μηχανικής αναπνοής (p-value 0<0.05) ήταν στατιστικά σημαντικές όταν η κάθε μία ήταν η μοναδική συμμεταβλητή στο μοντέλο.

Για να βρεθεί ένα κατάλληλο μοντέλο, αρχικά θεωρήσαμε το μοντέλο που περιελάμβανε μόνο μία μεταβλητή από τις τρεις κάθε φορά, έπειτα θεωρήσαμε εισάγαμε και δεύτερη μεταβλητή και ελέγξαμε με likelihood ratio test την σημαντικότητα και βρήκαμε ότι το μοντέλο με τις 2 μεταβλητές είναι σημαντικότερο (-2logLR: 45.75257 με 7 β.ε. όσες και οι παραπάνω μεταβάσεις και p-value 9.766561e-08) και τέλος εισάγαμε και την τρίτη μεταβλητή και πάλι μετά από έλεγχο likelihood ratio test καταλήξαμε στο μοντέλο που περιέχει και τις τρεις μεταβλητές (-2logLR: 83.91807 με 7 β.ε. όσες και οι παραπάνω μεταβάσεις και p-value 2.220446e-15).

Έτσι οι συναρτήσεις πιθανοτήτων μετάβασης μοντελοποιήθηκαν (από την Cox μορφή των μοντέλων πολλαπλών καταστάσεων της (3.20)) ως

$$\lambda_{ik}(t) = \lambda_{0k}(t) \exp\{\beta_1 \cdot \text{Apache2}_i + \beta_2 \cdot \text{infection}_i + \beta_3 \cdot \text{durven}_i\} \quad (4.1)$$

όπου $i = 1, 2, \dots, 294$ οι ασθενείς, k οι καταστάσεις 1,2,3,4,5 που αντιστοιχούν στις “ΜΕΘ”, “VAP”, “άλλη λοίμωξη”, “έξοδος από τη ΜΕΘ” και “θάνατος αντίστοιχα” και λ_{0k} ο αναφορικός κίνδυνος.

4.5 Αποτελέσματα

Από τους 294 ασθενείς, οι 164(55.8%) ήταν γυναίκες. Οι 140(47.6%) δεν απέκτησαν κάποια από τις τρεις λοιμώξεις όσο βρίσκονταν στη ΜΕΘ και από αυτούς οι 63(45%) απεβίωσαν. Η πρώτη (ή και μοναδική) λοίμωξη που απέκτησαν 94(32%) ασθενείς ήταν η VAP, για τους υπόλοιπους 60 ασθενείς ήταν κάποια άλλη λοίμωξη. Για τους 50 από τους 94 ασθενείς που απέκτησαν πρώτα VAP, η VAP ήταν και η μοναδική λοίμωξη που απέκτησαν. Και οι 22(44%) από αυτούς απεβίωσαν.

Η έξοδος είναι, στατιστικά, ένας ανταγωνιστικός κίνδυνος ως προς τον θάνατο στη ΜΕΘ, αφού ο ασθενής τελικά θα βγει από τη ΜΕΘ και το μόνο πράγμα που δεν γνωρίζουμε είναι ο λόγος της εξόδου του (βγήκε από τη ΜΕΘ λόγω θανάτου ή βγήκε ζωντανός;). Έτσι, οι αθροιστικές συναρτήσεις συχνότητας για τον θάνατο και για την έξοδο αθροίζουν τελικά στη μονάδα, επειδή κάθε ασθενής είτε αποχωρεί από τη ΜΕΘ είτε πεθαίνει σε αυτήν. Συνδυάζοντας τα δύο αυτά γεγονότα στοχεύουμε στο να εξετάσουμε την επίδραση της VAP στη διάρκεια παραμονής στη ΜΕΘ. Σε ένα τέτοιο μοντέλο όπου η έξοδος από τη ΜΕΘ (είτε ζωντανός είτε νεκρός) θεωρείται ως ένα συνδυασμένο γεγονός, η επίδραση της VAP σε αυτήν ως μια χρονικά εξαρτημένη συμμεταβλητή σε ένα μοντέλο αναλογικού κινδύνου είναι αρκετά σημαντική (εκτίμηση: -0.839, p-value < 0.001, Wald test). Άρα, η VAP όντως επιδρά σημαντικά στη διάρκεια παραμονής στη ΜΕΘ.

Στο μοντέλο για την εύρεση της αναλογούμενης θνητότητας της VAP όπου δεν συμπεριλαμβάνονται συμμεταβλητές παρατηρούμε ότι οι ασθενείς που έχουν προσβληθεί από τη VAP είναι δύο φορές πιο πιθανό να αναπτύξουν κι άλλες λοιμώξεις εκτός της VAP σε σύγκριση με τους ασθενείς που δεν απέκτησαν VAP, αλλά και σχεδόν δύο φορές πιο πιθανό να αναπτύξουν κι άλλες λοιμώξεις από το να βγουν από τη ΜΕΘ (είτε ζωντανοί είτε όχι). Επίσης, οι ασθενείς είναι πιθανότερο να αναπτύξουν VAP από το να αναπτύξουν BSI ή UTI ή να βγουν από τη ΜΕΘ (ζωντανοί ή νεκροί).

Σε ένα πολυμεταβλητό μοντέλο όπου συμπεριλαμβάνονται η λοίμωξη κατά την εισαγωγή, το σκορ Apache II και η διάρκεια μηχανικής αναπνοής ($\lambda_{ik}(t) = \lambda_{0k}(t) \exp\{\beta_1 \cdot$

$Apache2_i + \beta_2 \cdot infection_i + \beta_3 \cdot durven_i$) βλέπουμε ότι οι ασθενείς είναι λίγο πιο πιθανό να βγουν από τη ΜΕΘ ζωντανοί παρά νεκροί είτε έχουν αποκτήσει VAP είτε όχι αλλά και όπως παραπάνω οι ασθενείς που έχουν προσβληθεί από τη VAP είναι δύο φορές πιο πιθανό να αναπτύξουν κι άλλες λοιμώξεις εκτός της VAP σε σύγκριση με τους ασθενείς που δεν απέκτησαν VAP, καθώς και πιο πιθανό να αναπτύξουν κι άλλες λοιμώξεις από το να βγουν από τη ΜΕΘ(είτε ζωντανοί είτε όχι). Επίσης, οι ασθενείς είναι πιθανότερο να αναπτύξουν VAP από το να αναπτύξουν BSI ή UTI ή να βγουν από τη ΜΕΘ. (οι συμμεταβλητές βρίσκονται στις μέσες τιμές τους).

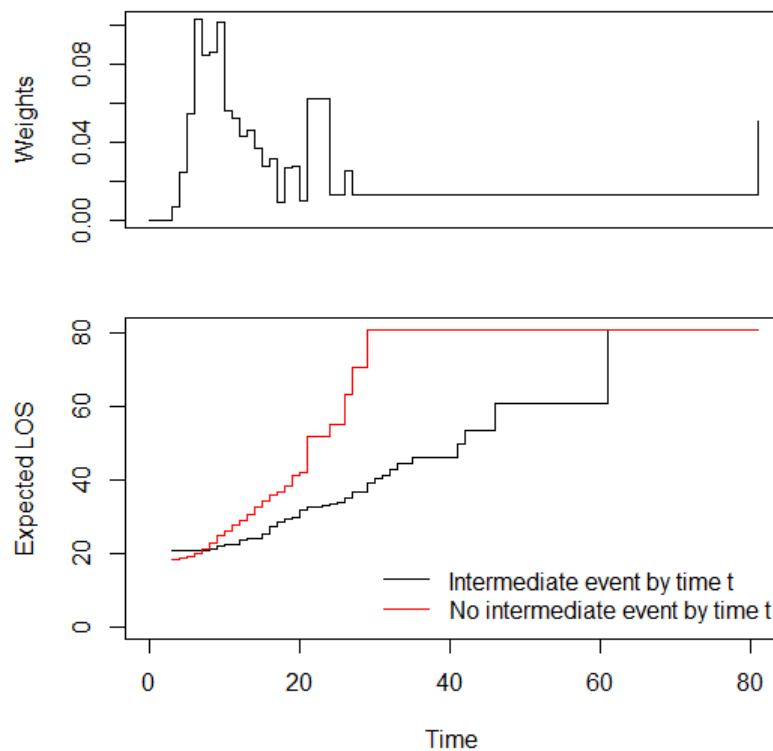
Και οι τρεις συμμεταβλητές είναι στατιστικά σημαντικές. Πιο συγκεκριμένα, μεγάλο σκορ Apache II δείχνει να κάνει πιο πιθανή την αποβίωση των ασθενών σε σχέση με την έξοδο τους όντας ζωντανοί, παρόμοια αποτελέσματα φαίνεται να έχει και η διάρκεια της μηχανικής αναπνοής των ασθενών. Η παρουσία υπάρχουσας λοίμωξης κατά την εισαγωγή για τους ασθενείς που δεν απέκτησαν VAP εντείνει τον κίνδυνο θανάτου, ενώ για αυτούς που δεν την απέκτησαν μεγαλύτερη πιθανότητα έχει η έξοδος από τη ΜΕΘ.

Όλοι οι ασθενείς θεωρήθηκαν ως αντικείμενα ελέγχου για όσο χρονικό διάστημα δεν είχαν αποκτήσει τη VAP ή (για τους ασθενείς που απέκτησαν πρώτα ή μόνο κάποια άλλη λοίμωξη) για όσο χρόνο διάνυσαν μέχρι να λογοκριθούν.

Η αναμενόμενη αλλαγή στη διάρκεια παραμονής λόγω της VAP εκτιμήθηκε (μέσω του πακέτου changeLOS) στις 5.9 ± 1.8 ημέρες, όπως φαίνεται και από το Σχήμα 4.4.

Η εκτίμηση έγινε χρησιμοποιώντας βάρη υπολογισμένα από την κατανομή του χρόνου έως τα μέλη των κατηγοριών των ασθενών να γίνουν συγκεκριμένα (κι όχι όσοι έχουν τη λοίμωξη με όσους δεν έχουν) και δείχουν την συνεισφορά των γεγονότων που συμβαίνουν τον αντίστοιχο χρόνο στη συνολική εκτίμηση. Διαφορετική επιλογή κατανομής χρόνου για τον υπολογισμό των βαρών δεν επιφέρει ριζοσπαστικές αλλαγές στην εκτιμώμενη αλλαγή της διάρκειας παραμονής (-5.395 και -6.119 οι εκτιμήσεις της αλλαγής της διάρκειας παραμονής υπολογισμένες με διαφορετικά βάρη). Οι ασθενείς που αποχώρησαν από τη ΜΕΘ ζωντανοί συνέβαλαν στη διαφορά της διάρκειας παραμονής κατά -3.12 ημέρες, ενώ οι ασθενείς που απεβίωσαν κατά -2.82 ημέρες. Η απόκτηση μόνο της BSI άλλαξε τη διάρκεια παραμονής στη ΜΕΘ με μέση αλλαγή τις 3.105 ημέρες.

Στην ανάλυση για την εκτίμηση της παραπάνω διάρκειας παραμονής των ασθενών στη ΜΕΘ εξαιτίας της VAP **δεν** συμπεριλαμβάνονται συμμεταβλητές αφού η μέθοδος



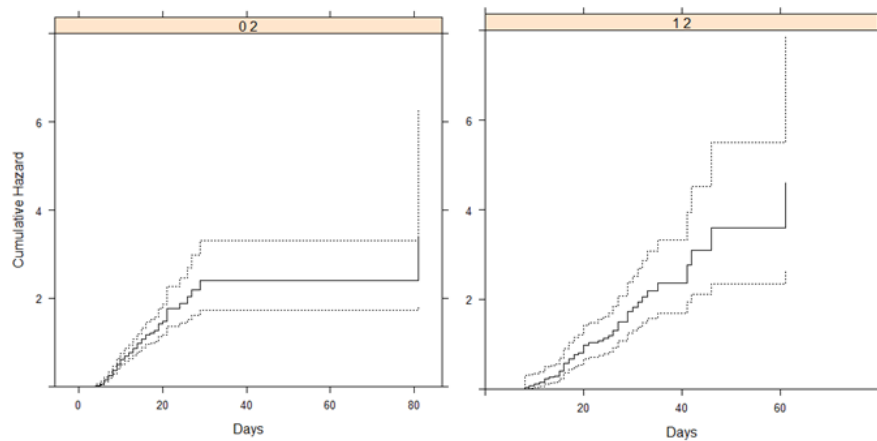
Σχήμα 4.4: Εκτιμώμενη αλλαγή στη διάρκεια παραμονής των ασθενών στη ΜΕΘ λόγω της VAP. Η διαφορά στις δύο γραμμές δεν εξαρτάται από τα βάρη.

μας δεν επιτρέπει προς το παρόν άμεση προσαρμογή σύμφωνα με την σοβαρότητα της ασθένειας και άλλους παράγοντες κινδύνου.

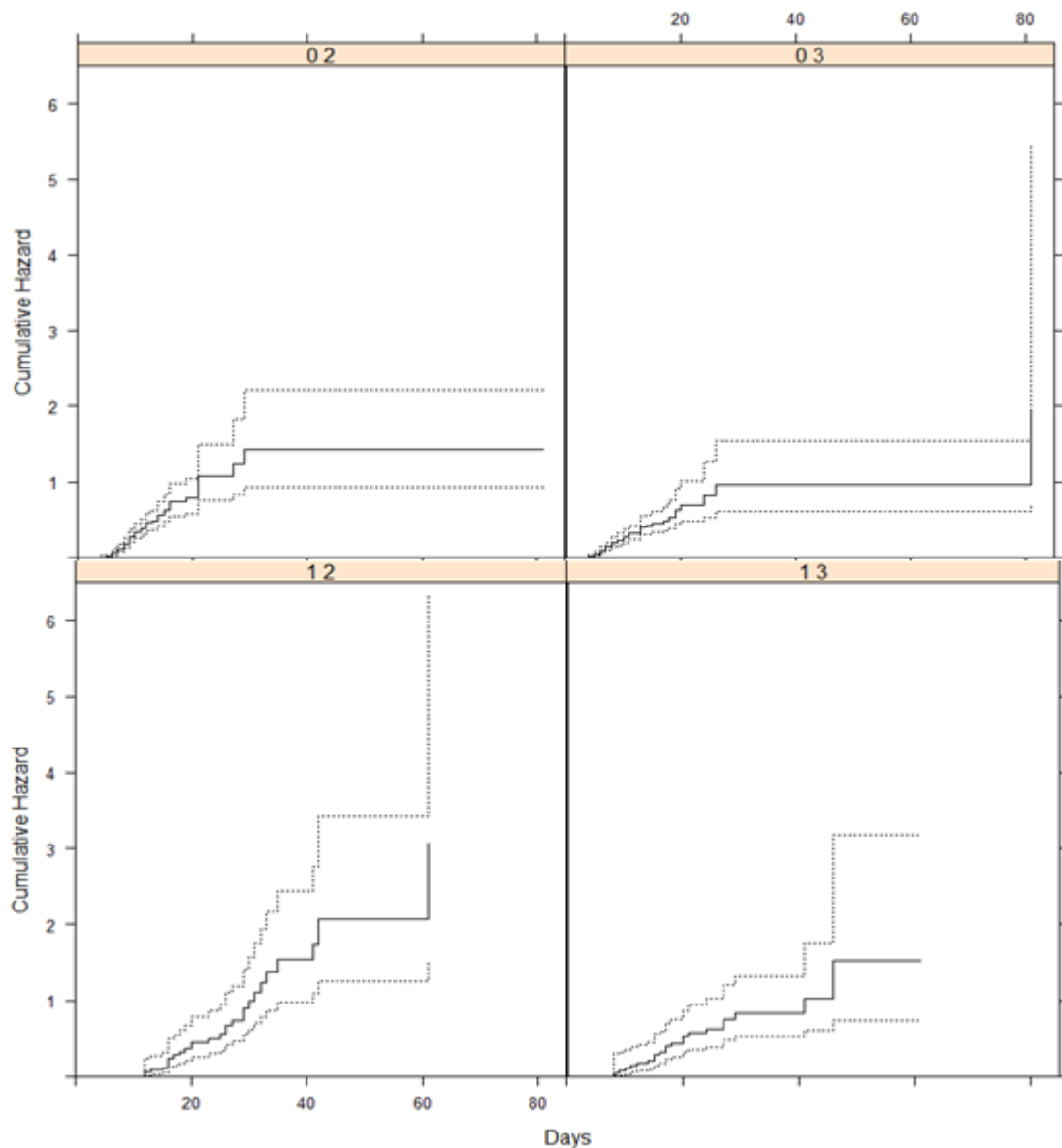
Στοχεύοντας στο να εξετάσουμε την επίδραση της VAP στο θάνατο των ασθενών στη ΜΕΘ προκύπτουν τα Σχήματα 4.5 και 4.6. Παρατηρούμε ότι η VAP φαίνεται να επιδρά αλλά όχι σε μεγάλο βαθμό στον κίνδυνο θανάτου. Αυτό δεν σημαίνει και ότι η VAP δεν επηρεάζει την θνητότητα, αλλά ότι:

- Η VAP εμφανίζεται να μειώνει τον κίνδυνο όλων των υπόλοιπων αιτιών στο τέλος της παραμονής στη ΜΕΘ
- Οι ασθενείς με VAP παραμένουν λιγότερο στη ΜΕΘ. Κατά τη διάρκεια αυτής της παραπάνω παραμονής τους οι ασθενείς που δεν απέκτησαν VAP, είναι εκτεθειμένοι σε έναν ουσιαστικά σταθερό κίνδυνο.

Επίσης από τα Σχήματα 4.5 και 4.6 παρατηρούμε ότι όσοι ασθενείς απέκτησαν VAP έχουν αυξημένη διάρκεια παραμονής στη ΜΕΘ αν λάβουμε ως διάστημα παρατήρησης τις πρώτες 38 ημέρες, επειδή ο αθροιστικός κίνδυνος για το τέλος της παραμονής της κατάστασης 1(όσοι ασθενείς είχαν VAP) είναι λιγότερος από αυτόν της αρχικής κατάστασης.

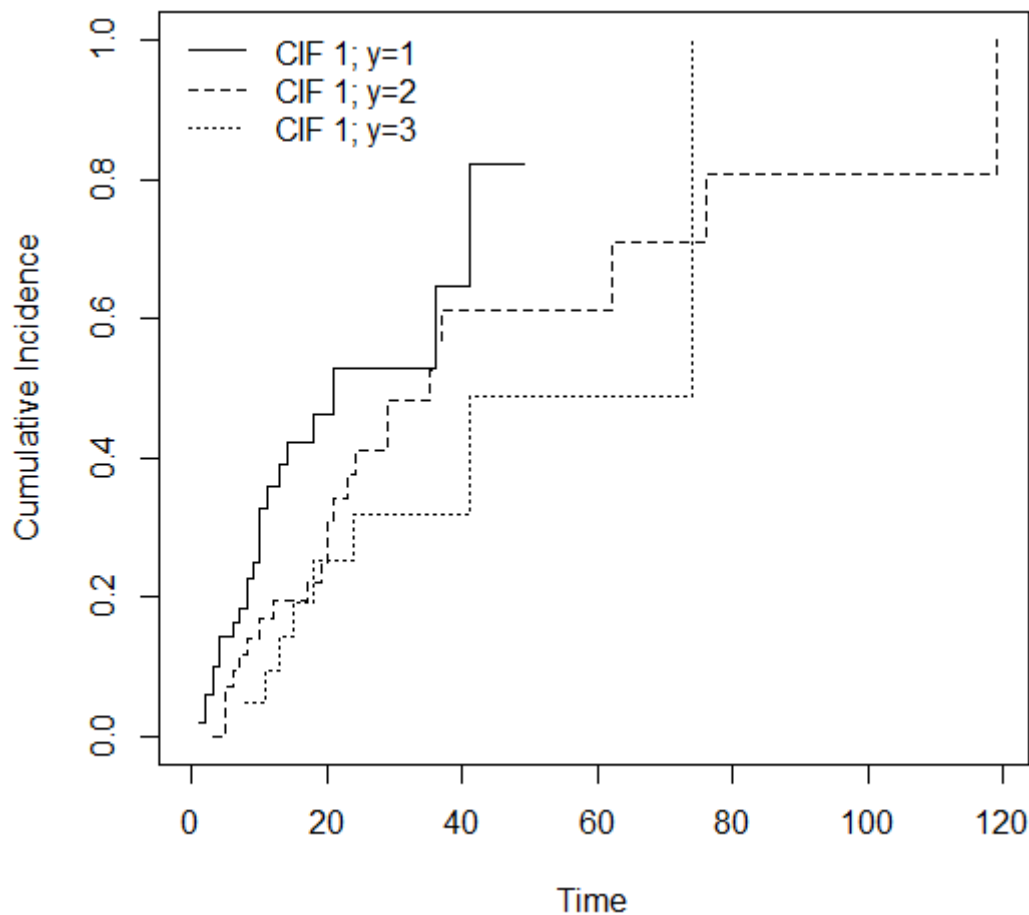


Σχήμα 4.5: Έστω να αποτελούν ένα γεγονός ο θάνατος στη ΜΕΘ και η έξοδος από τη ΜΕΘ, τότε αυτό το σχήμα συγκρίνει τους αθροιστικούς κινδύνους (παρουσιάζοντας και τα διαστήματα εμπιστοσύνης τους) των ασθενών με VAP (κάτω) και αυτών χωρίς VAP (πάνω) ως προς αυτό το γεγονός.

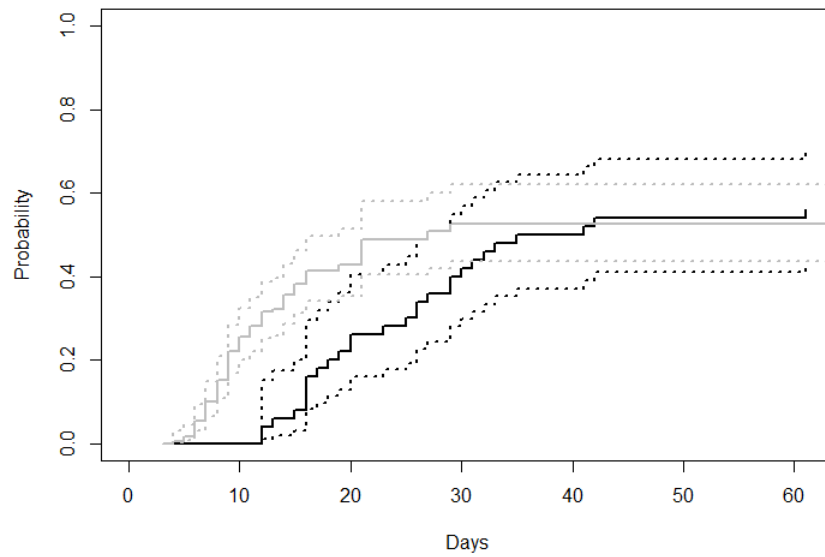


Σχήμα 4.6: Δεξιά στήλη: εκτιμητές Nelson-Aalen για τον αθροιστικό κίνδυνο θανάτου. Αριστερή στήλη: εκτιμητές Nelson-Aalen για τον αθροιστικό κίνδυνο εξόδου από τη ΜΕΘ. Η πάνω γραμμή αναφέρεται στους ασθενείς που δεν απέκτησαν VAP, ενώ η κάτω σε αυτούς που την απέκτησαν. Με τα 95% διαστήματα εμπιστοσύνης να αναπαρίστανται από τις διακεκομμένες γραμμές.

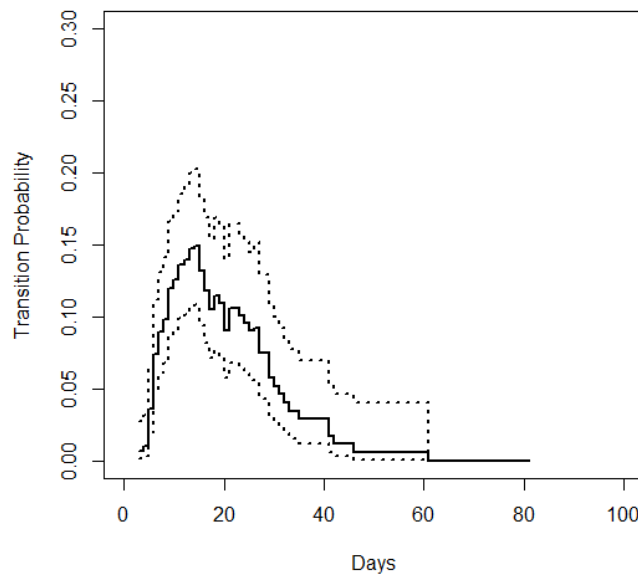
Από το παρακάτω Σχήμα 4.7 παρατηρούμε η αθροιστική επίπτωση της πρώτης κατηγορίας ασθενών είναι μεγαλύτερη από των υπόλοιπων δύο κατηγοριών ασθενών. Άρα συμπεραίνουμε ότι οι ασθενείς που απέκτησαν μόνο τη λοίμωξη VAP έχουν μικρότερη διάρκεια νοσηλείας στη ΜΕΘ συγκριτικά με τους ασθενείς που απέκτησαν πρώτα τη VAP και μετά κάποια άλλη λοίμωξη και αυτοί με τη σειρά τους μικρότερη διάρκεια νοσηλείας σε σύγκριση με τους ασθενείς που απέκτησαν άλλη λοίμωξη και ίσως να απέκτησαν τη VAP αργότερα.



Σχήμα 4.7: Η αθροιστική επίπτωση των ασθενών που απέκτησαν ($y = 1$:) αποκλειστικά VAP, ($y = 2$:) πρώτα VAP και μετά άλλη λοίμωξη και ($y = 3$:) αυτών που απέκτησαν άλλη λοίμωξη και μετά ίσως VAP.



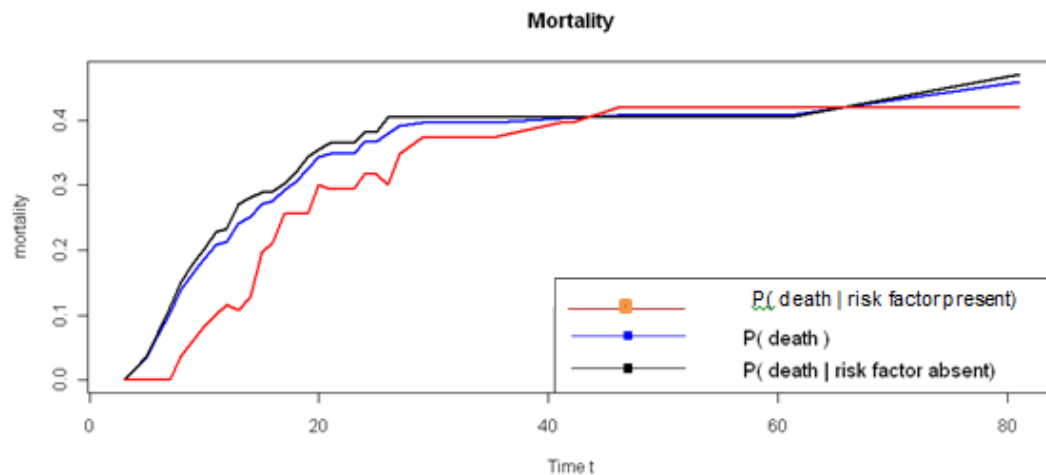
Σχήμα 4.8: Οι συμπαγείς γραμμές είναι οι εκτιμητές Aalen-Johansen για τον θάνατο. Οι μαύρες αφορούν τους ασθενείς με VAP, ενώ οι γκρι του ασθενείς που δεν απέκτησαν VAP. Οι διακεκομμένες γραμμές δηλώνουν τα 95% διαστήματα εμπιστοσύνης.



Σχήμα 4.9: Εκτιμητής Aalen-Johansen για την πιθανότητα της μετάβασης $0 \rightarrow 1$ με 95% διαστήμα εμπιστοσύνης βασισμένο στον συμπληρωματικό log-log μετασχηματισμό (ο οποίος προτιμάται να εφαρμόζεται για πραγματικές μεταβάσεις).

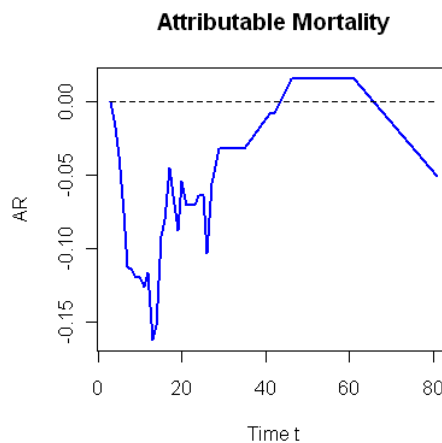
Το Σχήμα 4.9 μας παρουσιάζει την αναλογία των ασθενών με VAP κατά τη διάρκεια της μελέτης και παρατηρούμε ότι η αναλογία των ασθενών που αποκτούν VAP αυξάνεται τις πρώτες 20 μέρες και μετά μειώνεται πάλι.

Το Σχήμα 4.10 απεικονίζει την αθροιστική συχνότητα του θανάτου στη ΜΕΘ σε ολόκληρο τον πληθυσμό που μελετάμε ως συνάρτηση του χρόνου και την πιθανότητα θανάτου στη ΜΕΘ, δεδομένου ότι συνέβηκε η VAP ή όχι πριν από το t . Και οι τρεις συναρτήσεις δείχνουν να σταθεροποιούνται μετά τις 30 ημέρες.



Σχήμα 4.10: Θνητότητα

Στο Σχήμα 4.11 απεικονίζεται η αποδιδόμενη θνητότητα στη VAP και πάλι υπάρχει μια σταθεροποίηση στις 45 ημέρες που οδηγεί στη γενική εκτίμηση 5,5% αποδιδόμενης θνητότητας αλλά αργότερα μετά τις 60 ημέρες υπάρχει πάλι καθοδική πορεία.



Σχήμα 4.11: Αποδιδόμενη θνητότητα

Με μια πρώτη ματιά, παρατηρεί κανείς το "παράδοξο" των αρνητικών τιμών. Για τις αρχικές χρονικές στιγμές αυτό ίσως να εξηγείται από το γεγονός ότι οι ασθενείς που εισάγονται στη ΜΕΘ σε αρκετά άσχημη κατάσταση υγείας πιθανόν να πεθαίνουν χωρίς και απλά να μην έχουν τον χρόνο να αποκτήσουν τη VAP ($0 \rightarrow 1$ μετάβαση). Για τις τελευταίες χρονικές στιγμές μια πιθανή εξήγηση είναι ότι πια επιδρούν και άλλες λοιμώξεις στη θνητότητα των ασθενών.

4.6 Συζήτηση

Διάφορες πρόσφατες έρευνες υποστήριξαν ότι οι νοσοκομειακές λοιμώξεις αναδύονται ως ένα σημαντικό πρόβλημα σε αρκετές αναπτυσσόμενες χώρες με την νοσοκομειακή πνευμονία λόγω μηχανικής αναπνοής μέσω αναπνευστήρα (VAP) να είναι από τις κυριότερες και πιο θανατηφόρες. Γι αυτό εξετάσαμε την επίδραση της VAP στην θνητότητα των ασθενών στην Ελλάδα καθώς και στη διάρκεια νοσηλείας τους.

Η εφαρμογή των μοντέλων πολλαπλών καταστάσεων κατέληξε σε μια χρήσιμη και εύκολα κατανοητή προσέγγιση για την εκτίμηση της αποδιδόμενης θνητότητας και σχετικών ποσοτήτων λαμβάνοντας υπ' όψιν ανταγωνιστικούς κινδύνους και πιθανή λογοκρισία. Η συγκεκριμένη επιλογή μας, το λεγόμενο προοδευτικό μοντέλο ανικανότητας προσφέρει ερμηνευτικά πλεονεκτήματα από τη στιγμή που η πληροφορία για ένα άτομο που βρίσκεται σε μια κατάσταση την χρονική στιγμή t περιέχει όλη την πληροφορία των προηγούμενων μεταβάσεων με μοναδικό τρόπο.

Από την ανάλυση μας, τα αποτελέσματα που βρήκαμε δείχνουν ότι οι παρατηρούμενοι ασθενείς μας είχαν περισσότερες πιθανότητες να αποκτήσουν VAP σε σχέση με οποιοδήποτε άλλο γεγονός. Και μετά την απόκτηση της VAP ήταν επιρρεπείς στο να αποκτήσουν κι άλλες λοιμώξεις και να επιβαρυνθεί περισσότερο η κατάσταση τους. Όπως ήταν αναμενόμενο οι ασθενείς με μεγαλύτερο σκορ Apache II οδηγούνται με μεγαλύτερη πιθανότητα στο θάνατο. Η αρκετά μεγάλη μερίδα ασθενών που εισήχθησαν στη ΜΕΘ έχοντας ήδη λοίμωξη έχει μεγαλύτερη πιθανότητα να πεθάνει παρά να αποκτήσει VAP.

Κάποιες μελέτες έχουν δείξει ότι η αυξημένη αναλογούμενη θνητότητα στη VAP κυμαίνεται από 33% έως και 72% ενώ άλλες απέτυχαν να βρουν κάποια διαφορά στην θνητότητα. Μετά την παραπάνω ανάλυση μας, η αναλογούμενη θνητότητα στη VAP αποδίδεται σε ένα μικρό ποσοστό 5.5%.

Ο βαθμός της αποδιδόμενης θνητότητας στη VAP εξαρτάται από πολλούς παράγοντες όπως υποκείμενες ασθένειες, χαρακτηριστικά της λοίμωξης, η αντίδραση του ξενιστή-ασθενή ή και τυχόν δυσλειτουργίες οργάνων. Η ανακρίβεια των διαγνωστικών ελέγχων παίζει κι αυτή ένα σημαντικό ρόλο. Κλασσικές στατιστικές αναλύσεις της αποδιδόμενης θνητότητας στη VAP αγνοούσαν τα προαναφερθέντα ζητήματα των χρονικά εξαρτημένων μεταβλητών και ως εκ τούτου τα εμπειρικά αποτελέσματα είναι αρκετά αμφιλεγόμενα, με αρκετές μελέτες να αναφέρουν εκτιμήσεις για την επίδραση της VAP στη θνητότητα από μηδενικές έως εξαιρετικά επιβλαβείς.

Επίσης, εφαρμόσαμε κι ένα μοντέλο αναλογικού κινδύνου σε συνδυασμό με τα μοντέλα πολλαπλών καταστάσεων ώστε να εκτιμηθεί αν η εμφάνιση της VAP επιδρά στην διάρκεια παραμονής των ασθενών στη ΜΕΘ.

Αρκετές μελέτες μπορούν να βρεθούν στη βιβλιογραφία αναφορικά με την διάρκεια νοσηλείας των ασθενών στη ΜΕΘ λόγω νοσοκομειακών λοιμώξεων. Κάποιοι συγγραφείς εξέτασαν τις συνέπειες όλων των τύπων των νοσοκομειακών λοιμώξεων, ενώ άλλοι συγκεντρώσαν την προσοχή τους μόνο στη VAP ή στη BSI. Μελέτες έχουν επανηλειμένα αποδείξει ότι όταν μια λοίμωξη συμβεί, συμπεριλαμβανομένου και της VAP, τότε παρατείνεται και η διάρκεια της νοσηλείας στο νοσοκομείο από 4.3 έως 13 ημέρες. Αυτό το ευρύ φάσμα μπορεί μερικώς να εξηγηθεί από τις παρεκλίσεις των ασθενών αλλά και τις μεθοδολογικές διαφορές. Επιπλέον, λόγω του ότι ο αριθμός των ασθενών που θεωρήθηκαν για κάθε έρευνα κυμαινόταν από 46 ασθενείς έως 120.000, αναμένουμε και τα αποτελέσματα να διαφέρουν εξαιτίας αυτής της τεράστιας διαφοράς στο μέγεθος της κοορτής.

Υπάρχουν μελέτες που εξετάζουν την παράταση της διάρκειας παραμονής των ασθενών στη ΜΕΘ που να οφείλεται μόνο στη BSI λοίμωξη και τα αποτελέσματα τους κυμαίνονται από 8 έως 20 ημέρες. Σε αντίθεση με αυτά αλλά σύμφωνα με τους J.Beyersmann et al. [7], το αποτέλεσμα της δικής μας μελέτης ήταν αύξηση κατά 3.105 ημέρες κατά μέσο όρο. Η επίδραση των νοσοκομειακών λοιμώξεων μπορεί διαφέρει και εξαρτάται από το είδος της λοίμωξης.

Η διάρκεια παραμονής στη ΜΕΘ μπορεί να παραταθεί για τους ασθενείς που απέκτησαν λοίμωξη και τελικά απέρχονται από τη ΜΕΘ, αλλά η διάρκεια παραμονής μπορεί και να συντομευθεί για τους ασθενείς που απεβίωσαν στη ΜΕΘ. Αυτό ακριβώς παρουσιάζουν και τα αποτελέσματα μας που δείχνουν τη διάρκεια νοσηλείας των ασθενών που απεβίωσαν να είναι μικρότερη από των ασθενών που αποχώρησαν από

τη ΜΕΘ ζωντανοί ανεξάρτητα από το αν απέκτησαν VAP ή όχι. Αφού οι ασθενείς που δεν απέκτησαν καμία λοίμωξη στη ΜΕΘ απεβίωσαν μετά από 8 ημέρες ενώ αυτοί που βγήκαν ζωντανοί παρέμειναν στη ΜΕΘ 9 ημέρες. Το ίδιο συνέβηκε και για τους ασθενείς που απέκτησαν VAP, απεβίωσαν μετά από 17 ημέρες ενώ αποχώρησαν ζωντανοί μετά από 24 ημέρες.

Υπάρχει μεγάλη οικονομική επίδραση των νοσοκομειακών λοιμώξεων στα συστήματα υγείας των χωρών με συνέπεια πιθανή μεγάλη αποταμίευση αν παρθούν τα κατάλληλα μέτρα για την πρόληψη των λοιμώξεων. Όμως, τα δεδομένα για την επιμήκυνση της παραμονής των ασθενών στη ΜΕΘ δεν είναι σημαντικά μόνο για να εκτιμηθεί το οικονομικό βάρος των λοιμώξεων στο εθνικό σύστημα υγείας αλλά είναι σημαντικά και για τα νοσοκομεία ώστε να χρησιμοποιήσουν αυτή την πληροφορία για την καλύτερη πρόληψη.

Παράρτημα Α΄

Ονομασία και τιμές μεταβλητών

Αριθμός μεταβλητής	Ονομασία μεταβλητών (Όλες οι μεταβλητές που ήταν διαθέσιμες για ανάλυση)
1	Αύξοντα αριθμό
2	Όνομα ΜΕΘ
3	Ηλικία ασθενούς (συνεχής μεταβλητή)
4	Φύλο
5	Συνολικός Χρόνος παραμονής ή συνολική διάρκεια νοσηλείας στη ΜΕΘ (συνεχής)
6	Εισαγωγή στη ΜΕΘ από τα επείγοντα ΝΑΙ/ΟΧΙ
7	Εισαγωγή στη ΜΕΘ από το Χειρουργείο ΝΑΙ/ΟΧΙ
8	Εισαγωγή στη ΜΕΘ από άλλη ΜΕΘ ΝΑΙ/ΟΧΙ
9	Εισαγωγή στη ΜΕΘ από παθολογική κλινική ΝΑΙ/ΟΧΙ
10	Εισαγωγή στη ΜΕΘ από άλλο Νοσοκομείο ΝΑΙ/ΟΧΙ
11	Διάρκεια προηγούμενης νοσηλείας πάνω από 2 ημέρες ΝΑΙ/ΟΧΙ
12	Διάρκεια προηγούμενης νοσηλείας πάνω από 5 ημέρες
13	Τύπος περιστατικού ΝΑΙ/ΟΧΙ
Κύρια διάγνωση εισαγωγής	
15	Αναπνευστική νόσος ΝΑΙ/ΟΧΙ
16	Νευρολογική νόσος ΝΑΙ/ΟΧΙ
17	Χειρουργείο στη κοιλιακή χώρα ΝΑΙ/ΟΧΙ
18	Καρδιολογική νόσος ΝΑΙ/ΟΧΙ
19	Πολλαπλό τραύμα ΝΑΙ/ΟΧΙ
Χαρακτηριστικά ασθενών κατά την εισαγωγή	
21	Λοίμωξη κατά την εισαγωγή ΝΑΙ/ΟΧΙ
22	Διαβήτης κατά την εισαγωγή ΝΑΙ/ΟΧΙ
23	Κακοήθεια κατά την εισαγωγή ΝΑΙ/ΟΧΙ
24	Κώμα ΝΑΙ/ΟΧΙ
25	Δείκτης βαρύτητας περιστατικού
26	Δείκτης υποκείμενων νοσημάτων
27	Σοκ ΝΑΙ/ΟΧΙ

	Επεμβατικές τεχνικές που επηρεάζουν την πνευμονία από αναπνευστήρα-VAP, Τη βακτηριαιμία- BSI , και την ουρολοίμωξη από καθετήρα - UTI
29	Έγινε Επαναδιασωλήνωση-VAP NAI/OXI
30	Έγινε Βρογχοσκόπηση-VAP NAI/OXI
31	Τοποθετήθηκε Ρινογαστρικός σωλήνας- ημέρα εισαγωγής VAP NAI/OXI
32	Έγινε Τραχειοστομία –VAP NAI/OXI
33	Πότε έγινε η τραχειοστομία από την ημέρα εισαγωγής στη ΜΕΘ
34	Συνολική διάρκεια τραχειοστομίας
35	Διάρκεια τραχειοστομίας πριν τη διάγνωση της VAPκαι συνολική διάρκεια τραχειοστομίας των ασθενών που δεν ανέπτυξαν VAP
36	Διάρκεια τραχειοστομίας πριν τη διάγνωση της VAP
37	Διάρκεια τραχειοστομίας πριν τη διάγνωση της VAPκαι συνολική διάρκεια τραχειοστομίας των ασθενών που δεν ανέπτυξαν VAP, και συνολική διάρκεια νοσηλείας στη ΜΕΘ των υπόλοιπων ασθενών
38	Διάρκεια τραχειοστομίας μετά τη διάγνωση της VAP
39	Συνολική διάρκεια μηχανικής αναπνοής από την εισαγωγή στη ΜΕΘ
40	Διάρκεια μηχανικής αναπνοής πριν τη διάγνωση της VAP και συνολική διάρκεια μηχανικής αναπνοής των ασθενών χωρίς VAP
41	Διάρκεια μηχανικής αναπνοής πριν τη διάγνωση της VAP
42	Διάρκεια μηχανικής αναπνοής μετά τη διάγνωση της VAP
43	Περιλαμβάνει, τη Διάρκεια μηχανικής αναπνοής μετά τη διάγνωση της VAP για τους ασθενείς που ανέπτυξαν μόνο VAP, τη διάρκεια μηχανικής αναπνοής από τη διάγνωση της VAP μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την VAP, και τη διάρκεια μηχανικής αναπνοής από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της VAP, και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη
44	Χορηγήθηκε Παρεντερική διατροφή-VAPNAI/OXI
45	Διάρκεια μηχανικής αναπνοής πριν τη διάγνωση της BSI και συνολική διάρκεια μηχανικής αναπνοής των ασθενών χωρίς BSI
46	Διάρκεια μηχανικής αναπνοής πριν τη διάγνωση της BSI
47	Διάρκεια μηχανικής αναπνοής μετά τη διάγνωση της BSI
48	Κεντρικός φλεβικός καθετήρας- CVC ημέρα εισαγωγής NAI/OXI
49	Συνολική διάρκεια- CVC από την εισαγωγή στη ΜΕΘ
50	διάρκεια- CVC πριν τη BSI και συνολική διάρκεια CVC των ασθενών χωρίς BSI
51	Διάρκεια - CVC πριν τη BSI
52	Διάρκεια - CVC μετά τη BSI
53	Εισαγωγή Καθετήρα κύστεως ημέρα εισαγωγής NAI/OXI
54	Συνολική Διάρκεια καθετήρα κύστεως από την εισαγωγή στη ΜΕΘ

55	Διάρκεια καθετήρα κύστεως πριν τη UTI και συνολική διάρκεια καθετήρα κύστεως των ασθενών χωρίς UTI
56	Διάρκεια καθετήρα κύστεως πριν τη UTI
57	Διάρκεια καθετήρα κύστεως μετά τη UTI
58	Χορήγηση Παρεντερικής διατροφής- BSI NAI/OXI
59	Έγινε Διασωλήνωση Θώρακα NAI/OXI
60	Πότε έγινε η διασωλήνωσης θώρακα από την εισαγωγή στη ΜΕΘ
61	Συνολική διάρκεια διασωλήνωσης θώρακα
62	Διάρκεια διασωλήνωσης θώρακα πριν τη VAP, συνολική διάρκεια διασωλήνωσης θώρακα για τους ασθενείς με διασωλήνωση θώρακα και χωρίς VAP και συνολική διάρκεια νοσηλείας για τους υπόλοιπους ασθενείς
63	Διάρκεια διασωλήνωσης θώρακα μετά τη VAP
64	Διάρκεια διασωλήνωσης θώρακα πριν τη BSI, συνολική διάρκεια διασωλήνωσης θώρακα για τους ασθενείς με διασωλήνωση θώρακα και χωρίς BSI και συνολική διάρκεια νοσηλείας για τους υπόλοιπους ασθενείς
65	Διάρκεια διασωλήνωσης θώρακα μετά τη BSI
66	Παρουσία Αιμοδιάλυσης-VAP, NAI/OXI
67	Διάρκεια αιμοδιάλυσης πριν τη VAP και συνολική διάρκεια αιμοδιάλυσης των ασθενών χωρίς VAP
68	Διάρκεια αιμοδιάλυσης μετά τη VAP
69	Παρουσία Αιμοδιάλυσης- BSI NAI/OXI
70	Διάρκεια αιμοδιάλυσης πριν τη BSI και συνολική διάρκεια αιμοδιάλυσης των ασθενών χωρίς BSI
71	Διάρκεια αιμοδιάλυσης μετά τη BSI
Ειδικές θεραπείες	
73	Χορήγηση αίματος –VAP NAI/OXI
74	Αριθμός μονάδων αίματος VAP
75	Χορήγηση κορτιζόνης VAP NAI/OXI
76	Χορήγηση ινóτροπων φαρμάκων-VAP
77	Χορήγηση καταστολής-VAP
78	Αριθμός κατηγοριών αντιβιοτικών που χορηγήθηκαν πριν τη VAP
79	Χορήγηση γ γενεάς κεφαλοσπορίνη πριν τη VAP NAI/OXI
80	Χορήγηση κινολόνης πριν τη VAP NAI/OXI
81	Χορήγηση καρβαπενέμες πριν τη VAP NAI/OXI
82	Αριθμός κατηγοριών αντιβιοτικών που χορηγήθηκαν πριν τη BSI
83	Χορήγηση γ γενεάς κεφαλοσπορίνη πριν τη BSI NAI/OXI
84	Χορήγηση κινολόνης πριν τη BSI NAI/OXI
85	Χορήγηση καρβαπενέμες πριν τη BSI NAI/OXI
86	Χορήγηση πάνω από 4 κατηγορίες αντιβιοτικών NAI/OXI

Παράμετροι έκβασης και περιγραφικά στοιχεία	
88	Θάνατος στη ΜΕΘ. NAI/OXI
89	Ημέρα εμφάνισης θανάτου από την εισαγωγή στη ΜΕΘ
90	Ημέρα εμφάνισης θανάτου από την εισαγωγή στη ΜΕΘ για τους θανόντες και ημέρα εξόδου για τους επιβιώσαντες
91	Ασθενείς που εισήχθησαν χωρίς λοίμωξη και δεν απέκτησαν λοίμωξη στη ΜΕΘ NAI/OXI
92	Ασθενείς που απέκτησαν λοίμωξη στη ΜΕΘ NAI/OXI
93	Ασθενείς που απέκτησαν τουλάχιστο μία λοίμωξη NAI/OXI
94	Ασθενείς που απέκτησαν πάνω από μια λοίμωξη NAI/OXI
95	Διάρκεια νοσηλείας πριν τη πρώτη λοίμωξη
96	Διάρκεια νοσηλείας πριν τη πρώτη λοίμωξη για τους ασθενείς με λοίμωξη και συνολική διάρκεια νοσηλείας στους υπόλοιπους ασθενείς
97	Διάρκεια νοσηλείας μετά τη πρώτη λοίμωξη
98	Διάρκεια νοσηλείας μετά τη πρώτη λοίμωξη και συνολική διάρκεια νοσηλείας των υπόλοιπων ασθενών
99	Αριθμός λοιμώξεων
100	Είδος λοίμωξης
101	Διάρκεια νοσηλείας από την εισαγωγή μέχρι τη VAP
102	Διάρκεια νοσηλείας πριν τη VAP και συνολική διάρκεια νοσηλείας στους υπόλοιπους ασθενείς
103	Διάρκεια νοσηλείας μετά τη VAP
104	Διάρκεια νοσηλείας μετά τη VAP και συνολική διάρκεια νοσηλείας στους υπόλοιπους ασθενείς
105	Περιλαμβάνει, τη Διάρκεια νοσηλείας μετά τη διάγνωση της VAP για τους ασθενείς που ανέπτυξαν μόνο VAP, τη διάρκεια νοσηλείας από τη διάγνωση της VAP μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την VAP, και τη διάρκεια νοσηλείας από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της VAP, και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
106	Θάνατος μετά τη VAP NAI/OXI
107	Σε πόσες ημέρες πέθανε ο άρρωστος μετά τη VAP
108	Διάρκεια νοσηλείας από την εισαγωγή μέχρι τη BSI
109	Διάρκεια νοσηλείας πριν τη BSI και συνολική διάρκεια νοσηλείας στους υπόλοιπους ασθενείς
110	Διάρκεια νοσηλείας μετά τη BSI
111	Διάρκεια νοσηλείας μετά τη BSI και συνολική διάρκεια νοσηλείας στους υπόλοιπους ασθενείς

112	Περιλαμβάνει, τη διάρκεια νοσηλείας μετά τη διάγνωση της BSI για τους ασθενείς που ανέπτυξαν μόνο BSI , τη διάρκεια νοσηλείας από τη διάγνωση της BSI μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την BSI, και τη διάρκεια νοσηλείας από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της BSI, και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
113	Θάνατος μετά τη BSI NAI/OXI
114	Σε πόσες ημέρες πέθανε ο άρρωστος μετά τη BSI
115	Διάρκεια νοσηλείας με καθετήρα κύστεως μετά την UTI και συνολική διάρκεια νοσηλείας με καθετήρα κύστεως για τους υπόλοιπους ασθενείς που δεν ανέπτυξαν UTI
116	Περιλαμβάνει, τη διάρκεια νοσηλείας με καθετήρα κύστεως μετά τη διάγνωση της UTI για τους ασθενείς που ανέπτυξαν μόνο UTI, τη διάρκεια νοσηλείας με καθετήρα κύστεως από τη διάγνωση της UTI μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την UTI, και τη διάρκεια νοσηλείας με καθετήρα κύστεως από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της UTI, και τη συνολική διάρκεια νοσηλείας με καθετήρα κύστεως για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
117	Θάνατος μετά την UTI NAI/OXI
118	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την UTI
119	Ασθενείς που απέκτησαν λοίμωξη από <i>Acinetobacter Baumannii</i> παθογόνο NAI/OXI
120	Το είδος της λοίμωξης που αναπτύχθηκε από <i>Acinetobacter Baumannii</i> παθογόνο
121	Αριθμός λοιμώξεων από <i>Acinetobacter Baumannii</i> παθογόνο
122	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της λοίμωξης από <i>Acinetobacter Baumannii</i> παθογόνο
123	Διάρκεια νοσηλείας μετά την λοίμωξη από <i>Acinetobacter Baumannii</i> παθογόνο
124	Περιλαμβάνει, τη Διάρκεια νοσηλείας μετά τη διάγνωση της λοίμωξης από <i>Acinetobacter Baumannii</i> παθογόνο για τους ασθενείς που ανέπτυξαν μόνο λοίμωξη από <i>Acinetobacter Baumannii</i> παθογόνο , τη διάρκεια νοσηλείας από τη διάγνωση της λοίμωξης από <i>Acinetobacter Baumannii</i> παθογόνο μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την λοίμωξη από <i>Acinetobacter Baumannii</i> παθογόνο , και τη διάρκεια νοσηλείας από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της λοίμωξης από <i>Acinetobacter Baumannii</i> παθογόνο, και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
125	Θάνατος μετά την λοίμωξη από <i>Acinetobacter Baumannii</i> παθογόνο NAI/OXI

126	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την λοίμωξη από <i>Acinetobacter Baumannii</i> παθογόνο και τη συνολική διάρκεια νοσηλείας για τους υπόλοιπους ασθενείς
127	Ασθενείς που απέκτησαν λοίμωξη από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο NAI/OXI
128	Αριθμός λοιμώξεων από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο
129	Το είδος της λοίμωξης που αναπτύχθηκε από από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο
130	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της λοίμωξης από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο
131	Διάρκεια νοσηλείας μετά την λοίμωξη από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο
132	Περιλαμβάνει, τη Διάρκεια νοσηλείας μετά τη διάγνωση της λοίμωξης από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο για τους ασθενείς που ανέπτυξαν μόνο λοίμωξη από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο, τη διάρκεια νοσηλείας από τη διάγνωση της λοίμωξης από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την λοίμωξη από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο , και τη διάρκεια νοσηλείας από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της λοίμωξης από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο, και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
133	Θάνατος μετά την λοίμωξη από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο. NAI/OXI
134	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την λοίμωξη από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο και τη συνολική διάρκεια νοσηλείας για τους υπόλοιπους ασθενείς
135	Ανάπτυξη ουρολοίμωξης -UTI NAI/OXI
136	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της UTI
137	Το είδος του μικροβίου που προκάλεσε την UTI
138	Ανθεκτικά μικρόβια στις καρβαπενέμες που προκάλεσαν UTI
139	Ακατάλληλη εμπειρική αγωγή την ημέρα διάγνωσης της UTI
140	Ανάπτυξη Πνευμονίας από αναπνευστήρα- VAP
141	Είδος της VAP
142	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της VAP
143	Τα είδη των μικροβίων που προκάλεσαν την VAP
144	VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI
145	VAP από ανθεκτική στις καρβαπενέμες <i>Klebsiella</i> NAI/OXI
146	VAP από από ανθεκτική στις καρβαπενέμες <i>P aeruginosa</i>

147	Πολυμικροβιακή VAP NAI/OXI
148	Κωδικοποιημένα τα είδη μικροβίων που προκάλεσαν τη VAP
149	Καταγραφή της αντοχής σε ανθεκτικό ή ευαίσθητο στις καρβαπενέμες ανά παθογόνο που προκάλεσε τη VAP
150	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
151	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> και τη συνολική διάρκεια νοσηλείας των υπόλοιπων ασθενών
152	Διάρκεια νοσηλείας μετά την εμφάνιση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
153	Διάρκεια νοσηλείας μετά την εμφάνιση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> και τη συνολική διάρκεια νοσηλείας των υπόλοιπων ασθενών
154	Περιλαμβάνει, τη Διάρκεια νοσηλείας μετά τη διάγνωση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> για τους ασθενείς που ανέπτυξαν μόνο VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> , τη διάρκεια νοσηλείας από τη διάγνωση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> παθογόνο , και τη διάρκεια νοσηλείας από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> , και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
155	Διάρκεια μηχανικής αναπνοής μετά την εμφάνιση της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
156	Θάνατος μετά την VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI
157	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
158	VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI
159	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
160	Διάρκεια νοσηλείας μετά την εμφάνιση της VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
161	Διάρκεια μηχανικής αναπνοής μετά την εμφάνιση της VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
162	Θάνατος μετά την VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI

163	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
164	Ακατάλληλη εμπειρική αγωγή την ημέρα διάγνωσης της VAP από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI
165	Ακατάλληλη εμπειρική αγωγή την ημέρα διάγνωσης της VAP από ευαίσθητο στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI
166	Ακατάλληλη εμπειρική αγωγή την ημέρα διάγνωσης της VAP NAI/OXI
167	VAP από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα NAI/OXI
168	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της VAP από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
169	Διάρκεια νοσηλείας μετά την εμφάνιση της VAP από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
170	Θάνατος μετά την VAP από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα NAI/OXI
171	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την VAP από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
172	Ανάπτυξη βακτηριαιμίας -BSI NAI/OXI
173	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της BSI
174	Πρωτοπαθής -BSI NAI/OXI
175	Βακτηριαιμία από CVC NAI/OXI
176	Δευτεροπαθής -BSI NAI/OXI
177	Πηγή δευτεροπαθούς BSI
178	Τα είδη των μικροβίων που προκάλεσαν την BSI
179	Καταγραφή της αντοχής σε ανθεκτικό ή ευαίσθητο στις καρβαπενέμες ανά παθογόνο που προκάλεσε τη BSI
180	BSI από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i> NAI/OXI
181	BSI από ανθεκτική στις καρβαπενέμες <i>klebsiella</i>
182	BSI από ανθεκτική στις καρβαπενέμες Π αερυγινούσα
183	Πολυμικροβιακή BSI
184	Κωδικοποιημένα τα Είδη μικροβίων που προκάλεσαν τη BSI
185	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της BSI από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
186	Διάρκεια νοσηλείας μετά την εμφάνιση της BSI από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
187	Θάνατος μετά την BSI από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
188	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την BSI από ανθεκτικό στις καρβαπενέμες <i>Acinetobacter Baumannii</i>
189	BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα

190	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
191	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα και συνολική διάρκεια νοσηλείας των υπόλοιπων ασθενών
192	Διάρκεια νοσηλείας μετά την εμφάνιση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
193	Διάρκεια νοσηλείας μετά την εμφάνιση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα και τη συνολική διάρκεια νοσηλείας των υπόλοιπων ασθενών
194	Περιλαμβάνει, τη διάρκεια νοσηλείας μετά τη διάγνωση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα για τους ασθενείς που ανέπτυξαν μόνο BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα, τη διάρκεια νοσηλείας από τη διάγνωση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα μέχρι την εμφάνιση της επόμενης λοίμωξης για τους ασθενείς που ανέπτυξαν μη σχετιζόμενη λοίμωξη μετά την BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα, και τη διάρκεια νοσηλείας από την εισαγωγή μέχρι την ημέρα διάγνωσης της πρώτης λοίμωξης για τους ασθενείς που ανέπτυξαν κάποια λοίμωξη εκτός της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα, και τη συνολική διάρκεια νοσηλείας για τους ασθενείς που δεν ανέπτυξαν καμία λοίμωξη.
195	Διάρκεια μηχανικής αναπνοής μετά την εμφάνιση της BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
196	Θάνατος μετά την VAP από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
197	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την BSI από ανθεκτικά στις καρβαπενέμες GRAM negative παθογόνα
198	BSI από ευαίσθητα στις καρβαπενέμες GRAM negative παθογόνα
199	Διάρκεια νοσηλείας από την εισαγωγή μέχρι την εμφάνιση της BSI από ευαίσθητα στις καρβαπενέμες GRAM negative παθογόνα
200	Διάρκεια νοσηλείας μετά την εμφάνιση της BSI από ευαίσθητα στις καρβαπενέμες GRAM negative παθογόνα
201	Διάρκεια μηχανικής αναπνοής μετά την εμφάνιση της BSI από ευαίσθητα στις καρβαπενέμες GRAM negative παθογόνα
202	Θάνατος μετά την BSI από ευαίσθητα στις καρβαπενέμες GRAM negative παθογόνα
203	Σε πόσες ημέρες πέθανε ο άρρωστος μετά την BSI από ευαίσθητα στις καρβαπενέμες GRAM negative παθογόνα
204	Ακατάλληλη εμπειρική αγωγή την ημέρα διάγνωσης της BSI

aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
1	83	1	7	0	0	0	0	0	27	2	0	0	0	0	0
2	53	1	16	0	1	0	0	0	16	0	0	0	0	0	0
3	66	1	34	0	0	0	0	0	17	3	1	0	1	29	0
4	86	2	8	0	0	0	0	1	12	0	0	0	0	0	0
5	80	1	16	0	0	0	0	0	14	1	0	0	1	5	0
6	70	1	41	0	1	1	0	0	11	4	0	0	0	0	1
7	48	2	21	0	0	0	0	0	12	1	0	0	0	0	0
8	39	2	4	0	0	0	0	0	12	2	0	1	0	0	0
9	72	1	19	0	1	0	0	0	20	3	0	1	0	0	0
10	80	2	8	0	0	0	1	0	12	4	0	0	0	0	0
11	69	1	43	0	0	0	0	0	22	1	0	0	1	19	1
12	75	1	12	0	1	0	0	0	18	1	0	0	0	0	0
13	71	1	29	0	1	1	0	0	13	3	0	0	1	12	0
14	71	2	38	0	1	1	0	0	17	3	0	0	1	29	1
15	73	2	12	0	1	0	0	0	12	1	0	0	0	0	0
16	72	1	55	0	0	0	0	0	12	2	0	0	1	9	1
17	52	1	21	1	1	0	0	0	13	1	0	0	0	0	0
18	81	2	8	0	0	1	0	0	21	3	1	1	1	3	0
19	47	2	8	0	1	0	0	0	20	NA	1	1	0	0	0
20	79	2	6	0	1	0	0	0	25	2	0	1	0	0	0
21	31	2	41	1	0	0	0	0	13	NA	0	0	1	26	0
22	78	1	15	0	1	1	0	0	19	1	0	1	0	0	0
23	68	1	15	0	0	0	0	0	10	1	1	0	1	6	0
24	52	1	19	0	1	0	0	0	15	1	0	0	0	0	0
25	82	2	52	0	1	0	0	0	22	2	0	0	1	25	0
26	63	1	7	0	0	0	0	0	13	1	1	1	0	0	0
27	67	2	81	1	1	0	0	0	15	1	0	1	0	0	0
28	21	1	7	0	0	0	0	0	15	1	1	1	0	0	0
29	36	2	25	0	1	0	0	0	10	NA	0	0	1	5	0
30	78	2	9	0	1	0	0	0	20	1	0	1	0	0	0
31	67	1	57	0	0	1	0	0	18	2	0	1	0	0	0
32	81	1	8	0	1	0	0	0	20	3	0	1	0	0	0
33	78	1	15	1	0	0	0	0	20	NA	0	1	1	8	0
34	38	1	25	0	1	0	0	0	13	NA	1	1	1	5	0
35	83	1	6	0	1	0	0	0	12	2	0	1	0	0	0
36	74	1	8	0	1	0	0	0	11	1	0	1	0	0	0
37	76	1	24	0	1	0	1	0	19	5	1	1	0	0	1
38	73	2	10	0	0	0	0	0	21	1	0	1	1	6	0
39	81	2	17	0	0	0	0	0	17	1	0	0	0	0	0
40	65	1	7	1	0	0	1	0	18	3	0	1	0	0	0
41	73	1	8	0	1	0	0	0	15	2	0	1	0	0	0
42	64	1	7	0	0	0	0	0	17	2	0	1	0	0	0
43	79	1	7	0	0	0	1	0	20	4	0	1	0	0	0
44	75	1	14	0	1	0	0	0	17	1	0	0	1	8	0
45	35	1	11	1	0	0	0	0	16	2	0	1	1	9	0
46	71	1	19	0	0	0	0	0	18	1	0	1	1	14	0
47	29	1	13	0	1	0	0	0	9	6	0	1	0	0	0
48	70	2	41	0	0	0	0	1	9	1	1	1	1	5	0
49	67	2	24	0	0	0	0	0	16	2	0	1	1	21	0
50	79	1	26	0	0	0	0	0	19	3	0	0	1	6	0

aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
51	40	1	21	0	0	0	0	0	16	1	0	1	1	18	0
52	43	1	28	1	1	0	0	1	8	1	0	0	0	0	0
53	43	1	61	1	1	0	0	1	12	1	0	0	1	12	0
54	65	1	6	0	1	0	0	0	17	NA	0	1	0	0	0
55	73	2	5	0	0	0	0	0	18	3	0	0	0	0	1
56	80	2	10	0	1	0	0	0	16	1	1	1	0	0	0
57	33	2	13	1	0	0	0	1	12	NA	0	0	1	11	0
58	83	2	34	0	1	0	0	0	11	1	0	0	0	0	0
59	50	2	21	0	1	0	0	0	15	3	1	0	0	0	0
60	67	1	7	0	0	0	0	0	13	2	1	1	0	0	0
61	64	1	11	0	1	0	0	0	11	4	0	0	0	0	0
62	82	2	15	0	1	0	0	0	14	NA	1	1	1	6	0
63	48	1	19	1	0	0	0	0	11	NA	0	0	1	16	0
64	55	1	71	0	0	1	0	1	9	4	0	0	1	12	0
65	62	2	29	0	0	0	0	0	10	1	0	0	0	0	0
66	82	1	50	0	0	0	0	0	21	2	1	0	1	19	0
67	84	2	27	0	1	0	0	0	16	2	0	0	1	9	0
68	76	2	15	0	0	1	0	0	20	1	0	1	1	4	0
69	70	1	9	0	0	0	0	1	7	1	1	0	0	0	0
70	74	2	19	0	0	1	0	0	14	4	0	0	1	11	0
71	46	2	9	0	1	0	0	0	10	1	1	0	0	0	0
72	88	1	4	0	0	1	0	0	18	4	0	0	0	0	0
73	84	1	9	0	1	0	0	0	13	3	0	0	0	0	0
74	36	1	25	1	0	0	0	0	14	NA	0	0	1	7	0
75	42	1	33	0	0	0	0	0	14	1	0	0	1	12	0
76	73	2	16	0	0	0	0	0	15	1	0	0	1	6	0
77	77	2	12	0	1	0	0	0	15	1	0	0	1	3	0
78	75	2	32	0	0	0	1	0	10	4	0	0	1	11	0
79	44	1	14	0	1	0	1	0	20	4	1	0	0	0	0
80	59	1	30	0	1	0	0	0	13	1	0	1	1	13	0
aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
81	70	1	19	0	1	1	0	0	19	3	0	1	0	0	0
82	56	1	13	0	1	0	0	0	16	1	0	1	0	0	0
83	67	1	17	1	1	0	0	0	15	NA	0	1	0	0	0
84	91	1	8	0	1	0	0	0	20	1	0	1	0	0	0
85	82	2	18	0	0	0	1	0	19	3	0	1	1	10	0
86	72	1	13	0	1	0	0	0	30	2	0	1	0	0	0
87	70	2	4	0	0	0	0	0	25	4	1	1	0	0	0
88	65	1	33	1	0	1	1	0	16	5	1	1	1	30	0
89	65	1	4	1	0	0	1	0	31	2	0	1	0	0	0
90	66	1	51	0	1	1	0	0	9	3	0	1	1	31	1
91	44	1	26	0	1	0	0	0	10	2	0	1	0	0	0
92	75	2	16	0	1	0	0	0	26	1	0	1	1	12	0
93	80	1	68	0	1	1	0	0	22	2	1	1	1	6	0
94	73	1	20	1	0	0	0	1	19	NA	0	1	1	10	0
95	78	1	6	0	1	1	0	0	19	4	0	1	0	0	0
96	63	1	8	0	1	1	0	0	18	4	0	1	0	0	0
97	28	2	16	0	0	0	0	0	12	2	0	1	1	6	0
98	82	1	6	0	0	0	0	0	19	2	1	1	0	0	0
99	55	1	12	1	0	0	0	0	21	NA	0	0	1	5	0
100	80	2	11	0	1	0	0	0	17	1	0	1	0	0	0
101	86	1	48	0	0	0	0	0	14	3	0	1	1	11	0
102	76	1	11	1	1	0	1	0	20	4	1	1	0	0	0
103	79	1	9	1	0	0	0	0	26	3	0	1	0	0	0
104	70	2	71	0	1	1	1	0	10	4	0	0	1	45	1
105	69	2	80	0	1	0	1	0	22	1	0	1	1	4	0
106	67	1	27	0	0	0	0	0	33	2	0	1	1	14	0
107	78	1	44	0	0	1	0	0	20	3	0	1	1	9	0
108	80	2	13	0	0	0	0	0	21	2	1	1	0	0	0
109	67	2	17	0	1	1	0	0	16	4	0	1	1	7	0
110	72	2	30	0	1	0	0	0	18	1	0	1	1	24	1

aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
111	77	1	29	0	1	1	0	0	14	2	0	1	1	20	0
112	73	1	11	0	1	0	1	0	11	2	0	1	0	0	0
113	60	1	5	0	1	0	0	0	17	1	0	1	0	0	0
114	87	1	24	0	1	0	0	0	15	2	0	1	0	0	0
115	84	1	13	0	1	0	0	0	17	2	0	1	0	0	0
116	56	1	26	0	1	0	0	1	13	1	0	0	1	6	0
117	33	1	33	0	0	0	0	0	13	NA	0	0	1	18	0
118	73	1	36	1	0	0	0	0	15	1	1	0	1	15	1
119	82	1	52	0	1	0	0	0	14	1	0	1	1	36	0
120	80	1	21	0	1	0	0	0	14	1	0	0	0	0	0
121	76	1	27	1	0	1	0	0	15	3	0	1	0	0	0
122	75	1	12	0	1	0	0	0	13	2	1	1	0	0	1
123	75	2	28	1	0	0	0	0	17	2	0	1	1	5	1
124	35	1	10	1	0	0	0	1	7	1	0	0	0	0	0
125	39	2	20	0	0	0	0	0	8	NA	1	0	1	12	0
126	79	1	15	0	1	0	0	0	16	1	0	0	0	0	0
127	76	1	15	0	0	0	0	1	24	2	0	0	0	0	0
128	72	2	16	0	1	0	0	0	25	1	1	1	1	6	0
129	61	1	11	0	1	1	0	0	17	4	0	1	0	0	0
130	63	1	5	0	0	0	0	0	14	2	0	0	0	0	0
131	71	2	5	0	0	0	1	0	10	3	0	1	0	0	0
132	68	1	18	0	1	0	0	0	15	1	0	1	0	0	0
133	68	2	19	0	1	1	0	0	15	1	0	1	0	0	0
134	79	2	9	1	1	1	0	0	22	2	0	1	0	0	0
135	74	2	18	0	0	0	0	0	11	4	0	0	0	0	1
136	70	1	12	0	0	0	0	0	17	1	0	1	1	6	0
137	45	1	7	0	1	0	0	0	13	NA	0	1	0	0	0
138	62	2	5	0	0	0	0	0	13	2	0	1	0	0	0
139	64	2	9	1	1	0	0	0	24	2	1	1	0	0	0
140	81	1	9	0	1	1	0	0	18	3	0	1	1	5	0
aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
141	64	1	7	0	0	1	0	1	8	2	0	0	0	0	0
142	55	1	6	0	1	1	1	0	13	4	0	0	0	0	0
143	50	1	26	0	1	0	1	0	10	3	0	0	0	0	1
144	71	1	37	0	1	0	0	0	22	4	0	1	1	16	0
145	32	2	31	0	1	0	0	0	13	NA	0	0	1	7	0
146	47	1	43	0	1	0	0	0	9	1	1	0	0	0	1
147	78	2	10	0	0	0	0	0	14	3	0	0	0	0	0
148	85	2	7	0	1	0	0	0	17	2	0	1	0	0	0
149	81	2	18	0	0	0	0	0	15	2	0	1	0	0	0
150	76	1	8	0	0	0	0	0	14	2	1	0	0	0	0
151	41	1	30	1	1	0	0	0	14	1	0	0	1	9	0
152	81	1	10	0	1	0	0	0	26	1	0	1	0	0	0
153	72	1	13	0	1	0	0	0	10	2	0	1	1	6	0
154	73	1	8	0	1	0	0	0	18	2	0	1	0	0	0
155	63	2	42	0	0	0	0	1	8	1	0	0	1	5	0
156	76	1	19	0	1	0	0	0	23	2	1	1	1	7	0
157	58	2	11	1	0	0	0	1	22	NA	0	1	1	5	0
158	26	2	8	0	0	0	0	0	21	3	0	0	0	0	0
159	74	1	40	1	0	1	0	0	19	2	0	0	1	13	0
160	80	2	11	0	1	0	0	0	12	NA	0	0	0	0	0
161	83	2	23	1	0	1	0	0	32	3	1	0	1	16	0
162	39	1	35	0	0	0	0	1	24	NA	0	0	0	0	1
163	22	2	9	0	0	0	0	1	17	NA	0	0	0	0	0
164	23	2	29	0	0	0	0	1	15	NA	0	0	1	12	0
165	58	1	31	1	0	0	0	1	20	1	0	0	1	9	0
166	25	2	20	1	0	0	0	1	19	NA	0	1	0	0	0
167	67	2	46	1	0	1	0	1	12	2	0	0	1	9	0
168	18	2	5	0	0	0	0	0	18	NA	0	1	0	0	0
169	29	2	31	0	0	0	0	1	10	1	0	0	1	9	0
170	88	1	7	1	0	0	0	0	14	3	1	1	0	0	0

aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
171	38	1	46	1	0	0	1	1	19	2	0	1	1	5	0
172	72	2	43	1	0	0	1	0	13	3	0	1	1	14	1
173	58	2	17	0	1	1	0	0	12	1	0	0	1	7	0
174	50	1	122	0	1	0	0	1	26	NA	0	1	1	3	1
175	30	2	19	0	1	0	0	1	29	2	0	0	0	0	0
176	54	2	6	1	0	0	1	0	17	4	1	1	0	0	0
177	83	2	27	1	0	0	1	1	25	3	1	1	1	9	0
178	74	2	11	1	0	0	0	0	17	5	0	1	0	0	0
179	67	2	29	1	0	0	1	0	11	4	0	1	1	8	0
180	24	2	20	0	0	0	0	0	19	NA	1	0	0	0	0
181	18	2	9	1	0	0	0	0	8	NA	0	0	0	0	0
182	36	2	37	1	0	0	0	0	7	NA	0	0	1	13	0
183	56	2	8	1	0	0	1	0	19	5	1	1	0	0	0
184	49	2	13	0	0	0	0	0	6	NA	0	0	0	0	0
185	37	2	17	1	0	0	0	0	21	NA	0	1	1	3	0
186	82	2	10	1	0	0	1	0	18	3	0	1	0	0	0
187	65	2	6	0	1	0	0	1	29	1	1	1	0	0	0
188	54	2	14	1	1	0	1	0	14	3	1	1	0	0	0
189	57	2	74	1	0	0	1	0	7	4	0	0	1	61	0
190	72	2	11	1	0	0	1	0	9	3	0	0	0	0	0
191	34	2	12	1	0	0	0	0	4	NA	0	0	0	0	0
192	26	2	5	0	0	0	0	1	22	NA	1	1	0	0	0
193	41	2	9	1	0	0	0	0	6	NA	0	0	0	0	0
194	88	2	29	1	0	1	0	0	29	1	1	1	1	10	0
195	51	2	12	1	0	0	1	0	21	5	0	0	0	0	0
196	54	1	4	0	0	0	0	0	12	NA	1	1	0	0	0
197	78	2	8	1	0	0	1	1	35	2	1	1	1	6	0
198	35	2	9	1	0	0	0	0	4	NA	0	0	0	0	0
199	25	2	8	0	0	0	0	1	25	NA	0	1	0	0	0
200	36	2	30	1	0	0	0	0	11	NA	0	0	0	0	0
aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI
201	22	2	50	0	0	0	0	1	13	NA	0	0	1	11	1
202	55	1	7	1	0	0	0	0	11	NA	0	0	0	0	0
203	19	1	11	0	0	0	1	0	26	3	0	1	1	6	1
204	30	2	12	1	0	0	0	0	15	NA	1	1	1	5	0
205	60	1	102	1	1	0	0	0	11	1	0	0	0	0	0
206	64	1	107	0	0	0	0	1	28	NA	0	0	1	16	1
207	60	1	9	1	0	0	1	0	14	2	0	0	0	0	0
208	60	2	10	1	0	0	1	0	7	5	0	0	0	0	0
209	21	2	17	0	1	0	0	0	14	NA	1	0	1	9	0
210	42	2	21	1	0	0	0	0	12	NA	1	0	1	12	0
211	49	2	16	1	0	0	0	1	20	NA	0	0	1	6	0
212	23	2	16	1	1	0	0	0	11	NA	1	0	0	0	0
213	60	2	14	0	1	0	0	1	34	2	0	1	1	6	0
214	71	2	61	1	1	1	1	0	23	6	0	0	1	11	0
215	28	1	14	0	0	0	0	1	19	NA	1	1	0	0	0
216	71	2	15	0	0	0	1	0	12	3	0	0	0	0	0
217	77	2	24	1	0	NA	NA	0	11	2	1	0	1	11	1
218	82	1	58	1	1	NA	NA	0	20	5	1	1	0	0	1
219	52	2	10	0	1	NA	NA	1	9	3	1	0	0	0	0
220	41	1	11	0	1	NA	NA	1	9	NA	0	0	0	0	0
221	30	1	9	1	0	NA	NA	0	4	NA	0	0	0	0	0
222	43	1	6	1	0	NA	NA	0	10	2	0	0	0	0	0
223	47	1	20	0	1	NA	NA	0	15	5	0	0	0	0	0
224	79	1	9	0	0	NA	NA	0	17	2	0	0	0	0	0
225	84	1	8	1	0	NA	NA	0	27	3	0	0	0	0	0
226	41	2	21	1	0	NA	NA	1	10	NA	1	0	1	8	0
227	64	1	17	1	1	NA	NA	1	23	6	0	1	0	0	0
228	79	1	8	1	0	NA	NA	0	18	4	0	0	0	0	0
229	62	2	10	1	0	NA	NA	1	11	2	0	0	0	0	0
230	60	1	16	1	0	NA	NA	0	8	NA	0	0	0	0	0

aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI	t
231	59	1	20	1	0	NA	NA	1	8	2	0	0	1	13	0	0
232	23	1	6	0	0	NA	NA	0	5	NA	0	0	0	0	0	0
233	61	1	20	0	1	NA	NA	0	11	2	0	0	0	0	0	0
234	62	2	36	1	1	NA	NA	1	15	2	1	1	0	0	0	0
235	80	2	6	1	1	NA	NA	0	13	2	0	0	0	0	0	0
236	57	2	31	1	0	NA	NA	1	9	NA	0	0	1	13	0	0
237	60	1	9	0	1	NA	NA	1	4	4	0	1	0	0	0	0
238	78	1	20	0	1	NA	NA	1	29	2	0	1	1	12	0	0
239	17	2	8	1	0	NA	NA	0	18	2	0	0	0	0	0	0
240	61	2	6	0	0	NA	NA	1	7	NA	0	0	0	0	0	0
241	78	1	7	0	0	NA	NA	0	13	4	0	0	0	0	0	0
242	86	2	21	1	0	NA	NA	0	13	NA	0	0	1	14	0	0
243	75	1	5	1	0	NA	NA	0	24	4	0	0	0	0	0	0
244	35	2	14	0	0	NA	NA	1	13	NA	0	0	0	0	0	0
245	73	2	80	0	1	NA	NA	0	16	2	0	1	1	21	0	0
246	63	1	7	0	0	NA	NA	0	13	4	0	0	0	0	0	0
247	34	1	5	0	0	NA	NA	0	8	NA	0	0	0	0	0	0
248	64	1	21	1	1	NA	NA	1	16	5	0	0	1	8	0	0
249	69	1	44	0	1	NA	NA	1	18	2	1	1	1	15	0	0
250	20	1	39	1	0	NA	NA	1	10	NA	0	0	1	22	0	0
251	78	1	29	0	1	NA	NA	0	21	2	0	0	1	9	0	0
252	67	2	29	1	1	NA	NA	0	12	2	0	0	1	9	0	0
253	62	1	28	1	1	NA	NA	0	14	2	0	1	1	16	0	0
254	88	1	18	0	1	NA	NA	1	20	4	0	0	1	8	0	0
255	87	2	6	0	1	NA	NA	0	15	4	0	1	0	0	0	0
256	87	2	12	0	1	NA	NA	0	17	2	1	1	0	0	1	0
257	65	1	6	1	0	NA	NA	0	12	3	0	0	0	0	0	0
258	39	1	29	1	1	NA	NA	1	10	NA	0	1	1	5	0	0
259	57	1	4	0	0	NA	NA	1	15	2	0	1	0	0	0	0
260	16	1	8	0	0	NA	NA	0	4	NA	0	0	0	0	0	0
aa	age	sex	length	type	infection	diabetes	malign	koma	Apache2	Charlson	shock	death	VAP	VAPtime	UTI	t
261	33	1	8	1	0	NA	NA	0	3	NA	0	0	0	0	0	0
262	36	1	9	1	0	NA	NA	1	11	NA	0	0	0	0	0	0
263	35	1	6	0	0	NA	NA	0	4	NA	0	0	0	0	0	0
264	35	1	16	0	0	NA	NA	1	5	2	0	0	1	7	0	0
265	62	1	15	0	1	NA	NA	1	13	NA	0	0	0	0	0	0
266	59	1	14	1	0	NA	NA	0	19	5	0	0	0	0	0	0
267	27	2	12	0	1	0	0	1	28	2	0	0	0	0	0	0
268	33	1	6	1	0	0	1	0	7	2	0	0	0	0	0	0
269	83	1	12	1	1	0	0	0	24	3	1	0	0	0	0	0
270	79	1	15	0	1	1	0	1	34	4	1	1	0	0	0	0
271	78	1	14	1	1	1	0	0	14	4	1	0	0	0	0	0
272	17	2	27	0	0	0	0	0	20	2	0	0	0	0	0	0
273	55	1	9	1	1	1	0	0	14	3	0	0	0	0	0	0
274	48	2	8	1	1	0	0	0	10	2	0	0	0	0	0	0
275	50	2	6	0	0	0	1	0	19	2	0	1	0	0	0	0
276	60	2	6	0	0	1	0	0	18	4	0	1	0	0	0	0
277	63	1	44	0	0	0	1	0	12	2	0	0	1	11	0	0
278	64	1	15	0	0	1	0	0	22	2	1	1	0	0	0	0
279	78	2	7	0	1	1	0	0	15	3	1	0	0	0	0	0
280	65	2	7	1	0	0	1	0	12	3	0	0	0	0	0	0
281	79	1	7	1	0	0	0	0	16	6	1	0	0	0	0	0
282	76	1	13	1	0	0	0	0	12	3	0	0	0	0	0	0
283	80	2	48	1	0	0	0	0	13	3	0	0	0	0	0	0
284	74	1	14	0	0	0	1	0	18	5	0	1	1	13	0	0
285	47	1	10	1	0	0	1	0	12	3	1	0	0	0	0	0
286	68	2	24	0	0	1	0	0	17	2	0	0	1	15	0	0
287	75	1	11	1	0	0	0	0	13	2	0	0	0	0	0	0
288	56	1	21	0	0	1	0	0	11	4	1	1	0	0	0	0
289	82	2	7	1	0	0	1	0	11	4	0	0	0	0	0	0
290	74	2	35	0	1	1	1	0	23	6	1	0	1	21	0	0
291	84	1	10	0	0	0	0	0	23	8	0	1	0	0	0	0
292	75	1	11	1	0	1	1	1	33	4	1	1	0	0	0	0
293	89	2	9	1	0	0	0	0	19	2	1	0	0	0	0	0
294	82	1	41	0	0	1	1	0	22	6	0	1	1	5	0	0

aa	UTitime	BSI	BSitime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
1	0	0	0	0	0	0	0	1	0	0	0	6	0
2	0	0	0	0	0	0	0	0	0	0	1	16	0
3	0	1	17	2	1	0	0	1	0	0	0	29	1
4	0	0	0	0	0	0	0	0	1	0	0	8	0
5	0	0	0	6	1	0	0	0	0	0	1	5	1
6	4	0	0	0	0	1	0	0	0	0	1	41	0
7	0	0	0	0	0	0	0	0	0	0	1	21	0
8	0	0	0	0	0	1	0	0	0	0	1	4	0
9	0	0	0	0	0	1	1	0	0	0	0	19	0
10	0	0	0	0	0	1	0	0	0	0	1	8	0
11	9	0	0	4	0	0	0	0	0	0	1	19	0
12	0	0	0	0	0	0	0	0	0	0	1	12	0
13	0	1	20	4	1	0	0	0	0	0	1	12	1
14	31	1	9	4	0	0	0	0	0	0	1	29	0
15	0	0	0	0	0	0	0	0	0	0	1	10	0
16	24	1	10	1	0	0	0	0	0	0	1	9	0
17	0	0	0	0	0	1	0	0	0	0	1	18	0
18	0	1	7	2	1	0	0	1	0	0	0	3	1
19	0	0	0	0	0	1	0	0	0	0	1	8	0
20	0	1	5	0	0	0	0	0	0	0	1	6	0
21	0	0	0	4	1	0	1	0	0	0	0	26	1
22	0	0	0	0	0	1	0	0	0	0	1	15	0
23	0	0	0	1	1	0	0	0	0	0	1	6	1
24	0	0	0	0	0	0	0	0	0	0	1	19	0
25	0	1	23	1	1	1	0	0	0	0	1	25	1
26	0	0	0	0	0	0	0	0	0	0	1	7	0
27	0	0	0	0	0	1	1	0	0	0	0	81	0
28	0	0	0	0	0	1	0	0	0	0	1	7	0
29	0	0	0	1	1	0	0	0	0	0	1	5	1
30	0	0	0	0	0	0	0	0	0	0	1	9	0
aa	UTitime	BSI	BSitime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
31	0	1	5	0	0	1	0	0	0	0	1	57	0
32	0	0	0	0	0	0	0	1	0	0	0	8	0
33	0	0	0	4	0	0	0	0	0	0	1	8	0
34	0	1	11	4	0	1	0	0	0	0	1	5	0
35	0	0	0	0	0	0	0	1	0	0	0	6	0
36	0	0	0	0	0	0	0	0	0	0	1	8	0
37	11	0	0	0	0	0	0	0	0	0	1	24	0
38	0	0	0	1	1	0	0	0	0	0	1	6	1
39	0	1	9	0	0	0	0	0	0	0	1	17	0
40	0	0	0	0	0	0	1	0	0	0	0	7	0
41	0	0	0	0	0	1	0	0	0	0	1	6	0
42	0	0	0	0	0	1	0	0	0	0	1	5	0
43	0	0	0	0	0	1	0	0	0	0	1	7	0
44	0	1	10	1	1	1	0	0	0	0	1	8	1
45	0	0	0	4	1	1	0	0	0	0	1	9	1
46	0	1	14	1	0	0	0	0	0	0	1	14	0
47	0	0	0	0	0	1	0	0	0	0	1	13	0
48	0	1	27	1	1	0	0	0	0	1	0	41	1
49	0	0	0	1	1	0	0	0	0	0	1	21	1
50	0	0	0	1	1	1	0	0	0	0	1	6	1
51	0	0	0	1	0	0	0	0	0	0	1	18	0
52	0	1	15	0	0	0	0	0	1	0	0	27	0
53	0	0	0	3	1	1	0	0	1	0	0	12	1
54	0	0	0	0	0	0	0	0	0	0	1	6	0
55	4	0	0	0	0	0	0	0	0	0	1	4	0
56	0	0	0	0	0	0	0	0	0	0	1	10	0
57	0	0	0	1	0	0	0	0	1	0	0	11	0
58	0	1	6	0	0	0	0	0	0	0	1	34	0
59	0	0	0	0	0	0	1	0	0	0	0	21	0
60	0	1	4	0	0	0	1	0	0	0	0	7	0

aa	UTItime	BSI	BSItime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	iniththerapy
61	0	0	0	0	0	0	0	0	0	0	1	9	0
62	0	0	0	1	1	0	0	0	0	0	1	6	1
63	0	1	16	4	0	0	0	0	0	0	1	16	0
64	0	1	29	1	1	0	0	0	0	0	1	12	1
65	0	0	0	0	0	0	0	0	0	0	1	29	0
66	0	1	30	1	1	0	0	0	0	0	1	19	1
67	0	0	0	5	0	0	0	0	0	0	1	9	0
68	0	0	0	1	1	0	0	0	0	0	1	4	1
69	0	0	0	0	0	0	0	0	0	1	0	9	0
70	0	0	0	4	0	0	1	0	0	0	0	11	0
71	0	0	0	0	0	0	0	0	0	0	1	7	0
72	0	0	0	0	0	0	0	0	0	0	1	4	0
73	0	0	0	0	0	0	0	0	0	0	1	8	0
74	0	1	3	6	1	1	1	0	0	0	0	7	1
75	0	1	23	2	1	1	0	0	0	0	1	12	1
76	0	0	0	4	1	0	0	0	0	1	0	6	1
77	0	0	0	1	1	0	0	0	0	0	1	3	1
78	0	0	0	1	0	0	0	0	0	0	1	11	0
79	0	0	0	0	0	1	0	0	0	0	1	7	0
80	0	1	21	1	1	0	0	0	0	0	1	13	1
81	0	0	0	0	0	1	0	0	0	0	1	19	0
82	0	0	0	0	0	1	0	0	0	0	1	8	0
83	0	0	0	0	0	1	0	0	0	0	1	17	0
84	0	0	0	0	0	1	0	0	0	0	1	8	0
85	0	0	0	4	1	1	0	0	0	0	1	10	1
86	0	0	0	0	0	1	0	0	0	0	1	13	0
87	0	0	0	0	0	0	0	0	0	0	1	4	0
88	0	1	18	3	1	0	1	0	0	0	0	30	1
89	0	0	0	0	0	1	0	0	0	0	1	4	0
90	46	1	42	1	0	0	0	0	0	0	1	31	0
aa	UTItime	BSI	BSItime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	iniththerapy
91	0	0	0	0	0	1	0	0	0	0	1	26	0
92	0	1	5	1	0	1	0	0	0	0	1	12	0
93	0	1	25	0	1	0	0	0	0	0	1	6	1
94	0	0	0	1	0	1	0	0	1	0	0	10	0
95	0	0	0	0	0	1	0	0	0	0	1	6	0
96	0	0	0	0	0	0	0	0	0	0	1	8	0
97	0	1	6	1	1	0	0	0	0	0	1	6	1
98	0	0	0	0	0	1	0	0	0	0	1	6	0
99	0	0	0	1	1	1	1	0	0	0	0	5	1
100	0	0	0	0	0	1	0	0	0	0	1	11	0
101	0	1	14	6	1	0	0	0	0	0	1	11	1
102	0	1	6	0	0	1	0	0	0	0	1	11	0
103	0	0	0	0	0	1	0	0	0	0	1	9	0
104	30	1	46	6	1	0	0	0	0	0	1	45	1
105	0	1	53	1	1	0	0	0	0	0	1	4	1
106	0	0	0	1	1	0	1	0	0	0	0	14	1
107	0	1	23	1	1	0	0	0	0	0	1	9	1
108	0	0	0	0	0	0	0	1	0	0	0	13	0
109	0	0	0	1	1	1	0	0	0	0	1	7	1
110	6	0	0	1	1	0	0	0	0	0	1	24	1
111	0	1	11	4	1	0	0	0	0	0	1	20	1
112	0	0	0	0	0	0	0	0	0	0	1	11	0
113	0	0	0	0	0	0	0	0	0	0	1	5	0
114	0	0	0	0	0	0	0	1	0	0	0	24	0
115	0	0	0	0	0	0	0	1	0	0	0	13	0
116	0	0	0	1	1	0	0	0	0	1	0	6	1
117	0	0	0	1	1	0	0	0	0	0	1	18	1
118	11	0	0	1	0	0	0	1	0	0	0	15	0
119	0	1	11	1	0	0	0	0	0	0	1	36	0
120	0	0	0	0	0	0	0	0	0	0	1	21	0

aa	UTItime	BSI	BSItime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
121	0	1	10	0	0	0	0	1	0	0	0	27	0
122	6	0	0	0	0	0	0	1	0	0	0	12	0
123	14	1	9	1	1	0	1	0	0	0	0	5	1
124	0	0	0	0	0	0	0	0	1	0	1	10	0
125	0	0	0	1	1	0	0	0	0	0	1	12	1
126	0	0	0	0	0	0	0	0	0	0	1	15	0
127	0	1	4	0	0	0	0	0	0	1	0	15	0
128	0	0	0	4	1	0	0	0	0	0	1	6	1
129	0	0	0	0	0	0	0	0	0	0	1	11	0
130	0	0	0	0	0	0	0	0	0	0	1	5	0
131	0	0	0	0	0	0	0	0	0	0	1	5	0
132	0	0	0	0	0	0	0	0	0	0	1	18	0
133	0	1	11	0	0	0	0	0	0	0	1	19	0
134	0	0	0	0	0	0	1	0	0	0	0	9	0
135	7	0	0	0	0	0	0	0	0	0	1	11	0
136	0	0	0	1	1	0	0	0	0	0	1	6	1
137	0	0	0	0	0	0	0	0	0	0	1	7	0
138	0	0	0	0	0	0	0	1	0	0	0	5	0
139	0	0	0	0	0	0	0	0	0	0	1	9	0
140	0	0	0	4	1	0	0	0	0	0	1	5	1
141	0	0	0	0	0	0	0	0	0	0	1	6	0
142	0	0	0	0	0	0	0	1	0	0	0	5	0
143	16	1	14	0	0	0	0	0	0	0	1	26	0
144	0	1	26	1	1	0	0	1	0	0	0	16	1
145	0	0	0	1	0	0	0	0	0	0	1	7	0
146	9	1	11	0	0	0	0	0	0	0	1	43	0
147	0	0	0	0	0	0	0	0	0	0	0	9	0
148	0	0	0	0	0	0	0	0	0	0	1	7	0
149	0	1	4	0	0	0	0	0	0	0	1	18	0
150	0	0	0	0	0	0	0	0	0	0	1	8	0
aa	UTItime	BSI	BSItime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
151	0	0	0	1	1	0	1	0	0	0	0	9	1
152	0	0	0	0	0	0	0	0	0	0	1	10	0
153	0	1	5	1	1	0	0	0	0	0	1	6	1
154	0	1	4	0	0	0	0	1	0	0	0	8	0
155	0	0	0	1	1	0	0	1	0	0	0	5	1
156	0	1	7	4	1	0	0	0	1	0	0	7	1
157	0	1	7	1	0	0	0	0	0	1	0	5	0
158	0	0	0	0	0	1	0	0	0	0	1	8	0
159	0	1	13	1	0	1	1	0	0	0	0	13	0
160	0	0	0	0	0	0	0	0	0	1	0	11	0
161	0	0	0	5	0	0	0	1	0	0	0	16	0
162	9	1	21	0	0	0	0	0	1	0	0	35	0
163	0	0	0	0	0	0	0	0	1	0	0	8	0
164	0	1	12	1	0	0	0	0	1	0	0	12	0
165	0	1	26	1	0	0	0	0	0	1	0	9	0
166	0	0	0	0	0	1	0	0	1	0	0	20	0
167	0	1	15	1	0	0	0	0	0	1	0	9	0
168	0	0	0	0	0	0	0	0	0	1	0	5	0
169	0	1	9	4	0	0	0	0	1	0	0	9	0
170	0	0	0	0	0	0	1	0	0	0	0	7	0
171	0	0	0	6	0	1	0	0	0	1	0	5	0
172	33	0	0	3	0	1	1	0	0	0	0	14	0
173	0	0	0	1	0	1	0	0	0	1	0	7	0
174	22	1	3	4	0	1	0	0	0	1	0	3	0
175	0	1	14	0	0	0	0	0	0	0	1	16	0
176	0	0	0	0	0	1	0	0	0	0	1	6	0
177	0	0	0	1	0	1	1	0	0	0	0	9	0
178	0	0	0	0	0	0	0	1	0	0	0	9	0
179	0	0	0	1	0	0	0	1	0	0	0	8	0
180	0	1	8	0	0	0	0	0	1	0	0	19	0

aa	Uttime	BSI	BSitime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
181	0	0	0	0	0	0	0	0	1	0	0	9	0
182	0	1	8	4	0	0	0	0	1	0	0	13	0
183	0	0	0	0	0	1	0	0	0	1	0	8	0
184	0	1	7	0	0	0	0	0	1	0	0	13	0
185	0	0	0	5	0	0	0	0	1	0	0	3	0
186	0	0	0	0	0	1	1	0	0	0	0	10	0
187	0	0	0	0	0	1	0	0	0	1	0	6	0
188	0	1	13	0	0	1	0	0	0	0	1	14	0
189	0	1	15	4	0	1	0	0	0	0	1	61	0
190	0	0	0	0	0	0	0	0	0	0	1	11	0
191	0	0	0	0	0	0	0	0	1	0	0	12	0
192	0	0	0	0	0	0	0	0	1	0	0	5	0
193	0	0	0	0	0	1	0	0	0	0	1	9	0
194	0	1	22	1	0	1	1	0	0	0	0	10	0
195	0	0	0	0	0	0	0	0	0	0	1	12	0
196	0	0	0	0	0	0	0	1	0	0	0	4	0
197	0	0	0	3	0	1	0	0	0	1	0	6	0
198	0	0	0	0	0	0	0	0	1	0	0	9	0
199	0	0	0	0	0	0	0	0	0	1	0	8	0
200	0	1	6	0	0	0	0	0	1	0	0	30	0
201	43	0	0	4	0	0	0	0	1	0	0	11	0
202	0	0	0	0	0	0	0	0	0	1	0	7	0
203	6	1	6	4	0	1	0	1	0	0	0	6	0
204	0	1	5	4	0	0	0	0	1	0	0	5	0
205	0	1	52	0	0	1	1	0	0	0	0	80	0
206	28	1	87	4	0	1	0	0	0	1	0	16	0
207	0	0	0	0	0	0	0	0	0	0	1	9	0
208	0	0	0	0	0	1	1	0	0	0	0	10	0
209	0	1	4	4	0	0	0	0	1	0	0	9	0
210	0	1	14	1	0	0	0	0	0	1	0	12	0
aa	Uttime	BSI	BSitime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
211	0	0	0	1	0	0	0	0	1	0	0	6	0
212	0	0	0	0	0	1	0	0	1	0	0	16	0
213	0	1	9	1	0	0	0	0	0	0	1	6	0
214	0	1	24	4	0	0	0	0	0	0	1	11	0
215	0	0	0	0	0	0	0	0	0	1	0	14	0
216	0	0	0	0	0	0	0	0	0	0	1	15	0
217	5	0	0	4	1	1	0	1	0	0	0	11	1
218	3	1	16	0	0	1	0	0	0	0	1	58	0
219	0	0	0	0	0	1	0	0	0	0	1	7	0
220	0	1	6	0	0	1	0	0	0	0	1	10	0
221	0	0	0	0	0	0	0	0	1	0	0	9	0
222	0	0	0	0	0	0	0	0	0	1	0	4	0
223	0	1	9	0	0	0	0	0	0	0	1	15	0
224	0	0	0	0	0	1	0	1	0	0	0	7	0
225	0	0	0	0	0	1	0	0	0	0	1	6	0
226	0	1	4	1	0	0	0	0	1	0	0	8	0
227	0	1	6	0	0	1	0	0	0	0	1	17	0
228	0	0	0	0	0	1	0	0	0	0	1	4	0
229	0	0	0	0	0	0	0	0	0	1	0	9	0
230	0	0	0	0	0	0	0	0	1	0	0	11	0
231	0	0	0	2	0	0	0	0	0	1	0	13	0
232	0	0	0	0	0	0	0	0	1	0	0	4	0
233	0	1	11	0	0	1	0	0	0	0	1	15	0
234	0	1	20	0	0	0	0	0	1	0	0	25	0
235	0	0	0	0	0	1	0	0	1	0	0	4	0
236	0	1	14	4	1	1	0	0	1	0	0	13	1
237	0	0	0	0	0	0	0	0	0	0	1	9	0
238	0	0	0	1	1	1	0	1	0	0	0	12	1
239	0	0	0	0	0	0	0	0	0	0	1	5	0
240	0	0	0	0	0	0	0	1	0	0	0	4	0

aa	UTitime	BSI	BSitime	pathogensvap	antibiotics	prev	abdominal	cardio	multiply	neurological	pulmonary	durven	inittherapy
241	0	0	0	0	0	1	0	0	0	0	1	5	0
242	0	1	17	4	1	0	0	0	0	0	1	14	1
243	0	0	0	0	0	0	0	0	0	0	1	5	0
244	0	0	0	0	0	0	0	0	1	0	0	11	0
245	0	1	6	3	0	1	0	0	0	0	1	21	0
246	0	0	0	0	0	0	0	0	0	0	1	5	0
247	0	0	0	0	0	0	0	0	0	0	1	4	0
248	0	1	10	1	1	1	0	0	0	0	1	8	1
249	0	1	30	4	0	1	0	0	0	0	1	15	0
250	0	1	9	1	0	0	0	0	1	0	0	22	0
251	0	0	0	1	1	1	0	0	0	0	1	9	1
252	0	0	0	4	0	1	0	1	0	0	0	9	0
253	0	1	15	2	0	1	0	0	0	0	1	16	0
254	0	0	0	4	0	1	0	0	0	0	1	8	0
255	0	0	0	0	0	1	0	0	0	0	1	6	0
256	4	0	0	0	0	1	0	1	0	0	0	12	0
257	0	0	0	0	0	0	0	0	0	1	0	4	0
258	0	1	11	2	0	1	0	0	0	1	0	5	0
259	0	0	0	0	0	0	0	1	0	0	0	4	0
260	0	0	0	0	0	0	0	0	1	0	0	8	0
261	0	0	0	0	0	0	0	0	1	0	0	6	0
262	0	0	0	0	0	0	0	0	1	0	0	6	0
263	0	0	0	0	0	0	0	0	0	0	1	4	0
264	0	0	0	1	1	0	0	0	0	0	1	7	1
265	0	0	0	0	0	0	0	0	0	0	1	11	0
266	0	0	0	0	0	1	0	0	0	0	1	14	0
267	0	1	5	0	0	1	0	1	0	0	0	8	0
268	0	0	0	0	0	0	1	0	0	0	0	6	0
269	0	0	0	0	0	1	1	0	0	0	0	12	0
270	0	1	5	0	0	1	1	0	0	0	0	15	0
271	0	0	0	0	0	0	0	0	0	0	1	14	0
272	0	0	0	0	0	1	0	0	0	1	0	27	0
273	0	0	0	0	0	0	1	0	0	0	0	9	0
274	0	0	0	0	0	0	1	0	0	0	0	8	0
275	0	0	0	0	0	1	0	0	0	0	1	6	0
276	0	0	0	0	0	0	0	0	0	0	1	6	0
277	0	1	20	1	0	1	0	0	0	0	1	11	0
278	0	1	3	0	0	0	0	1	0	0	0	15	0
279	0	0	0	0	0	0	0	0	0	0	1	7	0
280	0	0	0	0	0	0	0	0	0	0	1	7	0
281	0	0	0	0	0	0	1	0	0	0	0	7	0
282	0	0	0	0	0	1	0	0	0	1	0	13	0
283	0	1	21	0	0	1	1	0	0	0	0	48	0
284	0	0	0	1	1	0	1	0	0	0	0	13	1
285	0	0	0	0	0	0	1	0	0	0	0	10	0
286	0	1	18	1	1	1	0	0	0	1	0	15	1
287	0	0	0	0	0	0	0	0	0	0	1	11	0
288	0	1	4	0	0	1	1	0	0	0	0	21	0
289	0	0	0	0	0	0	1	0	0	0	0	7	0
290	0	0	0	3	0	1	0	0	0	0	1	21	0
291	0	0	0	0	0	1	1	0	0	0	0	10	0
292	0	1	3	0	0	1	0	0	0	0	1	11	0
293	0	0	0	0	0	0	1	0	0	0	0	9	0
294	0	0	0	2	1	1	0	0	0	0	1	5	1

Παράρτημα Β΄

Κώδικας R

B.1 MSM

```
# DATA
data=read.csv("C:/My staff/newold/aaa/various/data/extra.csv",header=T)
data<-data.frame(extra[1],extra[3],extra[4],extra[5],extra[13],extra[21],extra[22],extra[23],
extra[24],extra[25],extra[26],extra[27],extra[86],extra[88],extra[140],extra[142],extra[172],
extra[108],extra[135],extra[136])
dataw=data
dataw$X=data$aa
dataw$none=as.numeric(data$VAP==0 & data$BSI==0 & data$UTI==0)
dataw$VAPalone=as.numeric(data$VAP==1 & data$BSI==0 & data$UTI==0)
dataw$BSIalone=as.numeric(data$VAP==0&data$BSI==1 & data$UTI==0)
dataw$UTIalone =as.numeric(data$VAP==0 & data$BSI==0 & data$UTI==1)
dataw$VAPBSI =as.numeric(data$VAP==1 & data$BSI==1 & data$UTI==0)
dataw$VAPUTI =as.numeric(data$VAP==1 & data$BSI==0 & data$UTI==1)
dataw$VAPBSIUTI=as.numeric(data$VAP==1 & data$BSI==1 & data$UTI==1)
dataw$BSIUTI =as.numeric(data$VAP==0 & data$BSI==1 & data$UTI==1)

# NONE
dataw0 =data[dataw$none==1,];
dataw0$status=1;
```

```

dataw0$time=0;
dataw0$aa=dataw0$aa+0.001;
dataw09=data[dataw$none==1,];
dataw09$status=dataw09$death*5+(1-dataw09$death)*4;
dataw09$time=dataw09$length;
dataw09$aa=dataw09$aa+0.009;

```

```

# VAP alone

```

```

dataw1=data[dataw$VAPalone==1,];
dataw1$status=1;
dataw1$time=0;
dataw1$aa=dataw1$aa+0.01;
dataw12=data[dataw$VAPalone==1,];
dataw12$status=2;
dataw12$time=dataw1$VAPtime;
dataw12$aa=dataw12$aa+0.0102;
dataw19=data[dataw$VAPalone==1,];
dataw19$status=dataw19$death*5+(1-dataw19$death)*4;
dataw19$time=dataw19$length;
dataw19$aa=dataw19$aa+0.019;

```

```

# BSI alone

```

```

dataw2 =data[dataw$BSIalone==1,];
dataw2$status=1;
dataw2$time=0;
dataw2$aa=dataw2$aa+0.02;
dataw23=data[data$BSIalone==1,];
dataw23$status=3;
dataw23$time=dataw23$BSItime;
dataw23$aa=dataw23$aa+0.023;

```

```

# UTI alone

```

```

dataw3 =data[data$UTIalone==1,];

```

```

dataw3$status=1;
dataw3$time=0;
dataw3$aa=dataw3$aa+0.03;
dataw33=data[data$UTIalone==1,];
dataw33$status=3;
dataw33$time=dataw33$UTItime;
dataw33$aa=dataw33$aa+0.033;

# VAP+BSI
dataw4 =data[dataw$VAPBSI==1,];
dataw4$status=1;
dataw4$time=0;
dataw4$aa=dataw4$aa+0.04;
dataw43=data[dataw$VAPBSI==1,];
dataw43$status=3;
dataw43$time=dataw43$BSItime;
dataw43$aa=dataw43$aa+0.043;
dataw42=data[dataw$VAPBSI==1,];
dataw42=dataw42[dataw42$VAPtime!=dataw42$BSItime,];
dataw42$status=2;
dataw42$time=dataw42$VAPtime;
dataw42$aa=dataw42$aa+0.0402;

# VAP+UTI
dataw5 =data[dataw$VAPUTI==1,];
dataw5$status=1;
dataw5$time=0;
dataw5$aa=dataw5$aa+0.05;
dataw53=data[data$VAPUTI==1,];
dataw53$status=3;
dataw53$time=dataw53$UTItime;
dataw53$aa=dataw53$aa+0.053;
dataw52=data[dataw$VAPUTI==1,];

```

```

dataw52=dataw52[dataw52$VAPtime;=dataw52$UTItime,];
dataw52$status=2;
dataw52$time=dataw52$VAPtime;
dataw52$aa=dataw52$aa+0.0502;

# VAP+BSI+UTI
dataw6 =data[dataw$VAPBSIUTI==1,];
dataw6$status=1;
dataw6$time=0;
dataw6$aa=dataw6$aa+0.06;
dataw63=data[dataw$VAPBSIUTI==1,];
dataw63$minBSIUTItime=apply(cbind(dataw63$BSItime,dataw63$UTItime),1,min)
dataw63$status=3;
dataw63$time=dataw63$minBSIUTItime;
dataw63$aa=dataw63$aa+0.063;
dataw63=dataw63[,-30];
dataw62=data[dataw$VAPBSIUTI==1,];
dataw62$minBSIUTItime=apply(cbind(dataw62$BSItime,dataw62$UTItime),1,min)
dataw62=dataw62[dataw62$VAPtime;=dataw62$minBSIUTItime,];
dataw62$status=2;
dataw62$time=dataw62$VAPtime;
dataw62$aa=dataw62$aa+0.0602;
dataw62=dataw62[,-30];

# BSI+UTI
dataw7 =data[dataw$BSIUTI==1,];
dataw7$status=1;
dataw7$time=0;
dataw7$aa=dataw7$aa+0.07;
dataw73=data[dataw$BSIUTI==1,];
dataw73$minBSIUTItime=apply(cbind(dataw73$BSItime,dataw73$UTItime),1,min);
dataw73$status=3;
dataw73$time=dataw73$minBSIUTItime;

```

```
dataw73$aa=dataw73$aa+0.072;
dataw73=dataw73[,-30]
```

```
vapw=rbind(dataw0,dataw09,dataw1,dataw12,dataw19,dataw2,dataw23,dataw3,dataw33,
dataw4,dataw43,dataw42,dataw5,dataw53,dataw52, dataw6,dataw63,dataw62,dataw7,dataw73)
ooo=order(vapw$aa)
vapw=vapw[ooo,]
vapw$sex=vapw$sex-1
```

```
# ANALYSIS
```

```
library(msm)
```

```
statetable.msm(status,X,data=vapw)
```

```
qqq=rbind(c(0,0.32,0.2,0.26,0.22),c(0,0,0.47,0.3,0.23),c(0,0,0,0,0),c(0,0,0,0,0),c(0,0,0,0,0));
```

```
qqq
```

```
crudeinits.msm(status time,X,data=vapw,qmatrix=qqq)
```

```
vapw.msm=msm(status time,subject=X,data=vapw,qmatrix=qqq,gen.inits=T,death=c(3,4,5),obst,
```

```
vapw.msm
```

```
summary(vapw.msm)
```

```
vapw.msm6=msm(status time,subject=X,data=vapw,qmatrix=qqq,gen.inits=T,death=c(3,4,5),obst,
covariates= infection);vapw.msm6
```

```
vapw.msm10=msm(status time,subject=X,data=vapw,qmatrix=qqq,gen.inits=T,death=c(3,4,5),obst,
covariates= Apache2);vapw.msm10;
```

```
vapw.msm13=msm(status time,subject=X,data=vapw,qmatrix=qqq,gen.inits=T,death=c(3,4,5),obst,
covariates= durven);vapw.msm13;
```

```
vapw.msm30=msm(status time,subject=X,data=vapw,qmatrix=qqq,gen.inits=T,death=c(3,4,5),obst,
covariates= infection+Apache2);vapw.msm30;
```

```
vapw.msm31=msm(status time,subject=X,data=vapw,qmatrix=qqq,gen.inits=T,death=c(3,4,5),obst,
covariates= infection+Apache2+durven);vapw.msm31; lrtest.msm(vapw.msm,vapw.msm6)
```

```
lrtest.msm(vapw.msm,vapw.msm10)
```

```
lrtest.msm(vapw.msm6,vapw.msm30,vapw.msm31)
```

```
lrtest.msm(vapw.msm10,vapw.msm30,vapw.msm31)
```

B.2 Cox regression

```

datax=data
datax$X=data$aa
datax$none =as.numeric(data$VAP==0 & data$BSI==0 & data$UTI==0)
datax$VAPalone =as.numeric(data$VAP==1 & data$BSI==0 & data$UTI==0)
datax$BSIalone =as.numeric(data$VAP==0 & data$BSI==1 & data$UTI==0)
datax$UTIALone =as.numeric(data$VAP==0 & data$BSI==0 & data$UTI==1)
datax$VAPBSI =as.numeric(data$VAP==1 & data$BSI==1 & data$UTI==0)
datax$VAPUTI =as.numeric(data$VAP==1 & data$BSI==0 & data$UTI==1)
datax$VAPBSIUTI=as.numeric(data$VAP==1 & data$BSI==1 & data$UTI==1)
datax$BSIUTI =as.numeric(data$VAP==0 & data$BSI==1 & data$UTI==1)
datax$minBSIUTItime=apply(cbind(datax$BSItime,datax$UTItime),1,min)

# VAP alone
datax1=data[datax$VAPalone==1,];
datax1$y=1;
datax1$duration=datax1$length-datax1$VAPtime;
# VAP+BSI
datax4=data[datax$VAPBSI==1,];
datax4=datax4[datax4$VAPtime<datax4$BSItime,];
datax4$duration=datax4$length-datax4$VAPtime;
datax4$y=2;
datax41=data[datax$VAPBSI==1,];
datax41=datax41[datax41$VAPtime<datax41$BSItime,];
datax41$duration=datax41$length-datax41$BSItime;
datax41$y=3;
# VAP+UTI
datax5=data[datax$VAPUTI==1,];
datax5=datax5[datax5$VAPtime<datax5$UTItime,];
datax5$duration=datax5$length-datax5$VAPtime;
datax5$y=2;
datax51=data[datax$VAPUTI==1,];

```



```

datax51=datax51[datax51$VAPtime<datax51$UTItime,];
datax51$duration=datax51$length-datax51$UTItime;
datax51$y=3;
# VAP+BSI+UTI
datax6=data[datax$VAPBSIUTI==1,];
datax6$minBSIUTItime=apply(cbind(datax6$BSItime,datax6$UTItime),1,min)
datax6=datax6[datax6$VAPtime<=datax6$minBSIUTItime,];
datax6$duration=datax6$length-datax6$VAPtime;
datax6$y=2;
datax6=datax6[,-30];
datax61=data[datax$VAPBSIUTI==1,];
datax61$minBSIUTItime=apply(cbind(datax61$BSItime,datax61$UTItime),1,min);
datax61=datax61[datax61$VAPtime<datax61$minBSIUTItime,];
datax61$duration=datax61$length-datax61$minBSIUTItime;
datax61$y=3;
datax61=datax61[,-30];
cx=rbind(datax1,datax4,datax41,datax5,datax51,datax6,datax61)
ooo=order(cx$aa)
cx=cx[ooo,]
cx$sex=cx$sex-1
cox<-coxph(Surv(cx$length, cx$death) factor(cx$y))

cif.ab <- etmCIF(Surv(duration, death) ~ y, cx, etype = death)
plot(cif.ab)

```

B.3 ChangeLOS-etm

```

# changeLOS-VAP
data.los<-data.frame(data$X,data$VAPalone*data$VAPtime,(data$none)*(1-data$death)
data$length,data$none*data$death*data$length,data$VAPalone*(1-data$death)*data$length,
data$VAPalone*data$death*data$length)

data.los[,7]<-0

```

```

data.los[,7]<—data$VAPBSI*data$minVAPBSItime+data$BSIalone*data$BSItime+
data$UTIalone*data$UTItime+data$VAPUTI*data$minVAPUTItime+data$VAPBSIUTI*
data$minVAPBSIUTItime+data$BSIUTI*data$minBSIUTItime
data.los<—data.frame(data.los,(1-data$VAPalone)*Inf+(1-data$none)*Inf+
data$VAPBSI*data$minVAPBSItime+data$BSIalone*data$BSItime+data$UTIalone*
data$UTItime+data$VAPUTI*data$minVAPUTItime+data$VAPBSIUTI*
data$minVAPBSIUTItime)
for (i in 1:294)
  for (j in 2:7)
    if (data.los[i,j]==0)
      data.los[i,j]<—Inf
names(data.los)[1]<—“j.id”
names(data.los)[2]<—“j.12”
names(data.los)[3]<—“j.13”
names(data.los)[4]<—“j.14”
names(data.los)[5]<—“j.23”
names(data.los)[6]<—“j.24”
names(data.los)[7]<—“cens”

tra<—rbind(c(FALSE, TRUE,TRUE,TRUE), c(FALSE,FALSE,TRUE,TRUE),
c(FALSE,FALSE, FALSE,FALSE), c(FALSE,FALSE, FALSE, FALSE))
my.model <— msmodel(c(“0”, “1”, “2”, “3”),tra,cens.name=“cens”)
los.data<—data.los
my.observ <— prepare.los.data(x=los.data)
losss<—clos(my.model, my.observ)
summary(losss)

loss <— etm(my.observ[,-5],c(“0”, “1”, “2”, “3”),tra,cens.name=“cens”,0);loss
los <—clos(loss);los
los2 <—clos(loss, aw=TRUE);los2
summary(los)
summary(loss)
plot(loss)

```

```
plot(los)
```

```
p <— prepare.progdismodel(data=los.data)
pdm <— progdismodel(p$model, p$observ)
plot(pdm)
```

```
plot(loss, tr.choice = 0 1", conf.int = TRUE, lwd = 2, legend = FALSE, ylim
= c(0, 0.3), xlim = c(0, 100), xlab = "Days", ci.fun = "cloglog")
```

```
# ChangeLOS-BSI
data.los2<—data.frame(data$X,data$BSIalone*data$BSItime,
(data$none)*(1-data$death)*data$length,data$none*data$death*data$length,
data$BSIalone*(1-data$death)*data$length,data$BSIalone*data$death*data$length)
data.los2[,7]<—data$VAPBSI*data$minVAPBSItime+data$VAPalone*data$VAPtime+
data$UTIalone*data$UTItime+data$BSIUTI*data$minBSIUTItime+data$VAPBSIUTI*
data$minVAPBSIUTItime+data$VAPUTI*data$minVAPUTItime;
data.los2<—data.frame(data.los,(1-data$BSIalone)*Inf+(1-data$none)*Inf+data$VAPBSI*
data$minVAPBSItime+data$VAPalone*data$VAPtime+data$UTIalone*data$UTItime+
data$BSIUTI*data$minBSIUTItime+data$VAPBSIUTI*data$minVAPBSIUTItime)
```

B'.4 Σχήματα

Σχήμα 4.5

```
my.observ2<—data.frame(my.observ,0)
for(i in 1:344)
{
  my.observ2[1:50,6]<—1
  my.observ2[191:240,6]<—1
}
my.data<—data.frame(my.observ2[,5])
names(my.data)[5]<—"VAP"
```

```

my.nelaal.nov <— mvna(my.data[my.data$VAP == 0, ], c("0", "1", "2", "3"), tra,
" cens")
my.nelaal.v <— mvna(my.data[my.data$VAP == 1, ], c("0", "1", "2", "3"), tra,
" cens")

xyplot(my.nelaal.nov, data = NULL, xlab = "Days", ylab = "Cumulative Haz-
ard", tr.choice = c("0 2", "0 3"), conf.int = TRUE)
xyplot(my.nelaal.v, data = NULL, xlab = "Days", ylab = "Cumulative Hazard",
tr.choice = c("1 2", "1 3"), conf.int = TRUE)

```

Σχήμα 4.6

```

my.data2 <— my.data
for(i in 1:344)
{if (my.data2[i,3]==3)

    my.data2[i,3] <— 2
}
names(my.data2)[5] <— "VAP"
tra2 <— rbind(c(FALSE, TRUE, TRUE), c(FALSE, FALSE, TRUE), c(FALSE, FALSE,
FALSE))
my.nelaal.nov2 <— mvna(my.data2[my.data2$VAP == 0, ], c("0", "1",
"2"), tra2, " cens")
my.nelaal.v2 <— mvna(my.data2[my.data2$VAP == 1, ], c("0", "1", "2"), tra2,
" cens")

xyplot(my.nelaal.nov2, data = NULL, xlab = "Days", ylab = "Cumulative Haz-
ard", tr.choice = c("0 2"), conf.int = TRUE)
xyplot(my.nelaal.v2, data = NULL, xlab = "Days", ylab = "Cumulative Hazard",
tr.choice = c("1 2"), conf.int = TRUE)

```

Σχήμα 4.8

```
my.etm.nov <— etm(my.data[my.data$VAP==0, ], c("0", "1", "2", "3"), tra, "cens",  
s = 0)  
my.etm.v <— etm(my.data[my.data$VAP==1, ], c("0", "1", "2", "3"), tra, "cens",  
s = 0)  
  
plot(my.etm.v, xlab="Days", tr.choice = '0 2', col = 1, lwd = 2, conf.int = TRUE,  
ci.fun = "cloglog", legend = FALSE, ylab="Probability")  
lines(my.etm.nov, tr.choice = '0 2', col = "gray", lwd = 2, conf.int = TRUE, ci.fun  
= "cloglog")
```

Βιβλιογραφία

- [1] A.Allignol, M.Schumacher, J..Beyersmann (2010a) *Empirical transition matrix of multistate models:the etm package*. Journal of Statistical Software, 38(4), 1-15. <http://www.jstatsoft.org/v38/i04/>.
- [2] A.Allignol, M.Schumacher, J..Beyersmann *Estimating summary functionals in multistate models with an application to hospital infection data* Comput Stat(2011) 26:181-197
- [3] Allignol A, Beyersmann J, Schumacher M (2008). *mvna:an R package for cumulative hazards and transition probabilities of multistate models* R News, 8, 48-50.
- [4] PK Andersen, SZ Abildstrom, S. Rosthoj *Competing risks as a multi-state model*. Statistical Methods in Medical Research 2002;11:203-15
- [5] PK Andersen, N. Keiding *Multi-state models for event history analysis*. Statistical Methods Medical Research 2002;11:91-115.
- [6] J.Beyersmann, M.Schumacher, A.Allignol *Competing Risks and Multistate Models with R* Springer;2012
- [7] J.Beyersmann, P.Gastmeier, H.Grunmann, S.Barwolff, C.Gefers, M.Behnke, H.Ruden, M.Schumacher, *Use of Multistate Models to Assess Prolongation of Intensive Care Unit Due to Nosocomial Infection* Infection Control and Hospital Epidemiology May 2006, Vol.27, No. 5
- [8] D.Commenges *Multi-state Models in Epidemiology* Lifetime Data Analysis 5,315-327(1999)

-
- [9] Cox,D R.; Miller,HD. *The theory of stochastic processes*. Chapman & Hall; 1965.
- [10] Cox, D.R. & D.Oakes. 1984. *Analysis of Survival Data*. London:Chapman and Hall
- [11] C.J.Crnich *Estimating Excess Length of Stay Due to Central Line-Associated Bloodstream Infection:Separating the Wheat from the Chaff* Infection Control and Hospital Epidemiology, November 20120, Vol.31, No.11
- [12] M J. Crowder *Classical competing risks*. Chapman & Hall/CRC:Boca raton, 2001.
- [13] J.L.Folks, R.S. Chhikara *The Inverse Gaussian Distribution and its Statistical Application-A Review* Journal of the Royal Statistical Society. Series B (Methological) Vol. 40, No. 3(1978), pp. 263-289
- [14] J.Fox *Cox Proportional-Hazards Regression for Survival Data* Appendix to An R and S-Plus Companion to Applied Regression
- [15] M.K.Goel, P.Khanna, J.Kishore *Understanding survival analysis:Kaplan-Meier* National Library of Medicine Institutes of Health 2010 Oct-Dec;1(4):274-278
- [16] P. Hougaard *Analysis of multivariate survival data*. Springer;2000.
- [17] Hougaard P. *Multi-state models:a review*.Lifetime Data Analysis 1999;5:239-64.
- [18] Jackson C (2010). *Multi-State Modelling with R: The msm Package*. R package version 1.0, <http://CRAN.R-project.org/package=msm>
- [19] Jackson C (2011). *Multi-State Models for Panel Data: The msm Package for R*. Journal of Statistical Software, 38(8), 1-29. <http://www.jstatsoft.org/v38/i08/>.
- [20] N.Keiding, P.K.Anderen *Nonparametric Estimation of Transition Intensities and Transition Probabilities: A Case Study of a Two-State Markov Process*

- Journal of the Royal Statistical Society. Series C (Applied Statistics), Vol. 38, No. 2 (1989), pp.319-329
- [21] M.-L. Lambert, C. Svetens, A.Savey, M.Palomar, M.Hiersmayr, I.Morales, A.Agodi, U. Frank, K.Martens, M. Schumacher, M. Wolkewitz *Clinical outcomes of health-care-associated infections and antimicrobial resistance in patients admitted to European intensive-care units: a cohort study*
- [22] B.H.Lindqvist *Competing Risks*
<http://www.math.ntnu.no/bo/forskning/CompRiskExtended2.pdf>
- [23] L. Meira-Machado, J. de Una-Alvarez, C. Cadarso-Sarez, P. k Andersen *Multi-state models for the analysis of time-to-event data* Statistical Methods Medical Research 2009 April;18(2):195-222.
- [24] M. Nguile-Makao, J.-R. Zahar, A.Francais, A.Tabah, M.Garrouste-Orgeas, B. Allaouchiche, D. Goldgran-Toledano, E.Azoulay, C.Adrie, S.Jamali, C.Clec'h, B.Souweine, J.-F. Timisit *Attributable mortality of ventilator-associated pneumonia: respective impact of main characteristics at ICU admission and VAP onset using conditional logistic regression and multi-state models* Intensive Care Med (2010) 36:781-789
- [25] H.Putter, M.Fiocco, R.B.Geskus *Tutorial in biostatistics: Competing risks and multi-state models* Statist. Med. 2007;26:2389-2430
- [26] R Development Core Team R: A Language and Environment for Statistical Computing. R Foundation for statistical computing, Vienna, Austria, <http://www.R-project.org>
- [27] G.Rodriguez *Non-Parametric Estimation in Survival Models* Spring;2001, revised 2005
- [28] M.Schumacher, M.Wangler, M.Wolkewitz, J.Beyersmann *Attributable Mortality due to Nosocomial Infections-A Simple and Useful Application of Multistate Models* Methods Inf Med 5/2007;46:595-600

-
- [29] G. Shulgen, M. Schumacher *Estimation of Prolongation of Hospital Stay Attributable to Nosocomial Infections: New Approaches Based on Multistate Models* Lifetime Data Analysis, 2, 219-240 (1996)
- [30] F. Siannis *Introduction to Survival Analysis* <http://users.uoa.gr/~fsiannis>
- [31] Therneau, T.M.; Grambsch, P.M. *Modeling survival data: extending the Cox model*. Springer-Verlag; 2000.
- [32] Therneau T, Lumley T (2010). *survival: Survival Analysis Including Penalised Likelihood*. R package version 2.36-1, <http://CRAN.R-project.org/package=survival>.
- [33] J.-F. Timsit, J.-R. Zahar, S. Chevret *Attributable mortality of ventilator-associated pneumonia* 1070-5295, 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins
- [34] M. Wangler, J. Beyersmann, M. Schumacher *Changelos: an R-package for change in length of hospital stay based on the Aalen-Johansen estimator*. R News 2006;6:31-5