

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

Θεοδώρας Δημητρακοπούλου
Μαθηματικού

Μοντέλα Ανάλυσης Επιβίωσης

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΙΩΑΝΝΙΝΑ 2006



ΒΙΒΛΙΟΘΗΚΗ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΙΩΑΝΝΙΝΩΝ



026000348920





ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΕΑΕΚ

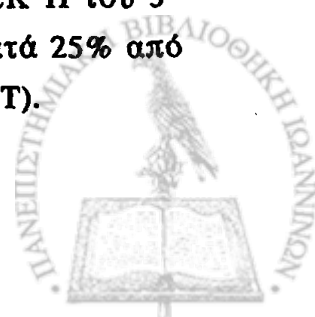


ΕΥΡΩΠΑΪΚΗ ΕΝΩΣΗ
ΣΥΓΧΡΗΜΑΤΟΔΟΤΗΣΗ
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



Η ΠΑΙΔΕΙΑ ΣΤΗΝ ΚΟΡΥΦΗ
Επιχειρησιακό Πρόγραμμα
Εκπαίδευσης και Αρχικής
Επαγγελματικής Κατάρτισης

Η εργασία αυτή συγχρηματοδοτήθηκε από την Ευρωπαϊκή Ένωση, στα πλαίσια του προγράμματος με τίτλο «ΗΡΑΚΛΕΙΤΟΣ», του προγράμματος ΕΠΕΑΕΚ II του 3^{ου} Κοινοτικού Πλαισίου Στήριξης του ΥΠ.Ε.Π.Θ., με χρηματοδότηση κατά 25% από εθνικούς πόρους και κατά 75% από το Ευρωπαϊκό Κοινωνικό Ταμείο (ΕΚΤ).



Θεοδώρας Δημητρακοπούλου
Μαθηματικού

Μοντέλα Ανάλυσης Επιβίωσης

Διδακτορική Διατριβή
υποβληθείσα στο Τμήμα Μαθηματικών
του Πανεπιστημίου Ιωαννίνων

Η Εξεταστική Επιτροπή που ορίστηκε από τη Γενική Συνέλευση Ειδικής Σύνθεσης (Γ.Σ.Ε.Σ.) του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, στην υπ' αριθμ. 437(A)/5-4-2006 συνεδρία, εγκρίνει την παρούσα διατριβή ως πληρούσα τις προϋποθέσεις για την απονομή του τίτλου του

Διδάκτορα των Μαθηματικών

Η Εξεταστική Επιτροπή

Σωτήριος Λουκάς, (Επιβλέπων Καθηγητής)

Καθηγητής Τμήματος Μαθηματικών Πανεπιστημίου Ιωαννίνων

Κωνσταντίνος Αδαμίδης, (Μέλος Τριμελούς Συμβουλευτικής Επιτροπής)

Επίκουρος Καθηγητής Τμήματος Οργάνωσης και Διαχείρισης Αγροτικών Εκμεταλλεύσεων
Πανεπιστημίου Ιωαννίνων

Αγλαΐα Καλαματιανού, (Μέλος Τριμελούς Συμβουλευτικής Επιτροπής)

Καθηγήτρια Παντείου Πανεπιστημίου

Σταύρος Κουρούκλης,

Καθηγητής Τμήματος Μαθηματικών Πανεπιστημίου Πατρών

Χρήστος Λάγκαρης,

Αναπληρωτής Καθηγητής Τμήματος Μαθηματικών Πανεπιστημίου Ιωαννίνων

Σωτήριος Παπαχρήστος,

Αναπληρωτής Καθηγητής Τμήματος Οργάνωσης και Διαχείρισης Αγροτικών
Εκμεταλλεύσεων Πανεπιστημίου Ιωαννίνων

Κοσμάς Φερεντίνος,

Καθηγητής Τμήματος Μαθηματικών Πανεπιστημίου Ιωαννίνων



Στους γονείς μου Δημήτρη και Αντωνία
και τις αδερφές μου Τζίνα και Σταυρούλα



Η διατριβή αυτή, εκπονήθηκε στον Τομέα Πιθανοτήτων, Στατιστικής και Επιχειρησιακών Ερευνών του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων. Στο σημείο αυτό, θα ήθελα να εκφράσω τις ευχαριστίες μου σε αυτούς που συνέβαλαν στην πραγματοποίησή της.

Στον κ. Σωτήριο Λουκά, Καθηγητή του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, για την επίβλεψη των ερευνητικών μου δραστηριοτήτων, την εξασφάλιση οικονομικών πόρων μέσω του προγράμματος 'Ηράκλειτος' για την απρόσκοπτη ενασχόλησή μου με τη διδακτορική μου διατριβή και την παντός είδους βοήθεια και συμπαράστασή του.

Στον κ. Κωνσταντίνο Αδαμίδα, Επίκουρο Καθηγητή του Τμήματος Οργάνωσης και Διαχείρισης Αγροτικών Εκμεταλλεύσεων του Πανεπιστημίου Ιωαννίνων, για την επιστημονική του καθοδήγηση, την αδιάκοπη ενθάρρυνση και τις χρήσιμες παρεμβάσεις του στη διαμόρφωση της διατριβής και στη συγγραφή της.

Στην κ. Αγλαΐα Καλαματιανού, Καθηγήτρια του Παντείου Πανεπιστημίου για την ενθάρρυνση που μου προσέφερε και τις χρήσιμες υποδείξεις της.

Τους παραπάνω ευχαριστώ ιδιαίτερα που δέχτηκαν να αποτελέσουν την Τριμελή Συμβουλευτική Επιτροπή. Επίσης, οφείλω να ευχαριστήσω τους κ. κ. Σταύρο Κουρούκλη, Καθηγητή του Τμήματος Μαθηματικών του Πανεπιστημίου Πατρών, Σωτήριο Παπαχρήστο, Αναπληρωτή Καθηγητή του Τμήματος Οργάνωσης και Διαχείρισης Αγροτικών Εκμεταλλεύσεων του Πανεπιστημίου Ιωαννίνων, Χρήστο Λάγκαρη, Αναπληρωτή Καθηγητή του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων και Κοσμά Φερεντίνο, Καθηγητή του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, για τη συμμετοχή τους στην Επταμελή Εξεταστική Επιτροπή και τις χρήσιμες υποδείξεις τους.

Περισσότερο από όλους, θα ήθελα να ευχαριστήσω τους γονείς μου Δημήτρη και Αντωνία και τις αδερφές μου Τζίνα και Σταυρούλα, που στάθηκαν αποφασιστικά δίπλα μου σε όλη τη διάρκεια αυτής μου της προσπάθειας. Τέλος, θα ήθελα να ευχαριστήσω τους θείους μου Μαργαρίτα και Λεάνδρο και τους φίλους μου Γιώργο, Κωσταντίνα, Έλενα και Βασίλη, για τη συμπαράστασή τους.

Θεοδώρα Δημητρακοπούλου

Ιωάννινα, Μάιος 2006



Περιεχόμενα

1. Εισαγωγή	1
1.1 Κατανομή του χρόνου ζωής, 3	
1.1.1 Η συνάρτηση κινδύνου και η σημασία της, 7	
1.1.2 Βασικές κατανομές ανάλυσης επιβίωσης, 9	
1.2 Ειδικά μοντέλα ανάλυσης επιβίωσης, 18	
1.2.1 Μοντέλα ανταγωνιστικών κινδύνων, 18	
1.2.2 Μοντέλα Ευπάθειας, 19	
1.3 Διδιάστατες κατανομές χρόνων ζωής, 20	
2. Η Γενικευμένη Εκθετική-Γεωμετρική Κατανομή	25
2.1 Η κατανομή, 26	
2.2 Φυσική ερμηνεία, 27	
2.3 Σχέσεις με άλλες κατανομές, 30	
2.4 Ροπές, 31	
2.5 Συναρτήσεις επιβίωσης και κινδύνου, 32	
2.6 Μέσος υπολειπόμενος χρόνος ζωής, 33	
2.7 Συμπερασματολογία, 34	
2.7.1 Εκτιμητική με τη μέθοδο της μέγιστης πιθανοφάνειας, 34	
2.7.2 Η μέθοδος EM, 35	
2.7.3 Ασυμπτωτική κατανομή των εκτιμητών μέγιστης πιθανοφάνειας, 38	
2.8 Χαρακτηρισμοί, 47	
2.9 Εφαρμογές, 49	
3. Ένα Μοντέλο Ανάλυσης Επιβίωσης με Μονότονες, Κυπελλοειδούς Μορφής και Μονοκόρυφες Συναρτήσεις Κινδύνου	53
3.1 Η κατανομή, 54	
3.2 Σχέσεις με άλλες κατανομές, 55	



3.3	Πιθανότητες και ροπές,	56
3.4	Συναρτήσεις επιβίωσης και κινδύνου,	57
3.5	Φυσική ερμηνεία,	58
3.6	Συμπερασματολογία,	60
3.6.1	Εκτιμητική με τη μέθοδο της μέγιστης πιθανοφάνειας,	60
3.6.2	Ασυμπτωτική κατανομή των εκτιμητών μέγιστης πιθανοφάνειας,	61
3.7	Εφαρμογές,	62
4.	Διδιάστατες Γενικευμένες Εκθετικές - Γεωμετρικές Κατανομές	65
4.1	Διδιάστατη <i>EEG</i> -1,	65
4.1.1	Περιθώριες και δεσμευμένες κατανομές,	66
4.1.2	Οι συναρτήσεις επιβίωσης και κινδύνου,	67
4.1.3	Ο συντελεστής συσχέτισης,	69
4.2	Διδιάστατη <i>EEG</i> -2,	70
4.2.1	Περιθώριες και δεσμευμένες κατανομές,	72
4.2.2	Οι συναρτήσεις επιβίωσης και κινδύνου,	73
4.2.3	Ο συντελεστής συσχέτισης,	75
4.3	Διδιάστατη <i>EEG</i> -3,	76
4.3.1	Περιθώριες και δεσμευμένες κατανομές,	78
4.3.2	Οι συναρτήσεις επιβίωσης και κινδύνου,	80
4.3.3	Ο συντελεστής συσχέτισης,	83
4.4	Διδιάστατη <i>EEG</i> - 4,	85
4.4.1	Περιθώριες και δεσμευμένες κατανομές,	86
4.4.2	Οι συναρτήσεις επιβίωσης και κινδύνου,	88
4.4.3	Ο συντελεστής συσχέτισης,	93
4.5	Φυσική ερμηνεία,	96
5.	Ιδιότητες Κατανομών Ανάλυσης Επιβίωσης - Μέθοδοι Παραγωγής και Γενίκευσής τους	97
5.1	Το Γάμα μοντέλο ευπάθειας,	98
5.2	Χαρακτηρισμοί,	102
5.3	Φράγματα της συνάρτησης επιβίωσης,	105



Παράρτημα	111
Abstract of the Dissertation	113
Βιβλιογραφία	117



Κεφάλαιο 1

Εισαγωγή

Στον όρο 'ανάλυση επιβίωσης', συνοψίζεται το σύνολο των στατιστικών μεθοδολογιών που έχουν ως στόχο την περιγραφή και τη μοντελοποίηση της διάρκειας ενός φαινομένου ή συμβάντος στο χρόνο. Ο χρόνος διάρκειας του φαινομένου αναφέρεται συνήθως ως 'επιβίωση' ή 'διάρκεια ζωής' ή 'χρόνος ζωής' ενώ η οριστική παύση του φαινομένου αναφέρεται συχνά ως 'αποτυχία' ή 'θάνατος'. Οι χρόνοι ανάνηψης, υποτροπής ή θανάτου από κάποια ασθένεια, ο χρόνος μέχρι την εκδήλωση εργατικού ατυχήματος ή δυσλειτουργικού σφάλματος σε κάποιο μηχανικό εξάρτημα, η βιωσιμότητα μιας επιχείρησης, η διάρκεια μιας απεργίας ή η περίοδος ανεργίας και το χρονικό διάστημα μέχρι την ολοκλήρωση μιας χημικής αντίδρασης, είναι μερικά μόνο παραδείγματα, που η μελέτη τους εμπίπτει στο γνωστικό αντικείμενο της ανάλυσης επιβίωσης.

Γενικά, η εκτίμηση, η πρόβλεψη, ή η βελτιστοποίηση πιθανοτήτων της 'επιβίωσης' ή της αναμενόμενης διάρκειας 'ζωής', αποτελεί το ζητούμενο σε πολλές εκφάνσεις της ανθρώπινης ζωής και δραστηριότητας. Κατά συνέπεια, οι μεθοδολογίες της ανάλυσης επιβίωσης αποτελούν χρήσιμο εργαλείο για ερευνητές από διάφορους χώρους, όπως της οικονομίας, της ιατρικής και των άλλων βιολογικών επιστημών.

Η διατριβή αυτή, πραγματεύεται θέματα που αφορούν σε συνεχείς κατανομές ανάλυσης επιβίωσης. Στο πρώτο κεφάλαιο, παρουσιάζονται οι εισαγωγικές έννοιες συνεχών κατανομών χρόνων ζωής, με ιδιαίτερη έμφαση στην έννοια της συνάρτησης κινδύνου και στη σημασία της, στην περιγραφή της ικανότητας επιβίωσης ενός πληθυσμού. Επίσης, παρουσιάζονται οι πιο δημοφιλείς κατανομές χρόνων ζωής με σταθερές, μονότονες,



μονοκόρυφες ή κυπελλοειδούς μορφής συναρτήσεις κινδύνου και γίνεται μια συνοπτική περιγραφή των εξειδικευμένων μοντέλων ανάλυσης της επιβίωσης (ανταγωνιστικών κινδύνων και ευπάθειας), που χρησιμοποιούνται στα επόμενα κεφάλαια. Στη συνέχεια, παρουσιάζονται οι εισαγωγικές έννοιες ανάλυσης επιβίωσης για την περίπτωση διδιάστατου πληθυσμού και επισημαίνονται οι όποιες διαφορές με τη μονοδιάστατη περίπτωση. Τέλος, γίνεται αναφορά στις πιο δημοφιλείς διδιάστατες κατανομές ανάλυσης επιβίωσης.

Στο **δεύτερο κεφάλαιο**, παρουσιάζεται μια κατανομή ανάλυσης επιβίωσης δύο παραμέτρων, με αύξουσα και φθίνουσα συνάρτηση κινδύνου. Μελετώνται διάφορες ιδιότητές της, όπως οι τρόποι παραγωγής της, οι σχέσεις της με άλλες κατανομές, η μονοτονία της συνάρτησης πυκνότητας πιθανότητάς της, οι ροπές της και οι αντίστοιχες συναρτήσεις επιβίωσης, κινδύνου και μέσου υπολειπόμενου χρόνου ζωής. Το πρόβλημα της εκτίμησης των παραμέτρων, αντιμετωπίζεται με τη μέθοδο της μέγιστης πιθανοφάνειας και τους πίνακες παρατηρούμενης και αναμενόμενης πληροφορίας ενώ για τον υπολογισμό των εκτιμητών, προτείνεται ένας EM αλγόριθμος (Dempster et al., 1977). Το κεφάλαιο κλείνει με χαρακτηρισμούς της κατανομής, όπως επίσης και με συγκριτικά - σε σχέση με τις κατανομές Weibull και Γάμα - αποτελέσματα από την προσαρμογή της σε πραγματικά δεδομένα. Η κατανομή και τα παραγόμενα από τη μελέτη της αποτελέσματα, παρουσιάζονται στην εργασία των Adamidis et al. (2005).

Στο **τρίτο κεφάλαιο**, παρουσιάζεται ένα μοντέλο ανάλυσης επιβίωσης τριών παραμέτρων, με μονότονες, κυπελλοειδούς μορφής και μονοκόρυφες συναρτήσεις κινδύνου. Μελετώνται διάφορες μαθηματικές και στατιστικές ιδιότητές του ενώ επίσης, παρουσιάζεται η φυσική διεργασία παραγωγής του, μέσα από τη θεωρία ανταγωνιστικών κινδύνων. Το πρόβλημα της εκτίμησης των παραμέτρων, αντιμετωπίζεται με τη μέθοδο της μέγιστης πιθανοφάνειας. Τέλος, παρουσιάζονται εφαρμογές του μοντέλου σε πραγματικά δεδομένα. Τα αποτελέσματα αυτού του κεφαλαίου, παρουσιάζονται στην εργασία των Dimitrakoroulou et al. (2006 b).

Στο **τέταρτο κεφάλαιο**, παρουσιάζονται διδιάστατες μορφές της κατανομής του δεύτερου κεφαλαίου και περιγράφονται οι φυσικές διεργασίες που τις παράγουν. Για κάθε μορφή, δίνονται οι περιθώριες και οι δεσμευμένες κατανομές και υπολογίζονται οι συναρτήσεις επιβίωσης και κινδύνου. Η συσχέτιση των τυχαίων μεταβλητών που περιγράφουν το διδιάστατο πληθυσμό, εξετάζεται μέσω του συντελεστή συσχέτισης του Pearson.



Οι κατανομές που παρουσιάστηκαν στα κεφάλαια 2 έως 4, στοχεύουν στη μοντελοποίηση χρόνων ζωής κάτω από την υπόθεση ομοιογενούς πληθυσμού. Όταν η υπόθεση αυτή δεν ευσταθεί, ένας από τους τρόπους περιγραφής της επίδρασης της ανομοιογένειας στην επιβίωση, είναι μέσω της έννοιας και των μοντέλων ευπάθειας (Vaupel et al., 1979). Στο **πέμπτο κεφάλαιο**, αποδεικνύεται ότι όλες οι κατανομές ανάλυσης επιβίωσης μπορούν να προκύψουν από το Γάμα μοντέλο ευπάθειας και παρουσιάζονται παραδείγματα χρήσης του στην παραγωγή και γενίκευσή τους (με προσθήκη μιας επιπλέον παραμέτρου υπό την έννοια που εξετάζεται στους Marsall και Olkin, 1977). Δίνονται ικανές και αναγκαίες συνθήκες χαρακτηρισμού του συναρτήσεως της μονοτονίας της συνάρτησης κινδύνου και παράγονται φράγματα για τη συνάρτηση επιβίωσης. Έτσι, μελετώνται με ενιαίο τρόπο οι ιδιότητες των κατανομών ανάλυσης επιβίωσης και εξάγονται συμπεράσματα για τις συναρτήσεις κινδύνου και επιβίωσης του πληθυσμού, σε σχέση με τη συμπεριφορά της ‘μέσης συνάρτησης κινδύνου’ των πληθυσμιακών μονάδων και του βαθμού ανομοιογένειας στην ικανότητα επιβίωσής τους. Τα αποτελέσματα του πέμπτου κεφαλαίου, παρουσιάζονται στην εργασία των Dimitrakopoulou et al. (2006 a).

1.1 Κατανομή του χρόνου ζωής

Στην παράγραφο αυτή, παρουσιάζονται ισοδύναμοι τρόποι ορισμού της κατανομής της συνεχούς, μη αρνητικής τυχαίας μεταβλητής T η οποία παριστά χρόνους ζωής, κάτω από την υπόθεση ομοιογενούς πληθυσμού.

Ως **συνάρτηση επιβίωσης** (survivor function ή reliability function), ορίζεται η πιθανότητα του ενδεχομένου η τυχαία μεταβλητή T να πάρει τιμές μεγαλύτερες του t :

$$S(t) = P(T > t), \quad t \geq 0.$$

Κατά συνέπεια, εκφράζει την πιθανότητα επιβίωσης μέχρι τη χρονική στιγμή t ή διαφορετικά, το ποσοστό των πειραματικών μονάδων που αναμένεται να έχουν επιβιώσει στο χρόνο t . Η συνάρτηση επιβίωσης είναι φθίνουσα και συνεχής, με $S(0) = 1$ και $\lim_{t \rightarrow \infty} S(t) = 0$.

Η σημασία της στην ανάλυση επιβίωσης είναι μεγάλη αφού μεταξύ άλλων, αφενός αποτελεί μέτρο σύγκρισης της βιωσιμότητας διαφορετικών πληθυσμών, αφετέρου διευκολύνει το στατιστικό χειρισμό δεξιά λογοκριμένων δεδομένων (right censored data).



Η συνάρτηση πυκνότητας πιθανότητας (σ.π.π.) της κατανομής του χρόνου ζωής, ορίζεται από τη σχέση:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{\Delta t} = -S'(t), \quad t \geq 0. \quad (1.1)$$

Για μικρό Δt , η ποσότητα $f(t)\Delta t$ εκφράζει κατά προσέγγιση την πιθανότητα 'θανάτου-αποτυχίας' στο διάστημα $[t, t + \Delta t)$. Η σ.π.π. κατανομών χρόνων ζωής, έχει τις παρακάτω ιδιότητες:

$$f(t) \geq 0, \quad \text{για κάθε } t \geq 0 \quad \text{και} \quad \int_0^{\infty} f(t) dt = 1.$$

Η δυική συνάρτηση της συνάρτησης επιβίωσης

$$F(t) = P(T \leq t) = 1 - S(t), \quad t \geq 0,$$

είναι η αθροιστική συνάρτηση κατανομής της $f(t)$ (cumulative distribution function). Από τις ιδιότητες της συνάρτησης επιβίωσης, έπεται ότι η F είναι αύξουσα και συνεχής συνάρτηση του t , με $F(0) = 0$ και $\lim_{t \rightarrow \infty} F(t) = 1$.

Η συνάρτηση κινδύνου (hazard function ή failure rate ή force of mortality), εκφράζει το στιγμιαίο ρυθμό 'θανάτου' ή 'αποτυχίας' τη χρονική στιγμή t , δοθείσης της επιβίωσης μέχρι το χρόνο t και δίνεται από τη σχέση:

$$h(t) = \lim_{\Delta t \rightarrow \infty} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \quad t \geq 0. \quad (1.2)$$

Για μικρό Δt , η ποσότητα $h(t)\Delta t$, εκφράζει κατά προσέγγιση την πιθανότητα θανάτου στο διάστημα $[t, t + \Delta t)$, δοθείσης της επιβίωσης μέχρι το χρόνο t . Η συνάρτηση κινδύνου για συνεχή κατανομή χρόνου ζωής, έχει τις παρακάτω ιδιότητες:

$$h(t) \geq 0, \quad \text{για κάθε } t \geq 0 \quad \text{και} \quad \int_0^{\infty} h(t) dt = \infty.$$

Από τις σχέσεις (1.1) και (1.2), προκύπτει ότι η συνάρτηση επιβίωσης γράφεται στη μορφή

$$S(t) = e^{-H(t)},$$



όπου η συνάρτηση H , δίνεται από τη σχέση:

$$H(t) = \int_0^t h(t) dt$$

και ονομάζεται **αθροιστική συνάρτηση κινδύνου** (cumulative hazard function ή integrated hazard function). Η τελευταία είναι μη φθίνουσα συνάρτηση του t , με $H(0) = 0$ και

$$\lim_{t \rightarrow \infty} H(t) = \infty.$$

Ένα μέτρο βιωσιμότητας (ή αξιοπιστίας) στη μελέτη της διάρκειας ζωής βιολογικών οργανισμών (ή εργοστασιακών προϊόντων), είναι ο **μέσος υπολειπόμενος χρόνος ζωής** (mean residual lifetime). Η συνάρτησή του, δίνει τις τιμές της αναμενόμενης διάρκειας ζωής από την στιγμή t μέχρι την 'αποτυχία', δοθείσης της επιβίωσης μέχρι το χρόνο t και δίνεται από τη σχέση:

$$m(t) = E(T - t | T \geq t) = \frac{\int_t^{\infty} S(t) dt}{S(t)}, \quad t \geq 0.$$

Για κατανομές ανάλυσης επιβίωσης με πεπερασμένη μέση τιμή, η συνάρτηση του μέσου υπολειπόμενου χρόνου ζωής, έχει τις ακόλουθες ιδιότητες:

$$m(t) \geq 0, \quad m'(t) \geq -1, \quad \text{για κάθε } t \geq 0 \quad \text{και} \quad \int_0^{\infty} \frac{1}{m(t)} dt = \infty.$$

Στον Πίνακα 1.1, παρουσιάζεται ο τρόπος προσδιορισμού καθεμίας από τις συναρτήσεις που προαναφέρθηκαν, αν μια από τις υπόλοιπες είναι γνωστή.

Αξίζει να σημειωθεί ότι υπάρχουν και άλλοι ισοδύναμοι τρόποι ορισμού της κατανομής της T , όπως για παράδειγμα, αυτοί της αντίστροφης συνάρτησης κινδύνου (reversed hazard rate, Gupta et al., 2004) και του αναμενόμενου χρόνου αδράνειας (expected inactivity time, Ghitany et al., 2005) που εμφανίζονται σχετικά πρόσφατα στη βιβλιογραφία ή οι προγενέστεροί τους, της ροπογεννήτριας συνάρτησης $E(e^{sT})$ (moment generating function) και της χαρακτηριστικής συνάρτησης $E(e^{isT})$ (characteristic function). Η επιλογή αυτών που παρουσιάστηκαν, έγινε στη βάση της δημοτικότητάς τους στην ανάλυση της επιβίωσης, λόγω της απτής φυσικής τους ερμηνείας.



	$f(t)$	$F(t)$	$S(t)$	$h(t)$	$H(t)$	$m(t)$
$f(t)$	•	$\int_0^t f(t) dt$	$\int_t^\infty f(t) dt$	$\frac{f(t)}{\int_t^\infty f(t) dt}$	$-\log \left[\int_t^\infty f(t) dt \right]$	$\frac{1}{\int_t^\infty f(t) dt} - 1$
$F(t)$	$F'(t)$	•	$1 - F(t)$	$-\left[\log\{1 - F(t)\} \right]'$	$-\log\{1 - F(t)\}$	$\frac{1}{1 - F(t)} \int_t^\infty \{1 - F(t)\} dt$
$S(t)$	$-S'(t)$	$1 - S(t)$	•	$-\left[\log\{S(t)\} \right]'$	$-\log\{S(t)\}$	$\frac{1}{S(t)} \int_t^\infty S(t) dt$
$h(t)$	$h(t)e^{-\int_0^t h(t) dt}$	$1 - e^{-\int_0^t h(t) dt}$	$e^{-\int_0^t h(t) dt}$	•	$\int_0^t h(t) dt$	$\frac{1}{e^{-\int_0^t h(t) dt}} - \int_0^t h(t) dt$
$H(t)$	$H'(t)e^{-H(t)}$	$1 - e^{-H(t)}$	$e^{-H(t)}$	$H'(t)$	•	$e^{H(t)} \int_t^\infty e^{-H(t)} dt$
$m(t)$	$\frac{1 + m'(t)}{m(t)} e^{-\int_0^t \frac{1 + m'(t)}{m(t)} dt}$	$1 - e^{-\int_0^t \frac{1 + m'(t)}{m(t)} dt}$	$e^{-\int_0^t \frac{1 + m'(t)}{m(t)} dt}$	$\frac{1 + m'(t)}{m(t)}$	$\int_0^t \frac{1 + m'(t)}{m(t)} dt$	•

Πίνακας 1.1 Σχέσεις προσδιορισμού καθέμιας από τις $f(t)$, $F(t)$, $S(t)$, $h(t)$, $H(t)$ και $m(t)$ (στήλες), συναρτήσει μιας των υπολοίπων (γραμμές).

Οι ροπές και τα εκατοστιαία σημεία δεν προσδιορίζουν μονοσήμαντα την κατανομή του χρόνου ζωής, αλλά συχνά αποτελούν ένα χρήσιμο τρόπο συνοπτικής περιγραφής της. Οι ροπές r τάξης περί το μηδέν, προσδιορίζονται από το ολοκλήρωμα:

$$E(T^r) = \int_0^{\infty} t^r f(t) dt.$$

Η μέση τιμή $E(T)$, αποτελεί ένα μέτρο της κεντρικής τάσης των τιμών της T ενώ η διακύμανση

$$Var(T) = E(T^2) - \{E(T)\}^2,$$

ένα μέτρο της διασποράς των τιμών της, γύρω από τη μέση τιμή.

Τα p εκατοστιαία σημεία, δίνονται από τη σχέση:

$$F(t_p) = p \quad \text{ή} \quad t_p = F^{-1}(p)$$

και προσδιορίζουν το χρόνο μέχρι τον οποίο δυνητικά επιβιώνει ένα προκαθορισμένο ποσοστό $1-p$ των πειραματικών μονάδων, ή διαφορετικά το χρόνο μέχρι τον οποίο η πιθανότητα επιβίωσης ισούται με $1-p$.

1.1.1 Η συνάρτηση κινδύνου και η σημασία της

Αν και καθεμία από τις συναρτήσεις που προαναφέρθηκαν φανερώνει μια διαφορετική πτυχή της επιβίωσης, αναμφίβολα, η συνάρτηση κινδύνου είναι το πιο αντιπροσωπευτικό μέτρο περιγραφής της ικανότητας επιβίωσης ενός πληθυσμού. Δεδομένου ότι περιγράφει τον τρόπο με τον οποίο μεταβάλλεται ο στιγμιαίος ρυθμός 'θανάτου-αποτυχίας' στο χρόνο, η εκ των προτέρων πληροφορία για τη μονοτονία της, βοηθά στην επιλογή κατάλληλου μοντέλου για την περιγραφή χρόνων ζωής.

Με εξαίρεση τις περιπτώσεις που η υπόθεση της σταθερής συνάρτησης κινδύνου είναι ρεαλιστική, κατά κανόνα, η επιβίωση περιγράφεται με κατανομές των οποίων η αντίστοιχη συνάρτηση κινδύνου είναι αύξουσα, φθίνουσα, κυπελλοειδούς μορφής ή μονοκόρυφη. Με βάση τις τέσσερις τελευταίες περιπτώσεις μονοτονίας, μια κατανομή ανάλυσης επιβίωσης χαρακτηρίζεται αντίστοιχα ως IFR (increasing failure rate), DFR (decreasing failure rate), BTFR (bathtub shaped failure rate) και UBTFR (upside down bathtub shaped failure rate). Οι μαθηματικοί ορισμοί των παραπάνω ιδιοτήτων, δίνονται στη συνέχεια (Glaser, 1980).



Ορισμός

Μια κατανομή ανάλυσης επιβίωσης, είναι:

- i) IFR (DFR) αν και μόνο αν $h'(t) > 0$ για κάθε $t > 0$ ($h'(t) < 0$ για κάθε $t > 0$).
- ii) BTFR (UBTFR) αν και μόνο αν υπάρχει $t_0 > 0$ τέτοιο ώστε $h'(t) < 0$ για κάθε $t \in (0, t_0)$ και $h'(t) > 0$ για κάθε $t \in (t_0, \infty)$, με $h'(t_0) = 0$ (αν και μόνο αν υπάρχει $t_0 > 0$ τέτοιο ώστε $h'(t) > 0$ για κάθε $t \in (0, t_0)$ και $h'(t) < 0$ για κάθε $t \in (t_0, \infty)$, με $h'(t_0) = 0$).

Συχνά, το ενδιαφέρον επικεντρώνεται σε μια περίοδο της συνολικής διάρκειας ζωής, στην οποία επέρχεται γήρανση ή φθορά χρήσης. Το γεγονός αυτό, δικαιολογεί εν μέρει, τη μεγάλη δημοτικότητα των κατανομών χρόνων ζωής με αύξουσα συνάρτηση κινδύνου.

Αντίθετα, όταν μελετάται η συνολική διάρκεια επιβίωσης, τα πλέον ρεαλιστικά μοντέλα για την περιγραφή χρόνων ζωής, είναι εκείνα με συνάρτηση κινδύνου κυπελλοειδούς μορφής. Για παράδειγμα, σε ότι αφορά στη διάρκεια ζωής βιολογικών οργανισμών, ο ρυθμός θανάτου τους αρχικά είναι σχετικά υψηλός (περίοδος νεογνικής θνησιμότητας), φθίνει προοδευτικά μέχρι μιας συγκεκριμένης τιμής όπου προσωρινά διατηρείται σχεδόν σταθερός (περίοδος νεότητας) και στη συνέχεια αυξάνει λόγω γήρατος (περίοδος γηρατειών). Ανάλογη συμπεριφορά παρουσιάζει και η συνάρτηση κινδύνου πολλών εργοστασιακών προϊόντων (π.χ. ηλεκτρικές συσκευές), με μεγάλη συχνότητα δυσλειτουργικών προβλημάτων με την έναρξη της λειτουργίας τους (λόγω κατασκευαστικών σφαλμάτων) ή μετά από χρήση (λόγω φθοράς) και σχετικά ομαλή λειτουργία στο μεσοδιάστημα. Βεβαίως, όταν διενεργείται από τον κατασκευαστή έλεγχος της ποιότητας της λειτουργίας τους, με σκοπό την απομάκρυνση των ελαττωματικών πριν τη διοχέτευσή τους στην αγορά, ο χρόνος ζωής των υπολοίπων όπως καταγράφεται από τον καταναλωτή, αναμένεται να περιγράφεται από αύξουσα συνάρτηση κινδύνου.

Τα μοντέλα με μονοκόρυφη συνάρτηση κινδύνου χρησιμοποιούνται κυρίως στην περίπτωση που παρατηρείται μεγάλη συχνότητα 'αποτυχιών' κοντά στη χρονική αφετηρία του πειράματος που παράγει τα δεδομένα ενώ εκείνα με φθίνουσα συνάρτηση κινδύνου, χρησιμοποιούνται ως επί το πλείστον για την περιγραφή της διάρκειας ζωής ηλεκτρονικών ή μηχανικών εξαρτημάτων, τα οποία παρουσιάζουν βελτίωση στη λειτουργία τους με την πάροδο του χρόνου.



Τέλος, η Εκθετική κατανομή με σταθερή συνάρτηση κινδύνου, παρουσιάζει μεγάλο εύρος εφαρμογών στη μοντελοποίηση χρόνων ζωής ενώ οι περιπτώσεις πληθυσμών με διαφορετικό -σε σχέση με αυτά που προαναφέρθηκαν- είδος μονοτονίας του στιγμιαίου ρυθμού 'θανάτου-αποτυχίας', είναι σχετικά σπάνιες (Lawless, 1982).

1.1.2 Βασικές κατανομές ανάλυσης επιβίωσης

Στην ενότητα αυτή, παρουσιάζονται οι βασικές κατανομές ανάλυσης της επιβίωσης. Εκτενής μελέτη της Εκθετικής κατανομής και των κατανομών Weibull, Γάμα και Lognormal, υπάρχει στα κλασικά βιβλία θεωρίας κατανομών (π.χ. Johnson et al., 1994) και στα περισσότερα εγχειρίδια ανάλυσης επιβίωσης (π.χ. Nelson, 1982, Lee, 1992 και Leemis, 1995). Επίσης, παρουσιάζονται οι κατανομές Inverse Gaussian (Johnson et al., 1994), Gompertz (Jordan, 1967), Lomax (Johnson et al., 1994), Modified Extreme Value (Johnson et al., 1995), EG (Adamidis και Loukas, 1998), IDB (Hjorth, 1980) και Exponential Power (Smith και Bain, 1975). Άλλες κατανομές ανάλυσης επιβίωσης, είναι οι κατανομές Rayleigh (Johnson et al., 1994), compound Rayleigh (Mostert et al., 1999) και γενικευμένη compound Rayleigh (Bekker et al., 2000), η γενικευμένη Γάμα κατανομή (Stacy, 1962), οι κατανομές Log-Logistic (Kalbfleisch και Prentice, 2002), Makeham (Jordan, 1967) και Muth (1977), οι κατανομές των Chen (2000) και Haupt και Schabe (1997) και η κατανομή Torpp-Leone (Ghitany et al., 2005). Άλλες, εκτός από αυτές που προαναφέρθηκαν, αναφέρονται στους Johnson et al. (1995) και Leemis (1995) ενώ οι Rajarshi και Rajarshi (1988) και πιο πρόσφατα οι Lai et al. (2001), παρουσιάζουν μια ανασκόπηση στα μοντέλα με κυπελλοειδούς μορφής συνάρτηση κινδύνου.

Η Εκθετική κατανομή

Η υπόθεση της σταθερής συνάρτησης κινδύνου, $h(t; \lambda) = \lambda > 0$ (λ σταθερά) ώστε η πιθανότητα επιβίωσης να είναι ανεξάρτητη από το χρόνο ζωής που έχει διανυθεί (ιδιότητα αμνησίας), προσδιορίζει μονοσήμαντα την Εκθετική κατανομή, με συναρτήσεις πυκνότητας πιθανότητας και επιβίωσης

$$f(t; \lambda) = \lambda e^{-\lambda t}, \quad t \geq 0,$$



$$S(t; \lambda) = e^{-\lambda t}, \quad t \geq 0,$$

αντίστοιχα και σταθερό μέσο υπολειπόμενο χρόνο ζωής

$$m(t; \lambda) = \frac{1}{\lambda}, \quad t \geq 0.$$

Οι αντίστοιχες ροπές r τάξης περί το μηδέν, δίνονται από τη σχέση:

$$E(T^r; \lambda) = \frac{\Gamma(r+1)}{\lambda^r},$$

όπου

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt,$$

είναι η συνάρτηση Γάμα. Κατά συνέπεια, η μέση τιμή και η διακύμανση της Εκθετικής κατανομής, είναι αντίστοιχα

$$E(T; \lambda) = \frac{1}{\lambda},$$

$$\text{Var}(T; \lambda) = \frac{1}{\lambda^2}.$$

Ιστορικά, η εκθετική κατανομή αποτελεί το πρώτο ευρέως χρησιμοποιούμενο μοντέλο ανάλυσης επιβίωσης, αφενός λόγω της απλής συναρτησιακής της μορφής και της ευκολίας στο στατιστικό χειρισμό της, αφετέρου λόγω της καταλληλότητάς της στην περιγραφή χρόνων ζωής σε πληθώρα περιπτώσεων. Οι εφαρμογές της κυμαίνονται από μελέτες στη διάρκεια ζωής εργοστασιακών προϊόντων, μέχρι έρευνες στους χρόνους ανάνηψης ή υποτροπής από χρόνιες ασθένειες. Βεβαίως, η υπόθεση της σταθερής συνάρτησης κινδύνου δεν παύει να είναι αρκετά περιοριστική.

Η κατανομή Weibull

Η κατανομή Weibull είναι ίσως η πιο δημοφιλής κατανομή ανάλυσης επιβίωσης, με ένα εξαιρετικά ευρύ φάσμα εφαρμογών τόσο στο χώρο της βιομηχανίας όσο και στο χώρο της ιατρικής και των άλλων βιολογικών επιστημών. Αφενός η καταλληλότητά της στην περιγραφή χρόνων ζωής και αφετέρου οι απλές εκφράσεις των αντίστοιχων συναρτήσεων πυκνότητας πιθανότητας, επιβίωσης και κινδύνου, εξηγούν εν μέρει τη δημοτικότητά της.



Η σ.π.π. της κατανομής Weibull, δίνεται από τη σχέση:

$$f(t; \theta) = \lambda a t^{a-1} \exp(-\lambda t^a), \quad t > 0,$$

με $\theta = (a, \lambda) \in R^+ \times R^+$, όπου a και λ είναι αντίστοιχα παράμετροι μορφής (shape) και κλίμακας (scale).

Η συνάρτηση επιβίωσής της, είναι

$$S(t; \theta) = \exp(-\lambda t^a), \quad t > 0,$$

ενώ η αντίστοιχη συνάρτηση κινδύνου έχει τη μορφή

$$h(t; \theta) = \lambda a t^{a-1}, \quad t > 0$$

και είναι αύξουσα για $a > 1$, με $\lim_{t \rightarrow 0^+} h(t; \theta) = 0$ και $\lim_{t \rightarrow \infty} h(t; \theta) = \infty$, φθίνουσα για $0 < a < 1$,

με $\lim_{t \rightarrow 0^+} h(t; \theta) = \infty$ και $\lim_{t \rightarrow \infty} h(t; \theta) = 0$ και σταθερή για $a = 1$ (Εκθετική).

Η συνάρτηση του μέσου υπολειπόμενου χρόνου ζωής της, είναι

$$m(t; \theta) = \frac{\exp(\lambda t^a)}{a \lambda^{1/a}} \Gamma\left(\frac{1}{a}\right) \left\{ 1 - I\left(\frac{1}{a}, \lambda t^a\right) \right\}, \quad t > 0,$$

όπου

$$I(a, t) = \frac{1}{\Gamma(a)} \int_0^t u^{a-1} e^{-u} du,$$

είναι η μη πλήρης συνάρτηση Γάμα (incomplete Gamma function).

Οι ροπές r τάξης περί το μηδέν της κατανομής Weibull, δίνονται από τη σχέση:

$$E(T^r; \theta) = \frac{r}{a \lambda^{r/a}} \Gamma\left(\frac{r}{a}\right),$$

όπου $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$. Κατά συνέπεια, η μέση τιμή της είναι

$$E(T; \theta) = \frac{1}{a \lambda^{1/a}} \Gamma\left(\frac{1}{a}\right)$$

ενώ η διακύμανσή της, προσδιορίζεται ως

$$\text{Var}(T; \theta) = \frac{1}{\lambda^{2/a}} \left[\Gamma\left(1 + \frac{2}{a}\right) - \left\{ \Gamma\left(1 + \frac{1}{a}\right) \right\}^2 \right].$$



Η κατανομή Γάμα

Η κατανομή Γάμα αποτελεί τη δεύτερη σημαντική γενίκευση της Εκθετικής κατανομής. Χρησιμοποιείται ευρύτατα στην ανάλυση επιβίωσης αλλά όχι στον ίδιο βαθμό με την κατανομή Weibull αφού οι συναρτήσεις επιβίωσης και κινδύνου της, δεν μπορούν να γραφούν σε απλή κλειστή μορφή και η στατιστική συμπερασματολογία ειδικότερα στην περίπτωση λογοκριμένων δειγμάτων, παρουσιάζει ιδιαίτερες δυσκολίες.

Η σ.π.π. της κατανομής Γάμα, δίνεται από τη σχέση:

$$f(t; \theta) = \frac{1}{\Gamma(a)} \lambda (\lambda t)^{a-1} \exp(-\lambda t), \quad t > 0,$$

με $\theta = (a, \lambda) \in R^+ \times R^+$, όπου a είναι παράμετρος μορφής και λ παράμετρος κλίμακας.

Οι συναρτήσεις επιβίωσης και κινδύνου της, γράφονται συναρτήσει της μη πλήρους συνάρτησης Γάμα, στη μορφή:

$$S(t; \theta) = 1 - I(a, \lambda t), \quad t > 0,$$

$$h(t; \theta) = \frac{f(t; \theta)}{1 - I(a, \lambda t)}, \quad t > 0.$$

Η συνάρτηση κινδύνου είναι αύξουσα για $a > 1$, με $\lim_{t \rightarrow 0^+} h(t; \theta) = 0$ και $\lim_{t \rightarrow \infty} h(t; \theta) = \lambda$,

φθίνουσα για $0 < a < 1$, με $\lim_{t \rightarrow 0^+} h(t; \theta) = \infty$ και $\lim_{t \rightarrow \infty} h(t; \theta) = \lambda$ και σταθερή για $a = 1$

(Εκθετική).

Οι ροπές r τάξης περί το μηδέν της κατανομής Γάμα, δίνονται από τη σχέση:

$$E(T^r; \theta) = \frac{a(a+1)\dots(a+r-1)}{\lambda^r}$$

και κατά συνέπεια, η μέση τιμή και η διακύμανσή της, είναι αντίστοιχα

$$E(T; \theta) = \frac{a}{\lambda},$$

$$\text{Var}(T; \theta) = \frac{a}{\lambda^2}.$$

Αξίζει να σημειωθεί ότι όταν η παράμετρος μορφής παίρνει θετικές ακέραιες τιμές, $a = n$, οι συναρτήσεις πυκνότητας πιθανότητας και επιβίωσης της κατανομής Γάμα, γράφονται ως



$$f(t; \theta) = \frac{1}{(n-1)!} \lambda (\lambda t)^{n-1} \exp(-\lambda t), \quad t > 0,$$

$$S(t; \theta) = \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad t > 0,$$

αντίστοιχα (κατανομή Erlang).

Η κατανομή Log-normal

Στα πλαίσια ανάλυσης της επιβίωσης, η κατανομή Log-normal χρησιμοποιείται σε ποικίλες περιπτώσεις. Για παράδειγμα, έχει χρησιμοποιηθεί στην περιγραφή του χρόνου μέχρι την εμφάνιση καρκίνου των πνευμόνων σε καπνιστές.

Η τυχαία μεταβλητή T ακολουθεί κατανομή Log-normal, αν η $Y = \log T$ ακολουθεί κανονική κατανομή. Αν η μέση τιμή και η διακύμανση της τελευταίας είναι αντίστοιχα μ και σ^2 , τότε η σ.π.π. της κατανομής Log-normal, δίνεται από τη σχέση:

$$f(t; \theta) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log t - \mu)^2}{2\sigma^2}\right\}, \quad t > 0,$$

όπου $\theta = (\sigma^2, \mu) \in R^+ \times R$.

Οι συναρτήσεις επιβίωσης και κινδύνου της, είναι αντίστοιχα

$$S(t; \theta) = 1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right), \quad t > 0,$$

$$h(t; \theta) = \frac{f(t; \theta)}{S(t; \theta)}, \quad t > 0,$$

όπου

$$\Phi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt,$$

είναι η αθροιστική συνάρτηση της τυπικής κανονικής κατανομής.

Η μέση τιμή και η διακύμανσή της, είναι αντίστοιχα

$$E(T; \theta) = e^{\mu + \sigma^2/2},$$

$$\text{Var}(T; \theta) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1).$$

Το βασικό μειονέκτημα της κατανομής Log-normal, αφορά στη μονοτονία της



συνάρτησης κινδύνου της, η οποία αυξάνει από την τιμή μηδέν μέχρι να πάρει τη μέγιστη τιμή της και στη συνέχεια φθίνει τείνοντας ασυμπτωτικά στο μηδέν. Η φθίνουσα τάση των τιμών της συνάρτησης κινδύνου από κάποια χρονική στιγμή και μετά, περιορίζει το εύρος εφαρμογών της κατανομής, στις περιπτώσεις που μεγάλες τιμές του t δεν παρουσιάζουν ενδιαφέρον.

Η κατανομή Inverse Gaussian

Η κατανομή Inverse Gaussian, προέκυψε από την ανάγκη περιγραφής της κίνησης Brown. Όπως και η κατανομή Log-normal, χρησιμοποιείται κυρίως όταν παρατηρείται μεγάλη συχνότητα πρόωρων 'αποτυχιών'.

Η σ.π.π. της κατανομής Inverse Gaussian, δίνεται από τη σχέση:

$$f(t; \theta) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left\{-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right\}, \quad t > 0,$$

όπου $\theta = (\mu, \lambda) \in R^+ \times R^+$ ενώ οι αντίστοιχες συναρτήσεις επιβίωσης και κινδύνου είναι

$$S(t; \theta) = \Phi\left\{\sqrt{\frac{\lambda}{t}}\left(1 - \frac{t}{\mu}\right)\right\} - e^{2\lambda/\mu} \Phi\left\{-\sqrt{\frac{\lambda}{t}}\left(1 + \frac{t}{\mu}\right)\right\}, \quad t > 0,$$

$$h(t; \theta) = \frac{f(t; \theta)}{S(t; \theta)}, \quad t > 0,$$

όπου Φ είναι η αθροιστική συνάρτηση της τυπικής κανονικής κατανομής.

Η συνάρτηση κινδύνου αυξάνει από το μηδέν μέχρι να πάρει τη μέγιστη τιμή της και στη συνέχεια φθίνει, τείνοντας ασυμπτωτικά στην τιμή $\lambda(2\mu^2)^{-1}$. Βεβαίως, για τιμές του λ πολύ μεγαλύτερες από εκείνες του μ , είναι ουσιαστικά μη φθίνουσα για κάθε τιμή του t (Chhikara και Folks, 1977).

Η μέση τιμή και η διακύμανσή της, είναι αντίστοιχα

$$E(T; \theta) = \mu,$$

$$Var(T; \theta) = \frac{\mu^3}{\lambda}.$$

Η κατανομή Gompertz

Η κατανομή Gompertz, έχει χρησιμοποιηθεί κυρίως στη μοντελοποίηση του χρόνου ζωής ενηλίκων, σε μελέτες ασφαλιστικών επιστημών.



Η σ.π.π. της, δίνεται από τη σχέση:

$$f(t; \theta) = \lambda a^t \exp\left\{-\frac{\lambda(a^t - 1)}{\log a}\right\}, \quad t \geq 0,$$

με $\theta = (a, \lambda)$, όπου $a > 1$ και $\lambda > 0$ είναι παράμετροι μορφής. Οι αντίστοιχες συναρτήσεις επιβίωσης και κινδύνου, είναι

$$S(t; \theta) = \exp\left\{-\frac{\lambda(a^t - 1)}{\log a}\right\}, \quad t \geq 0,$$

$$h(t; \theta) = \lambda a^t, \quad t \geq 0.$$

Η συνάρτηση κινδύνου είναι αύξουσα συνάρτηση του χρόνου, με ελάχιστη τιμή την $h(0; \theta) = \lambda$.

Η κατανομή Pareto δεύτερου τύπου

Η κατανομή Pareto δεύτερου τύπου (γνωστή και ως κατανομή του Lomax), έχει χρησιμοποιηθεί ως επί το πλείστον, στη μελέτη της βιωσιμότητας επιχειρήσεων.

Η σ.π.π. της, δίνεται από τη σχέση:

$$f(t; \theta) = \frac{a\lambda}{(1 + \lambda t)^{a+1}}, \quad t \geq 0,$$

με $\theta = (a, \lambda) \in \mathbb{R}^+ \times \mathbb{R}^+$, όπου a και λ είναι παράμετροι μορφής και κλίμακας αντίστοιχα.

Οι συναρτήσεις επιβίωσης και κινδύνου της, δίνονται αντίστοιχα από τις σχέσεις:

$$S(t; \theta) = \frac{1}{(1 + \lambda t)^a}, \quad t \geq 0,$$

$$h(t; \theta) = \frac{a\lambda}{1 + \lambda t}, \quad t \geq 0,$$

όπου η τελευταία, είναι φθίνουσα με $h(0; \theta) = a\lambda$ και $\lim_{t \rightarrow \infty} h(t; \theta) = 0$.

Οι ροπές r τάξης περί το μηδέν της κατανομής, δίνονται από τη σχέση:

$$E(T^r; \theta) = \frac{\Gamma(1+r)\Gamma(a-r)}{\Gamma(a)}, \quad -1 < r < a$$

και επομένως, η μέση τιμή και η διακύμανσή της, είναι αντίστοιχα

$$E(T; \theta) = \frac{1}{\lambda(a-1)}, \quad a > 1,$$



$$\text{Var}(T; \theta) = \frac{a}{(a-1)^2(a-2)\lambda^2}, \quad a > 2.$$

Η κατανομή Modified Extreme Value

Η κατανομή Modified Extreme Value, χρησιμοποιείται για την περιγραφή της επιβίωσης πληθυσμών που γερνάνε ή φθείρονται με ραγδαίο ρυθμό με την πάροδο του χρόνου.

Η σ.π.π. της Modified Extreme Value κατανομής, είναι

$$f(t; \theta) = a\lambda \exp\{\lambda t + a(1 - e^{\lambda t})\}, \quad t \geq 0,$$

με $\theta = (a, \lambda) \in R^+ \times R^+$, όπου η a είναι παράμετρος μορφής και η λ παράμετρος κλίμακας ενώ η συνάρτηση επιβίωσής της, δίνεται από τη σχέση:

$$S(t; \theta) = \exp\{a(1 - e^{\lambda t})\}, \quad t \geq 0.$$

Η αντίστοιχη συνάρτηση κινδύνου

$$h(t; \theta) = a\lambda \exp(\lambda t), \quad t \geq 0,$$

αυξάνει εκθετικά με το χρόνο, με $h(0; \theta) = a\lambda$ και $\lim_{t \rightarrow \infty} h(t; \theta) = \infty$.

Η Εκθετική-Γεωμετρική κατανομή

Η Εκθετική-Γεωμετρική κατανομή (EG), έχει χρησιμοποιηθεί για τη μοντελοποίηση του χρόνου μεταξύ διαδοχικών ατυχημάτων σε ανθρακωρυχεία και επίσης, στην περιγραφή του χρόνου μεταξύ διαδοχικών δυσλειτουργικών σφαλμάτων, στη λειτουργία συστημάτων κλιματισμού αεροπλάνων.

Η σ.π.π. της κατανομής EG , δίνεται από τη σχέση:

$$f(t; \theta) = \frac{\lambda(1-a)e^{-\lambda t}}{(1-ae^{-\beta t})^2}, \quad t \geq 0,$$

με $\theta = (a, \lambda) \in (0, 1) \times R^+$, όπου η a είναι παράμετρος μορφής και η λ , κλίμακας.

Οι αντίστοιχες συναρτήσεις επιβίωσης και κινδύνου, είναι

$$S(t; \theta) = \frac{(1-a)e^{-\lambda t}}{1-ae^{-\lambda t}}, \quad t \geq 0,$$

$$h(t; \theta) = \frac{\lambda}{1-ae^{-\lambda t}}, \quad t \geq 0.$$



Η συνάρτηση κινδύνου είναι φθίνουσα, με $h(0; \theta) = \lambda(1-a)^{-1}$ και $\lim_{t \rightarrow \infty} h(t; \theta) = \lambda$.

Η συνάρτηση του μέσου υπολειπόμενου χρόνου ζωής είναι

$$m(t; \theta) = -\frac{e^{\lambda t}(1-ae^{-\lambda t})\log(1-ae^{-\lambda t})}{a\lambda}, \quad t \geq 0.$$

Οι ροπές r τάξης περί το μηδέν της κατανομής EG , δίνονται από τη σχέση:

$$E(T^r; \theta) = \frac{(1-a)r!}{a\lambda^r} Q(a; r),$$

όπου $Q(a; r) = \sum_{j=1}^{\infty} a^j j^{-r}$ είναι η πολυλογαριθμική συνάρτηση (Erdelyi et al., 1953, σελ. 31).

Κατά συνέπεια, η μέση τιμή και η διακύμανσή της, είναι αντίστοιχα

$$E(T; \theta) = \frac{-(1-a)\log(1-a)}{a\lambda},$$

$$Var(T; \theta) = \frac{1-a}{a\lambda^2} \left\{ 2Q(a; 2) - \frac{(1-a)\log^2(1-a)}{a} \right\}.$$

Η κατανομή του Hjorth

Η κατανομή του Hjorth (IDB), έχει σ.π.π.

$$f(t; \theta) = \frac{\{(1+\lambda t)\beta t + a\} \exp(-\beta t^2 / 2)}{(1+\lambda t)^{a/\lambda+1}}, \quad t \geq 0,$$

με $\theta = (a, \lambda, \beta) \in R^+ \times R^+ \times R^+$, όπου λ και β είναι παράμετροι κλίμακας και a παράμετρος μορφής.

Οι αντίστοιχες συναρτήσεις επιβίωσης και κινδύνου είναι

$$S(t; \theta) = \frac{\exp(-\beta t^2 / 2)}{(1+\lambda t)^{a/\lambda}}, \quad t \geq 0,$$

$$h(t; \theta) = \beta t + \frac{a}{1+\lambda t}, \quad t \geq 0.$$

Η συνάρτηση κινδύνου είναι γραμμικά αύξουσα για $a = 0$ (κατανομή Rayleigh), σταθερή για $\beta = \lambda = 0$ (Εκθετική κατανομή), φθίνουσα για $\beta = 0$, αύξουσα για $\beta \geq a\lambda$ και κυπελλοειδούς μορφής για $0 < \beta < a\lambda$.



Η κατανομή Exponential Power

Η κατανομή Exponential Power είναι από τις λίγες κατανομές ανάλυσης επιβίωσης δύο παραμέτρων, που μπορούν να δώσουν κυπελλοειδούς μορφής συνάρτηση κινδύνου.

Η σ.π.π. της, είναι:

$$f(t; \theta) = a\lambda t^{a-1} \exp(1 + \lambda t^a - e^{\lambda t^a}), \quad t > 0,$$

με $\theta = (a, \lambda) \in R^+ \times R^+$ όπου a και λ είναι παράμετροι μορφής και κλίμακας αντίστοιχα.

Οι συναρτήσεις επιβίωσης και κινδύνου της, είναι

$$S(t; \theta) = \exp(1 - e^{\lambda t^a}), \quad t > 0,$$

$$h(t; \theta) = a\lambda t^{a-1} \exp(\lambda t^a), \quad t > 0.$$

Η συνάρτηση κινδύνου είναι κυπελλοειδούς μορφής για $0 < a < 1$, με ελάχιστη τιμή για $t = \{(1-a)/(a\lambda)\}^{1/a}$ και αύξουσα για $a \geq 1$.

1.2 Ειδικά μοντέλα ανάλυσης επιβίωσης

Υπάρχουν πολλοί τρόποι διαχείρισης και συνδυασμού των παραπάνω κατανομών, για την περιγραφή της επιβίωσης σε πιο σύνθετες καταστάσεις, δύο από τους οποίους οδηγούν στα μοντέλα ανταγωνιστικών κινδύνων και ευπάθειας. Τα μοντέλα αυτά, που παρουσιάζονται συνοπτικά στη συνέχεια, περιγράφουν αντίστοιχα, το χρόνο ζωής πληθυσμών που είναι εκτεθειμένοι σε πολλούς παράγοντες 'αποτυχίας' ή είναι ανομοιογενείς ως προς την ικανότητα επιβίωσης των μελών τους.

1.2.1 Μοντέλα ανταγωνιστικών κινδύνων

Τα μοντέλα ανταγωνιστικών κινδύνων (competing risks models, Leemis, 1995), περιγράφουν το χρόνο ζωής πληθυσμών, εκτεθειμένων σε k διαφορετικά αίτια 'αποτυχίας' (κινδύνους), τα οποία λειτουργούν ανταγωνιστικά ως προς το πιο θα επιφέρει 'αποτυχία'. Για παράδειγμα, ένας ασφαλιστής θα μπορούσε να χρησιμοποιήσει τα καρδιακά νοσήματα, τον καρκίνο, τα δυστυχήματα και όλα τα άλλα αίτια θανάτου, ως τέσσερις κινδύνους για το χρόνο ζωής του ανθρώπου.

Αν T_i , ($i = 1, \dots, k$) είναι ο χρόνος ζωής που θα παρατηρούνταν αν ο πληθυσμός ήταν



εκτεθειμένος μόνο στον $i^{\text{οστό}}$ κίνδυνο, τότε ο παρατηρούμενος χρόνος ζωής είναι $T = \min\{T_i\}_{i=1}^k$. Υποθέτοντας ότι οι τυχαίες μεταβλητές T_i είναι ανεξάρτητες, με συναρτήσεις πυκνότητας πιθανότητας, επιβίωσης και κινδύνου f_i , S_i και h_i αντίστοιχα, τότε:

$$f_T(t) = \sum_{i=1}^k f_i(t) \prod_{\substack{j=1 \\ j \neq i}}^k S_j(t), \quad t \geq 0,$$

$$S_T(t) = \prod_{i=1}^k S_i(t), \quad t \geq 0,$$

$$h_T(t) = \sum_{i=1}^k h_i(t), \quad t \geq 0.$$

Μια διαφορετική φυσική ερμηνεία των μοντέλων ανταγωνιστικών κινδύνων, μπορεί να δοθεί μέσω των σειριακών συστημάτων (series systems) αφού κάθε κίνδυνος μπορεί να θεωρηθεί ταυτόσημος με το ενδεχόμενο αποτυχίας μιας από τις συνιστώσες ενός τέτοιου συστήματος, στο οποίο επέρχεται αποτυχία όταν ένα οποιοδήποτε εξάρτημά του πάψει να λειτουργεί.

1.2.2 Μοντέλα Ευπάθειας

Ένας από τους τρόπους περιγραφής της ανομοιογένειας ενός πληθυσμού σε μελέτες ανάλυσης επιβίωσης, είναι μέσω μιας μη αρνητικής τυχαίας μεταβλητής η οποία καλείται μεταβλητή ευπάθειας (frailty variable, Vaupel et al., 1979). Η τελευταία δεν παρατηρείται και αντιπροσωπεύει παράγοντες κινδύνου μετρήσιμους ή μη, οι οποίοι δε συμπεριλαμβάνονται στο μοντέλο, αλλά διαφοροποιούν τις πειραματικές μονάδες ως προς την ικανότητα επιβίωσής τους.

Σύμφωνα με τη θεωρία ευπάθειας, σε κάθε πειραματική μονάδα αντιστοιχεί ο χρόνος ζωής της T_i και η τιμή της μεταβλητής ευπάθειας Z_i . Δοθείσης της τιμής $Z_i = z$, η συνάρτηση κινδύνου της $i^{\text{οστής}}$ πειραματικής μονάδας σε χρόνο t , δίνεται από τη σχέση:

$$h(t | z) = zh(t | z = 1) = zh_1(t), \quad t \geq 0, \quad z \geq 0,$$

όπου η h_1 περιγράφει την επίδραση του χρόνου στην επιβίωση και είναι κοινή για όλες τις



πειραματικές μονάδες (Vaupel et al., 1979).

Όταν σε επίπεδο πειραματικών μονάδων, η επίδραση της μεταβλητής της ευπάθειας στην επιβίωση περιγράφεται από την τελευταία σχέση, η συνάρτηση επιβίωσης του πληθυσμού είναι:

$$S(t) = \int_0^{\infty} \exp\{-zH_1(t)\}f(z)dz = \mathcal{L}\{H_1(t)\}, \quad t \geq 0,$$

όπου $\mathcal{L}(s) = E\{\exp(-sz)\}$ είναι ο μετασχηματισμός Laplace της κατανομής της ευπάθειας και

$H_1(t) = \int_0^{\infty} h_1(t)dt$. Κατά συνέπεια, οι συναρτήσεις πυκνότητας πιθανότητας και κινδύνου του πληθυσμού, είναι αντίστοιχα

$$f(t) = -\mathcal{L}'\{H_1(t)\}, \quad t \geq 0,$$

$$h(t) = -\left[\log \mathcal{L}\{H_1(t)\}\right]', \quad t \geq 0, \quad (1.3)$$

όπου η παραγώγιση είναι ως προς t .

Εμπειρικές αποδείξεις συνηγορούν υπέρ της άποψης ότι καταλληλότερες κατανομές για την περιγραφή της συμπεριφοράς της μεταβλητής της ευπάθειας, είναι οι μονοκόρυφες με θετική λοξότητα (Aalen, 1988). Ανάμεσα σε αυτές, η Γάμα είναι ίσως η δημοφιλέστερη, κυρίως για λόγους μαθηματικής απλότητας. Αν η κατανομή της ευπάθειας είναι η Γάμα, τότε και η δεσμευμένη κατανομή της, δοθείσης επιβίωσης ή δοθέντος 'θανάτου' σε χρόνο t είναι επίσης η Γάμα, αν και η εν λόγω ιδιότητα δε χαρακτηρίζει με μοναδικό τρόπο τη συγκεκριμένη οικογένεια κατανομών. Τέλος, η υπόθεση της γάμα κατανεμημένης ευπάθειας, ισοδυναμεί με την παραδοχή ότι η σχετική ανομοιογένεια του πληθυσμού διατηρείται σταθερή, εν αντιθέσει με την περίπτωση άλλων επιλογών μοντελοποίησής της -όπως για παράδειγμα η κατανομή Inverse Gaussian- που καθιστούν τον πληθυσμό πιο ομοιογενή με την πάροδο του χρόνου (Hougaard, 1984).

1.3 Διδιάστατες κατανομές χρόνων ζωής

Δεδομένα χρόνων ζωής στα οποία η μεταβλητή T είναι διδιάστατη, προκύπτουν σε διάφορες περιπτώσεις, όπως για παράδειγμα σε συστήματα αποτελούμενα από δύο



συνιστώσες (π.χ. δικινητήρια αεροπλάνα), σε καταστάσεις όπου κάθε πειραματική μονάδα είναι εκτεθειμένη σε δύο τύπους αποτυχίας (σειριακά συστήματα με δύο συνιστώσες), ή στη λειτουργία βιολογικών οργάνων (π.χ. μάτια, αυτιά, νεφρά) που εμφανίζονται σε ζεύγη. Βεβαίως, οι συνιστώσες της διδιάστατης τυχαιάς μεταβλητής μπορεί να αναφέρονται σε διαφορετικές πειραματικές μονάδες. Για παράδειγμα, σε μελέτες που αφορούν στην επίδραση της κληρονομικότητας στην επιβίωση, αναφέρονται στους χρόνους ζωής γονέα και παιδιού.

Αν T_1 και T_2 είναι συνεχείς, μη αρνητικές τυχαιές μεταβλητές οι οποίες παριστούν χρόνους ζωής, η από κοινού συνάρτηση πυκνότητας πιθανότητάς τους, δίνεται από τη σχέση:

$$f(t_1, t_2) = \lim_{\substack{\Delta t_1 \rightarrow 0 \\ \Delta t_2 \rightarrow 0}} \frac{P(t_1 \leq T_1 < t_1 + \Delta t_1, t_2 \leq T_2 < t_2 + \Delta t_2)}{\Delta t_1 \Delta t_2}, \quad t_1, t_2 \geq 0$$

ενώ η αντίστοιχη συνάρτηση επιβίωσης είναι

$$S(t_1, t_2) = P(T_1 \geq t_1, T_2 \geq t_2) = \int_{t_1}^{\infty} \int_{t_2}^{\infty} f(t_1, t_2) dt_1 dt_2, \quad t_1, t_2 \geq 0$$

και εκφράζει την πιθανότητα οι χρόνοι ζωής της πρώτης και δεύτερης συνιστώσας του πληθυσμού να υπερβαίνουν αντίστοιχα τις τιμές t_1 και t_2 .

Η συνάρτηση κινδύνου διδιάστατης κατανομής χρόνων ζωής, ορίζεται ως

$$\begin{aligned} h(t_1, t_2) &= \lim_{\substack{\Delta t_1 \rightarrow 0 \\ \Delta t_2 \rightarrow 0}} \frac{P(t_1 \leq T_1 < t_1 + \Delta t_1, t_2 \leq T_2 < t_2 + \Delta t_2 \mid T_1 \geq t_1, T_2 \geq t_2)}{\Delta t_1 \Delta t_2} \\ &= \frac{f(t_1, t_2)}{S(t_1, t_2)}, \quad t_1, t_2 \geq 0. \end{aligned} \quad (1.4)$$

Κατά συνέπεια, για μικρές τιμές των $\Delta t_1, \Delta t_2$, η ποσότητα $h(t_1, t_2) \Delta t_1 \Delta t_2$ εκφράζει κατά προσέγγιση την πιθανότητα 'θανάτου' των δύο συνιστωσών στα διαστήματα $[t_1, t_1 + \Delta t_1)$ και $[t_2, t_2 + \Delta t_2)$ αντίστοιχα, δοθέντος ότι δεν απέτυχαν προηγουμένως. Εν αντιθέσει με τη μονοδιάστατη συνάρτηση κινδύνου, η διδιάστατη δεν προσδιορίζει μονοσήμαντα την κατανομή.

Σημαντικό ρόλο σε μελέτες ανάλυσης της επιβίωσης παίζουν οι ακόλουθες συναρτήσεις κινδύνου:

$$\bar{h}(t_i \mid t_j) = \lim_{\Delta t \rightarrow 0} \frac{P(t_i \leq T_i < t_i + \Delta t \mid T_1 \geq t_1, T_2 \geq t_2)}{\Delta t}$$



$$= -\frac{\partial}{\partial t_i} \log S(t_1, t_2), \quad i=1,2, \quad i \neq j, \quad (1.5)$$

$$\begin{aligned} h(t_i | t_j) &= \lim_{\Delta t \rightarrow 0} \frac{P(t_i \leq T_i < t_i + \Delta t | T_i \geq t_i, T_j = t_j)}{\Delta t} \\ &= -\frac{\partial}{\partial t_i} \log \left\{ -\frac{\partial}{\partial t_j} S(t_i, t_j) \right\}, \quad i, j = 1, 2, \quad i \neq j, \end{aligned} \quad (1.6)$$

οι οποίες προσδιορίζουν πλήρως την από κοινού κατανομή των T_1 και T_2 (Lawless, 1982).

Επίσης, αποδεικνύεται ότι η διδιάστατη συνάρτηση επιβίωσης γράφεται στη μορφή:

$$S(t_1, t_2) = \exp\{-H(t_1, t_2)\},$$

όπου

$$H(t_1, t_2) = \int_0^{t_1} h_1(t_1) dt_1 + \int_0^{t_2} h_2(t_2) dt_2 - \int_0^{t_1} \int_0^{t_2} \{h(t_1, t_2) - \bar{h}(t_1 | t_2) \bar{h}(t_2 | t_1)\} dt_1 dt_2$$

και h_1 και h_2 είναι οι περιθώριες συναρτήσεις κινδύνου που αντιστοιχούν στις κατανομές των T_1 και T_2 (Finkelstein, 2003).

Το πιο κλασικό μέτρο συσχέτισης των τυχαίων μεταβλητών που περιγράφουν ένα διδιάστατο πληθυσμό, είναι ο συντελεστής συσχέτισης του Pearson, ο οποίος ορίζεται ως

$$\rho(T_1, T_2) = \frac{Cov(T_1, T_2)}{\{Var(T_1)Var(T_2)\}^{1/2}}, \quad (1.7)$$

όπου

$$Cov(T_1, T_2) = E\{(T_1 - ET_1)(T_2 - ET_2)\} = \int_0^{\infty} \int_0^{\infty} \{S(t_1, t_2) - S_1(t_1)S_2(t_2)\} dt_1 dt_2. \quad (1.8)$$

Άλλα μέτρα ολικής ή τοπικής συσχέτισης δίνονται στον Hougaard (2000).

Μια από τις πιο δημοφιλείς διδιάστατες κατανομές ανάλυσης επιβίωσης, είναι η διδιάστατη γενίκευση της Εκθετικής κατανομής του Freund (1961), η οποία χρησιμοποιείται κυρίως, στην περιγραφή της λειτουργίας συστημάτων με δύο συνιστώσες, τα οποία μπορούν να λειτουργούν ακόμα και όταν μια από τις συνιστώσες τους πάψει να λειτουργεί. Η σ.π.π. της είναι

$$f(t_1, t_2) = \begin{cases} a\beta' \exp\{-\beta' t_2 - (a + \beta - \beta') t_1\} & 0 < t_1 < t_2 \\ \beta a' \exp\{-a' t_1 - (a + \beta - a') t_2\} & 0 < t_2 < t_1, \end{cases}$$



όπου $a, a', \beta, \beta' > 0$. Οι περιθώριες κατανομές των T_1 και T_2 δίνονται από τις σχέσεις:

$$f(t_1) = \frac{(a - a')(a + \beta) \exp\{-(a + \beta)t_1\}}{a + \beta - a'} + \frac{a' \beta \exp(-a't_1)}{a + \beta - a'}, \quad \text{για } a + \beta - a' \neq 0,$$

$$f(t_2) = \frac{(\beta - \beta')(a + \beta) \exp\{-(a + \beta)t_2\}}{a + \beta - \beta'} + \frac{a \beta' \exp(-\beta't_2)}{a + \beta - \beta'}, \quad \text{για } a + \beta - \beta' \neq 0$$

και

$$f(t_1) = (a + a'\beta t_1) \exp(-a't_1), \quad \text{για } a + \beta - a' = 0,$$

$$f(t_2) = (\beta + \beta'at_2) \exp(-\beta't_2), \quad \text{για } a + \beta - a' = 0.$$

Στην ειδική περίπτωση που $a = a'$ και $\beta = \beta'$, οι T_1 και T_2 είναι ανεξάρτητες και ακολουθούν Εκθετικές κατανομές με παραμέτρους a και β αντίστοιχα.

Επίσης δημοφιλής στην ανάλυση επιβίωσης, είναι η διδιάστατη Εκθετική κατανομή των Marshall και Olkin (1967), με συνάρτηση επιβίωσης

$$S(t_1, t_2) = \exp\{-a_1 t_1 - a_2 t_2 - a_{12} \max(t_1, t_2)\}, \quad t_1, t_2 > 0.$$

Οι περιθώριες κατανομές των T_1 και T_2 είναι Εκθετικές, με παραμέτρους $a_1 + a_{12}$ και $a_2 + a_{12}$ αντίστοιχα.



Κεφάλαιο 2

Η Γενικευμένη Εκθετική-Γεωμετρική Κατανομή

Ένα σημαντικό πρόβλημα στην ανάλυση της επιβίωσης, είναι η επιλογή του κατάλληλου παραμετρικού μοντέλου για την περιγραφή του χρόνου ζωής του υπό εξέταση πληθυσμού. Μια τέτοια επιλογή, γίνεται είτε με γνώμονα τις φυσικές διεργασίες που οδηγούν στην ‘αποτυχία’, είτε με κριτήριο την ποιότητα προσαρμογής του εκάστοτε μοντέλου στα δεδομένα χρόνων ζωής.

Η μαθηματική περιγραφή της πιθανοθεωρητικής συμπεριφοράς συστημάτων (ή οργανισμών) που παρουσιάζουν βελτίωση στη λειτουργία (ή τη βιωσιμότητά τους) με την πάροδο του χρόνου, οδήγησε στη διπαραμετρική Εκθετική-Γεωμετρική κατανομή (Exponential-Geometric ή *EG* για συντομία), με φθίνουσα συνάρτηση κινδύνου (Adamidis and Loukas, 1998). Η τελευταία ανήκει στην οικογένεια κατανομών των Marshall και Olkin (1997), η οποία προτάθηκε ως μέθοδος γενίκευσης κατανομών με προσθήκη παραμέτρου.

Στο κεφάλαιο αυτό, γίνεται μια επέκταση της κατανομής *EG*, χωρίς την προσθήκη επιπλέον παραμέτρων ώστε η προκύπτουσα κατανομή να συμπεριλαμβάνει αύξουσα συνάρτηση κινδύνου. Η τελευταία θα αναφέρεται στη συνέχεια, ως η γενικευμένη Εκθετική-Γεωμετρική κατανομή (Extended Exponential-Geometric κατανομή ή *EEG* για συντομία). Μελετώνται διάφορες μαθηματικές και στατιστικές ιδιότητές της, όπως οι σχέσεις της με άλλες κατανομές, η μονοτονία της σ.π.π. της, οι ροπές της, και οι αντίστοιχες συναρτήσεις επιβίωσης, κινδύνου και μέσου υπολειπόμενου χρόνου ζωής. Το πρόβλημα της εκτίμησης των παραμέτρων, αντιμετωπίζεται με τη μέθοδο της μέγιστης πιθανοφάνειας και τους πίνακες παρατηρούμενης και αναμενόμενης πληροφορίας. Τέλος, παρουσιάζονται χαρακτηρισμοί της



κατανομής, όπως επίσης και συγκριτικά - σε σχέση με τις κατανομές Weibull και Γάμα - αποτελέσματα, από την προσαρμογή της σε πραγματικά δεδομένα.

2.1 Η κατανομή

Η σ.π.π. της γενικευμένης Εκθετικής-Γεωμετρικής κατανομής (*EEG*), δίνεται από τη σχέση:

$$f(t; \theta) = \frac{\beta \gamma e^{-\beta t}}{\{1 - (1 - \gamma)e^{-\beta t}\}^2}, \quad t \geq 0, \quad (2.1)$$

με $\theta = (\beta, \gamma) \in R^+ \times R^+$, όπου οι παράμετροι β και γ , είναι παράμετροι κλίμακας (scale) και μορφής (shape) αντίστοιχα. Περιορίζοντας τον παραμετρικό χώρο στο $R^+ \times (0, 1)$, προκύπτει η κατανομή *EG* ως ειδική περίπτωση.

Η αντίστοιχη συνάρτηση κατανομής, είναι

$$F(t; \theta) = \frac{1 - e^{-\beta t}}{1 - (1 - \gamma)e^{-\beta t}}, \quad t \geq 0. \quad (2.2)$$

Αποδεικνύεται ότι για $0 < \gamma \leq 2$, η σ.π.π. της κατανομής *EEG* είναι γνησίως φθίνουσα με μέγιστο στο 0, την τιμή $\beta \gamma^{-1}$. Για $\gamma > 2$, αυξάνει από την τιμή $\beta \gamma^{-1}$ προς τη μέγιστη τιμή της $\beta \gamma \{4(\gamma - 1)\}^{-1}$ για $t = \beta^{-1} \log(\gamma - 1)$ και στη συνέχεια φθίνει τείνοντας ασυμπτωτικά στο μηδέν. Το είδος μονοτονίας της f διαπιστώνεται εν μέρει από τη συμπεριφορά της παραγώγου της, ως προς t :

$$f'(t; \theta) = -\frac{\beta^2 \gamma e^{-\beta t} \{1 + (1 - \gamma)e^{-\beta t}\}}{\{1 - (1 - \gamma)e^{-\beta t}\}^3}.$$

Για $\gamma > 2$, η $f'(t; \theta) = 0$ έχει μοναδική ρίζα στο $[0, \infty)$ την $t_0 = \beta^{-1} \log(\gamma - 1)$. Όμως,

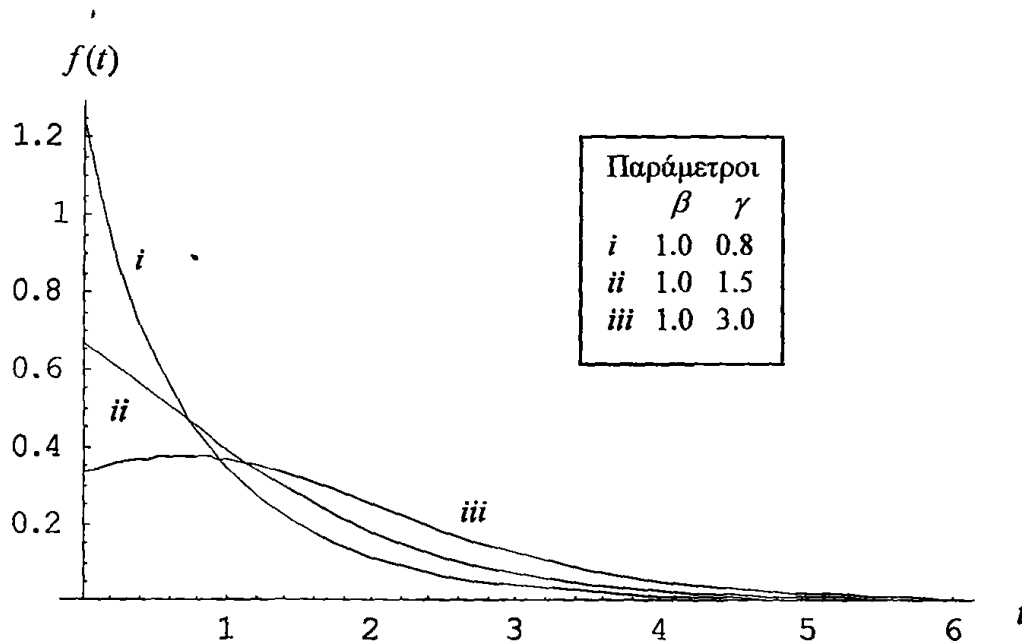
$$f''(t; \theta) = \frac{\beta^3 \gamma e^{-\beta t} \{1 + 4(1 - \gamma)e^{-\beta t} + (1 - \gamma)^2 e^{-2\beta t}\}}{\{1 - (1 - \gamma)e^{-\beta t}\}^4}$$

και αφού

$$f''(t; \theta)|_{t=t_0} = -\frac{\beta^3 \gamma}{8(\gamma - 1)} < 0,$$



η f είναι μονοκόρυφη, με μέγιστη τιμή στο t_0 αυτή που προαναφέρθηκε. Για $0 < \gamma \leq 2$, είναι $f'(t; \theta) < 0$ για κάθε $t \geq 0$ και άρα, η f είναι φθίνουσα συνάρτηση του t . Στο Σχήμα 2.1 απεικονίζονται πυκνότητες της κατανομής EEG για επιλεγμένες τιμές των παραμέτρων β και γ .



Σχήμα 2.1 Πυκνότητες της κατανομής EEG .

2.2. Φυσική ερμηνεία

Οι φυσικές διεργασίες παραγωγής της κατανομής EEG περιγράφονται από τα ακόλουθα δύο θεωρήματα.

Θεώρημα 2.1

Έστω ότι η τυχαία μεταβλητή $T | A$ ακολουθεί την τροποποιημένη Extreme Value κατανομή με σ.π.π.

$$f(t | \alpha; \beta) = \alpha \beta \exp\{\beta t + \alpha(1 - e^{\beta t})\}, \quad t \geq 0,$$

όπου $\alpha, \beta \in R^+$. Αν η τυχαία μεταβλητή A ακολουθεί Εκθετική κατανομή με πυκνότητα

$$f(\alpha; \gamma) = \gamma e^{-\gamma \alpha}, \quad \alpha \geq 0,$$

όπου $\gamma \in R^+$, τότε η T ακολουθεί την κατανομή EEG .



Απόδειξη

Από τις υποθέσεις του θεωρήματος, έπεται ότι η από κοινού κατανομή των T και A δίνεται από τη σχέση:

$$f(t, a; \beta, \gamma) = a\beta\gamma \exp\{\beta t + a(1 - \gamma - e^{\beta t})\}.$$

Κατά συνέπεια, η περιθώρια κατανομή της T υπολογίζεται από το ολοκλήρωμα

$$\begin{aligned} f(t; \theta) &= \beta\gamma e^{\beta t} \int_0^{\infty} a \exp\{a(1 - \gamma - e^{\beta t})\} da \\ &= \frac{\beta\gamma e^{\beta t}}{1 - \gamma - e^{\beta t}} \int_0^{\infty} ad \left[\exp\{a(1 - \gamma - e^{\beta t})\} \right] \\ &= \frac{\beta\gamma e^{\beta t}}{1 - \gamma - e^{\beta t}} \left\{ a \exp\{a(1 - \gamma - e^{\beta t})\} \Big|_0^{\infty} - \int_0^{\infty} \exp\{a(1 - \gamma - e^{\beta t})\} da \right\} \\ &= \frac{\beta\gamma e^{\beta t}}{1 - \gamma - e^{\beta t}} \left\{ a \exp\{a(1 - \gamma - e^{\beta t})\} - \frac{\exp\{a(1 - \gamma - e^{\beta t})\}}{1 - \gamma - e^{\beta t}} \right\} \Big|_0^{\infty}. \end{aligned}$$

Με χρήση του κανόνα του L'Hospital προκύπτει ότι

$$\lim_{a \rightarrow \infty} a \exp\{a(1 - \gamma - e^{\beta t})\} = 0$$

και αποδεικνύεται το ζητούμενο.

Κατά συνέπεια, ο χρόνος μέχρι την 'αποτυχία' μοντελοποιείται με βάση την κατανομή EEG , όταν ο υπό εξέταση πληθυσμός ακολουθεί την τροποποιημένη Extreme Value κατανομή και είναι ανομοιογενής ως προς μια συνεχή παράμετρο του, η οποία συμπεριφέρεται ως εκθετική τυχαία μεταβλητή.

Θεώρημα 2.2

Έστω $\{T_i\}_{i=1}^N$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με πυκνότητα

$$f(t; \beta) = \beta e^{-\beta t}, \quad t \geq 0,$$

όπου $\beta \in R^+$ και N γεωμετρική τυχαία μεταβλητή με συνάρτηση πιθανότητας

$$f(n; p) = (1 - p)p^{n-1}, \quad n \in N$$

με $p \in (0, 1)$. Τότε:



α) η τυχαία μεταβλητή $Z = \min\{T_i\}_{i=1}^N$ ακολουθεί την κατανομή EEG με $\gamma = 1 - p$,

β) η τυχαία μεταβλητή $W = \max\{T_i\}_{i=1}^N$ ακολουθεί την κατανομή EEG με $\gamma = (1 - p)^{-1}$.

Απόδειξη

α) Από τις υποθέσεις του θεωρήματος προκύπτει ότι η δεσμευμένη κατανομή της τυχαίας μεταβλητής $Z | N$ είναι Εκθετική με παράμετρο $n\beta$. Κατά συνέπεια, η από κοινού κατανομή των Z και N δίνεται από τη σχέση:

$$f(z, n; \beta, p) = n\beta(1 - p)p^{n-1}e^{-n\beta z}.$$

Άρα,

$$\begin{aligned} f(z; \beta, p) &= \beta(1 - p)e^{-\beta z} \sum_{n=1}^{\infty} n(pe^{-\beta z})^{n-1} \\ &= \beta(1 - p)e^{-\beta z} \left(\sum_{n=0}^{\infty} x^n \right)' \Bigg|_{x=pe^{-\beta z}} \\ &= \frac{\beta(1 - p)e^{-\beta z}}{(1 - pe^{-\beta z})^2}. \end{aligned}$$

Συνεπώς, η Z ακολουθεί την κατανομή EEG με $\gamma = 1 - p$.

β) Η δεσμευμένη κατανομή της τυχαίας μεταβλητής $W | N$ δίνεται από τη σχέση:

$$f(w | n; \beta) = n\beta e^{-\beta w} (1 - e^{-\beta w})^{n-1}.$$

Κατά συνέπεια, η από κοινού κατανομή των W και N , είναι

$$f(w, n; \beta, p) = n\beta(1 - p)p^{n-1}e^{-\beta w} (1 - e^{-\beta w})^{n-1}.$$

Άρα,

$$\begin{aligned} f(w; \beta, p) &= \beta(1 - p)e^{-\beta w} \sum_{n=1}^{\infty} n\{p(1 - e^{-\beta w})\}^{n-1} \\ &= \beta(1 - p)e^{-\beta w} \left(\sum_{n=0}^{\infty} x^n \right)' \Bigg|_{x=p(1 - e^{-\beta w})} \\ &= \frac{\beta(1 - p)^{-1}e^{-\beta w}}{\{1 + p(1 - p)^{-1}e^{-\beta w}\}^2}. \end{aligned}$$



Επομένως, η W ακολουθεί την κατανομή EEG με $\gamma = (1 - p)^{-1}$.

Με βάση το τελευταίο θεώρημα, η φυσική ερμηνεία του προτεινόμενου μοντέλου προκύπτει μέσα από τη θεώρηση μιας κατάστασης όπου η ‘αποτυχία’ (π.χ. μιας συσκευής), οφείλεται στην παρουσία άγνωστου αριθμού N αρχικών ελαττωμάτων ίδιου τύπου (π.χ. κατασκευαστικά σφάλματα εξαρτημάτων από την ίδια παρτίδα παραγωγής). Στα πλαίσια αυτά, T_i , $i = 1, \dots, N$, είναι οι χρόνοι ζωής των επιμέρους στοιχείων-εξαρτημάτων που φέρουν ελάττωμα. Αν τα ελαττώματα ανιχνεύονται μόνο εφόσον προκαλέσουν δυσλειτουργία στο ‘σύστημα’, οπότε και επιδιορθώνονται πλήρως και κάτω από τις υποθέσεις του θεωρήματος για τις κατανομές των T_i και N , η κατανομή του χρόνου ζωής του ‘συστήματος’ είναι εκείνη της τυχαίας μεταβλητής Z . Αξίζει να σημειωθεί ότι η διαδικασία της πλήρους επιδιόρθωσης έχει ως αποτέλεσμα το ‘σύστημα’ να γίνεται προοδευτικά ανθεκτικότερο και επομένως η διάρκεια ζωής του να περιγράφεται από την κατανομή EEG με φθίνουσα συνάρτηση κινδύνου ($0 < \gamma < 1$). Σε όρους μηχανολογίας, η παραπάνω κατάσταση μπορεί να προκύψει σε συστήματα των οποίων τα επιμέρους εξαρτήματα έχουν συνδεσμολογία σειράς. Αντιθέτως, αν το ‘σύστημα’ αποτυγχάνει τη στιγμή ακριβώς που και το τελευταίο ελάττωμα ανιχνεύεται και κάτω από τις υποθέσεις του θεωρήματος, η κατανομή του χρόνου ζωής του είναι εκείνη της τυχαίας μεταβλητής W . Σε μια τέτοια περίπτωση ο χρόνος λειτουργεί επιβαρυντικά στη διάρκεια ζωής του συστήματος, η οποία κατά συνέπεια περιγράφεται από την κατανομή EEG με αύξουσα συνάρτηση κινδύνου ($\gamma > 1$). Σε όρους μηχανολογίας, μια τέτοια κατάσταση προκύπτει σε συστήματα των οποίων τα επί μέρους εξαρτήματα έχουν παράλληλη σύνδεση.

2.3 Σχέσεις με άλλες κατανομές

Η κατανομή EEG περιλαμβάνει ως ειδικές περιπτώσεις την Εκθετική κατανομή για $\gamma = 1$ και την περικομμένη στο R^+ Λογιστική κατανομή ($\text{logistic}(\lambda, \mu)$, Johnson et al., 1995), με $\lambda = \beta^{-1}$ και $\mu = \beta^{-1} \log(\gamma - 1)$ για $\gamma > 1$. Επίσης, για τιμές της παραμέτρου γ στο διάστημα $(0, 1)$, ταυτίζεται με την κατανομή EG , με $p = 1 - \gamma$.

Ακόμα, αποδεικνύεται ότι αν $T \sim EEG$, τότε:



i) η τυχαία μεταβλητή $Y = |1 - \gamma|^{-1} (e^{\beta T} - 1)$ ακολουθεί κατανομή Pareto τύπου II (γνωστή και ως κατανομή του Lomax), με παραμέτρους μορφής και κλίμακας 1 και $\gamma^{-1}|1 - \gamma|$, αντίστοιχα,

ii) η τυχαία μεταβλητή $Z = \{1 - (1 - \gamma)e^{-\beta T}\}^{-1}$ ακολουθεί ομοιόμορφη κατανομή $U(\min(1, \gamma^{-1}), \max(1, \gamma^{-1}))$.

Τα τελευταία αποτελέσματα προκύπτουν χρησιμοποιώντας κάποια από τις μεθόδους αλλαγής μεταβλητών.

2.4 Ροπές

Οι ροπές r τάξης περί το μηδέν της κατανομής EEG , υπολογίζονται από το ολοκλήρωμα:

$$\begin{aligned} \mu_r &= E(T^r; \theta) = \int_0^{\infty} \frac{t^r \beta \gamma e^{-\beta t}}{\{1 - (1 - \gamma)e^{-\beta t}\}^2} dt \\ &= \frac{\gamma}{\beta^r} \int_0^{\infty} \frac{x^r e^{-x}}{\{1 - (1 - \gamma)e^{-x}\}^2} dx \\ &= \frac{-\gamma}{(1 - \gamma)\beta^r} \int_0^{\infty} x^r d\left\{ \frac{1}{1 - (1 - \gamma)e^{-x}} \right\} \\ &= \frac{-\gamma}{(1 - \gamma)\beta^r} \cdot \frac{x^r}{1 - (1 - \gamma)e^{-x}} \Big|_0^{\infty} + \frac{\gamma}{(1 - \gamma)\beta^r} \int_0^{\infty} \frac{rx^{r-1}}{1 - (1 - \gamma)e^{-x}} dx \\ &= \frac{-\gamma}{(1 - \gamma)\beta^r} \cdot \frac{x^r}{1 - (1 - \gamma)e^{-x}} \Big|_0^{\infty} + \frac{r\gamma}{(1 - \gamma)\beta^r} \int_0^{\infty} x^{r-1} dx + \frac{r\gamma}{\beta^r} \int_0^{\infty} \frac{x^{r-1} e^{-x}}{1 - (1 - \gamma)e^{-x}} dx \\ &= \frac{-\gamma}{(1 - \gamma)\beta^r} \cdot x^r \left\{ \frac{1}{1 - (1 - \gamma)e^{-x}} - 1 \right\} \Big|_0^{\infty} + \frac{r\gamma}{\beta^r} \int_0^{\infty} \frac{x^{r-1} e^{-x}}{1 - (1 - \gamma)e^{-x}} dx \\ &= \frac{-\gamma}{\beta^r} \cdot \frac{x^r e^{-x}}{1 - (1 - \gamma)e^{-x}} \Big|_0^{\infty} + \frac{r\gamma}{\beta^r} \int_0^{\infty} \frac{x^{r-1} e^{-x}}{1 - (1 - \gamma)e^{-x}} dx. \end{aligned}$$

Εφαρμόζοντας τον κανόνα του L'Hospital, προκύπτει ότι



$$\lim_{x \rightarrow \infty} \frac{x^r e^{-x}}{1 - (1-\gamma)e^{-x}} = \lim_{x \rightarrow \infty} \frac{x^r}{e^x - (1-\gamma)} \quad (\text{μορφή } \frac{\infty}{\infty}) = \dots = \lim_{x \rightarrow \infty} \frac{r!}{e^x} = 0.$$

Άρα,

$$E(T^r; \theta) = \frac{r! \gamma}{\beta^r \Gamma(r)} \int_0^{\infty} \frac{x^{r-1} e^{-x}}{1 - (1-\gamma)e^{-x}} dx = \frac{r! \gamma}{\beta^r} \Phi(1-\gamma, r, 1),$$

όπου

$$\Phi(z, s, \alpha) = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{t^{s-1} e^{-\alpha t}}{1 - z e^{-t}} dt,$$

με $\alpha > 0$, $z < 1$ και $s > 0$ είναι η συνάρτηση του Lerch (Lerch's transcendent function, Magnus et al., 1966, σελ. 32). Η τελευταία είναι διαθέσιμη στα κλασικά μαθηματικά-στατιστικά λογισμικά όπως Mathematica ή Maple.

Κατ'επέκταση, η μέση τιμή της κατανομής είναι

$$\begin{aligned} E(T; \theta) &= \frac{\gamma}{\beta} \Phi(1-\gamma, 1, 1) \\ &= \frac{\gamma}{\beta} \int_0^{\infty} \frac{e^{-t}}{1 - (1-\gamma)e^{-t}} dt \\ &= \frac{\gamma}{\beta(1-\gamma)} \int_0^{\infty} d[\log\{1 - (1-\gamma)e^{-t}\}] \\ &= \frac{-\gamma \log \gamma}{\beta(1-\gamma)} \end{aligned} \tag{2.3}$$

ενώ η διακύμανσή της, δίνεται από τη σχέση:

$$\text{Var}(T; \theta) = \frac{\gamma}{\beta^2} \left\{ 2\Phi(1-\gamma, 2, 1) - \frac{\gamma \log^2 \gamma}{(1-\gamma)^2} \right\}. \tag{2.4}$$

2.5 Συναρτήσεις επιβίωσης και κινδύνου

Η πιθανότητα επιβίωσης πέραν του χρόνου t , γνωστή και ως συνάρτηση επιβίωσης, έχει τη μορφή:

$$S(t; \theta) = \frac{\gamma e^{-\beta t}}{1 - (1-\gamma)e^{-\beta t}}, \quad t \geq 0,$$



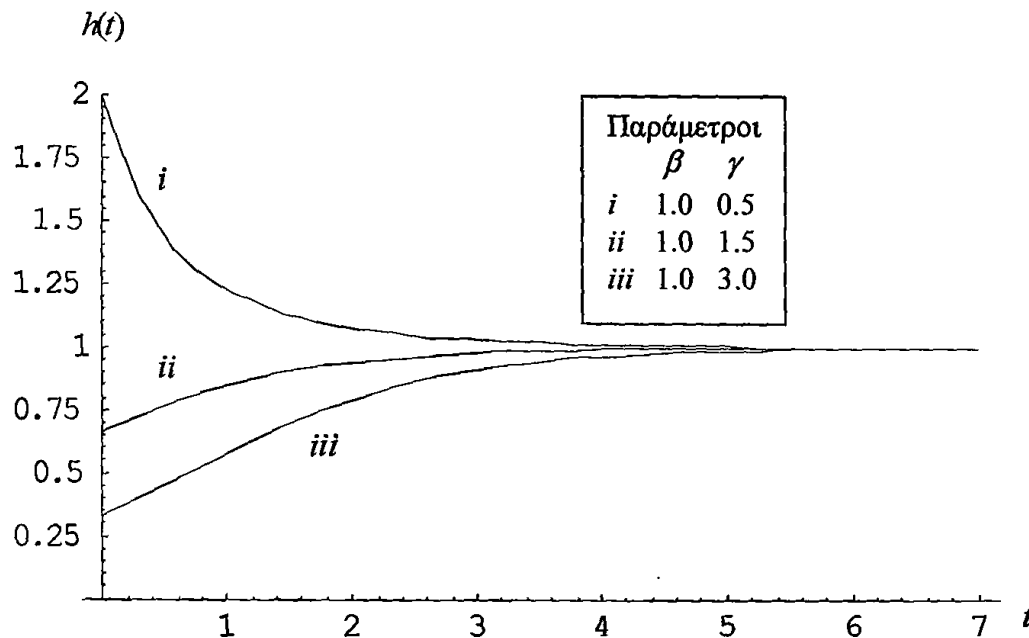
ενώ από τη (2.1) και την τελευταία σχέση, η συνάρτηση κινδύνου της κατανομής *EEG* υπολογίζεται ως

$$h(t; \theta) = \frac{\beta}{1 - (1 - \gamma)e^{-\beta t}}, \quad t \geq 0.$$

Εξετάζοντας το πρόσημο της παραγώγου της *h* ως προς *t*

$$h'(t; \theta) = \frac{\beta^2(\gamma - 1)e^{-\beta t}}{\{1 - (1 - \gamma)e^{-\beta t}\}^2}, \tag{2.5}$$

διαπιστώνεται ότι η συνάρτηση κινδύνου είναι γνησίως φθίνουσα για $0 < \gamma < 1$, γνησίως αύξουσα για $\gamma > 1$ και σταθερή για $\gamma = 1$ (Εκθετική). Σε κάθε περίπτωση, ξεκινώντας από την τιμή $\beta\gamma^{-1}$ για $t = 0$, τείνει στην τιμή β . Στο Σχήμα 2.2 απεικονίζονται συναρτήσεις κινδύνου της κατανομής *EEG* για επιλεγμένες τιμές των παραμέτρων της.



Σχήμα 2.2 Συναρτήσεις κινδύνου της κατανομής *EEG*.

2.6 Μέσος υπολειπόμενος χρόνος ζωής

Δοθείσης της επιβίωσης μέχρι τη χρονική στιγμή *t*, ο μέσος υπολειπόμενος χρόνος ζωής για την κατανομή *EEG* υπολογίζεται ως εξής:

$$m(t; \theta) = \frac{1}{S(t; \theta)} \int_t^{\infty} \frac{\gamma e^{-\beta t}}{1 - (1 - \gamma)e^{-\beta t}} dt$$



$$\begin{aligned}
&= \frac{\gamma}{\beta(1-\gamma)S(t;\theta)} \int_t^{\infty} d\left[\log\left\{1-(1-\gamma)e^{-\beta t}\right\}\right] \\
&= \frac{-\gamma \log\left\{1-(1-\gamma)e^{-\beta t}\right\}}{\beta(1-\gamma)S(t;\theta)} \\
&= -\frac{e^{\beta t}}{\beta(1-\gamma)} \left\{1-(1-\gamma)e^{-\beta t}\right\} \log\left\{1-(1-\gamma)e^{-\beta t}\right\}.
\end{aligned}$$

Η τελευταία συνάρτηση είναι γνησίως αύξουσα για $0 < \gamma < 1$, γνησίως φθίνουσα για $\gamma > 1$ και σταθερή για $\gamma = 1$ (Εκθετική κατανομή) αφού όταν η συνάρτηση κινδύνου είναι γνησίως μονότονη, η συνάρτηση του μέσου υπολειπόμενου χρόνου ζωής παρουσιάζει το αντίθετο είδος μονοτονίας ενώ είναι ανεξάρτητη του χρόνου στην περίπτωση της Εκθετικής κατανομής. Η τιμή της $m(t;\theta)$ στο μηδέν, ταυτίζεται με εκείνη του μέσου χρόνου ζωής ενώ η οριακή τιμή της για $t \rightarrow \infty$ είναι β^{-1} και αποκτάται εφαρμόζοντας τον κανόνα του L'Hospital σύμφωνα με την ακόλουθη διαδικασία:

$$\lim_{t \rightarrow \infty} e^{\beta t} \log\left\{1-(1-\gamma)e^{-\beta t}\right\} = \lim_{t \rightarrow \infty} \frac{\log\left\{1-(1-\gamma)e^{-\beta t}\right\}}{e^{-\beta t}} = \lim_{t \rightarrow \infty} \frac{-(1-\gamma)}{1-(1-\gamma)e^{-\beta t}} = \gamma - 1.$$

Άρα,

$$\lim_{t \rightarrow \infty} m(t;\theta) = \lim_{t \rightarrow \infty} \frac{1-(1-\gamma)e^{-\beta t}}{\beta(\gamma-1)} \cdot \lim_{t \rightarrow \infty} e^{\beta t} \log\left\{1-(1-\gamma)e^{-\beta t}\right\} = \frac{1}{\beta}.$$

Από το είδος μονοτονίας της συνάρτησης $m(t;\theta)$ και την τιμή της στο μηδέν, έπεται ότι για $0 < \gamma < 1$, ο μέσος υπολειπόμενος χρόνος ζωής υπερβαίνει το μέσο χρόνο ζωής για κάθε t ενώ υπολείπεται του δευτέρου για $\gamma > 1$. Τέλος, για $\gamma = 1$ οι δύο αναμενόμενοι χρόνοι ζωής ταυτίζονται, λαμβάνοντας την τιμή β^{-1} .

2.7 Συμπερασματολογία

2.7.1 Εκτιμητική με τη μέθοδο της μέγιστης πιθανοφάνειας

Για τυχαίο δείγμα μεγέθους n , $y_{obs} = (t_i; i=1, \dots, n)$ από την (2.1), η συνάρτηση πιθανοφάνειας και ο λογάριθμός της, δίνονται αντίστοιχα από τις σχέσεις:

$$L = L(\theta; y_{obs}) = (\beta\gamma)^n \exp(-\beta \sum_{i=1}^n t_i) \prod_{i=1}^n \frac{1}{\left\{1-(1-\gamma)e^{-\beta t_i}\right\}^2},$$



$$\ell = \ell(\theta; y_{obs}) = \log L = n \log(\beta\gamma) - \beta \sum_{i=1}^n t_i - 2 \sum_{i=1}^n \log \{1 - (1-\gamma)e^{-\beta t_i}\}.$$

Θέτοντας τις μερικές παραγώγους $\partial \ell / \partial \theta_i$ $i = 1, 2$ ίσες με το μηδέν, προκύπτουν οι εξισώσεις μέγιστης πιθανοφάνειας

$$\begin{aligned} \frac{\partial \ell}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^n t_i - 2 \sum_{i=1}^n \frac{t_i}{1 - (1-\gamma)e^{-\beta t_i}} = 0, \\ \frac{\partial \ell}{\partial \gamma} &\propto \frac{n(1+\gamma)}{\gamma} - 2 \sum_{i=1}^n \frac{1}{1 - (1-\gamma)e^{-\beta t_i}} = 0. \end{aligned} \tag{2.6}$$

Είναι προφανές ότι για την επίλυση του συστήματος των παραπάνω μη γραμμικών ως προς τις παραμέτρους εξισώσεων, θα πρέπει να χρησιμοποιηθεί κάποια από τις μεθόδους αριθμητικής ανάλυσης όπως για παράδειγμα η μέθοδος Newton-Raphson ή η μέθοδος scoring.

2.7.2 Η μέθοδος EM

Στα πλαίσια των αριθμητικών μεθόδων με δεδομένη σύγκλιση που αναπτύχθηκαν για τη διαχείριση στατιστικών προβλημάτων, οι Adamidis και Loukas (1998) προτείνουν έναν EM αλγόριθμο (Dempster et al., 1977) για την εκτίμηση του θ όταν $\gamma \in (0,1)$. Δυστυχώς, για $\gamma > 1$, είναι δύσκολο να κατασκευάσει κανείς το υποθετικό δείγμα προκειμένου να εφαρμόσει τη μέθοδο, αποφεύγοντας την εύρεση μεγίστου στο βήμα μεγιστοποίησης (maximization step ή M-step) σε κάθε κλήση της μεθόδου. Εντούτοις, όπως φαίνεται στη συνέχεια η μέθοδος EM εξακολουθεί να είναι ανταγωνιστική.

Η εφαρμογή της μεθόδου διευκολύνεται κάτω από τη θεώρηση ότι το ‘πλήρες’ δείγμα αποτελείται από τις παρατηρούμενες τιμές της τυχαίας μεταβλητής T και τις τιμές μιας μη παρατηρούμενης τυχαίας μεταβλητής A , με από κοινού κατανομή:

$$f(t, a; \theta) = a\beta\gamma e^{\beta t + a(1-\gamma - e^{-\beta t})}, \quad t, a \geq 0,$$

όπου $\beta, \gamma \in R^+$. Αν $\ell_c = \ell(\theta; Y_c)$ είναι ο λογάριθμος της πιθανοφάνειας των υποθετικών δεδομένων $Y_c = (T_i, A_i; i = 1, \dots, n)$ με την παραπάνω πυκνότητα, δοθείσης μιας τιμής του θ , $\theta^{(t)} = (\beta^{(t)}, \gamma^{(t)})$, το βήμα αναμενόμενης τιμής (expectation step ή E-step) μιας κλήσης της μεθόδου EM, απαιτεί τον υπολογισμό



$$E(l_c) \propto n \log(\beta\gamma) + \beta \sum_{i=1}^n t_i + (1-\gamma) \sum_{i=1}^n E(A_i) - \sum_{i=1}^n e^{\beta t_i} E(A_i), \quad (2.7)$$

όπου οι αναμενόμενες τιμές είναι υπό συνθήκη δοθέντων των παρατηρούμενων τιμών t_i $i = 1, \dots, n$ και μιας τιμής της παραμέτρου θ , $\theta^{(t)}$. Αποδεικνύεται ότι δεσμευμένη κατανομή της τυχαίας μεταβλητής $A|T$, έχει πυκνότητα

$$f(a|t; \theta) = a \{e^{\beta t} + (1-\gamma)\}^2 e^{a(1-\gamma-e^{\beta t})}, \quad a \geq 0,$$

όπου $t, \beta, \gamma \in R^+$ και κατά συνέπεια, η αναμενόμενη τιμή της, είναι

$$E(A|T; \theta) = \frac{2e^{-\beta t}}{1 - (1-\gamma)e^{-\beta t}}.$$

Στο βήμα μεγιστοποίησης, η ποσότητα στη σχέση (2.7) μεγιστοποιείται ως προς θ , αντιμετωπίζοντας τις δεσμευμένες αναμενόμενες τιμές ως δεδομένα και τελικά, μια κλήση της μεθόδου EM ορίζεται από τις σχέσεις:

$$\frac{n}{\beta^{-(t+1)}} = -\sum_{i=1}^n t_i + 2 \sum_{i=1}^n \left[\frac{t_i e^{\{\beta^{(t+1)} - \beta^{(t)}\} t_i}}{1 - \{1 - \gamma^{(t)}\} e^{-\beta^{(t)} t_i}} \right], \quad (2.8)$$

$$\frac{n}{\gamma^{-(t+1)}} = 2 \sum_{i=1}^n \left[\frac{e^{-\beta^{(t)} t_i}}{1 - \{1 - \gamma^{(t)}\} e^{-\beta^{(t)} t_i}} \right].$$

Το θεώρημα που ακολουθεί εγγυάται την ύπαρξη και την μοναδικότητα της λύσης σε κάθε κλήση της μεθόδου και αφορά την περίπτωση που $\gamma \geq 1$. Ανάλογα αποτελέσματα για την περίπτωση $\gamma \in (0, 1)$, δίνονται από τους Adamidis και Loukas (1998).

Θεώρημα 2.3

Έστω $g(\beta; \gamma)$ η πρώτη από τις εξισώσεις μέγιστης πιθανοφάνειας στο σύστημα (2.6) και

$\bar{t} = n^{-1} \sum_{i=1}^n t_i$. Τότε, δοθείσης της τιμής του γ στο διάστημα $(1, \infty)$, η εξίσωση $g(\beta; \gamma) = 0$ έχει

μοναδική λύση στο διάστημα $(1/\bar{t}, \infty)$.

Απόδειξη

Για $\gamma > 1$, είναι



$$\lim_{\beta \rightarrow \infty} g(\beta; \gamma) = -\sum_{i=1}^n t_i < 0,$$

$$\lim_{\beta \rightarrow 1/\bar{t}} g(\beta; \gamma) = 2 \sum_{i=1}^n t_i \left[1 - \{1 - (1 - \gamma) \exp(-t_i / \bar{t})\}^{-1} \right] > 0.$$

Άρα, η εξίσωση $g(\beta; \gamma) = 0$ έχει μια τουλάχιστον λύση στο διάστημα $(1/\bar{t}, \infty)$ και αφού

$$\frac{\partial g(\beta; \gamma)}{\partial \beta} = -n\beta^{-2} + 2(1 - \gamma) \sum_{i=1}^n t_i^2 e^{-\beta t_i} \left\{ 1 - (1 - \gamma) e^{-\beta t_i} \right\}^{-2} < 0,$$

η λύση είναι μοναδική.

Επίσης, μπορεί να αποδειχθεί ότι για $\gamma \in (1, 2)$ η ρίζα της εξίσωσης $g(\beta; \gamma) = 0$, βρίσκεται στο διάστημα μεταξύ \bar{t}^{-1} και $\left\{ (2\gamma^{-1} - 1)\bar{t} \right\}^{-1}$. Η απόδειξη είναι ανάλογη εκείνης του θεωρήματος 4.1 των Adamidis και Loukas (1998), για $\gamma \in (0, 1)$.

Η ταχύτητα σύγκλισης του προτεινόμενου EM αλγόριθμου διερευνήθηκε με προσομοίωση. Χρησιμοποιώντας το Θεώρημα 2.2 και ρουτίνες της NAG (1991) για τη γέννηση τυχαίων αριθμών από την ομοιόμορφη κατανομή, παράχθηκαν 200 τυχαία δείγματα μεγέθους 200 από την κατανομή *EEG* για καθένα από εννέα ζεύγη τιμών (β, γ) . Για κάθε δείγμα υπολογίστηκαν οι τιμές των εκτιμητών από τις σχέσεις (2.8), χρησιμοποιώντας ως αρχικές τιμές τους εκτιμητές που λαμβάνονται με τη μέθοδο των ροπών. Δεν τέθηκε περιορισμός όσον αφορά στο μέγιστο αριθμό κλήσεων της μεθόδου ενώ θεωρήθηκε ότι η σύγκλιση επιτεύχθηκε όταν η απόλυτη τιμή της διαφοράς μεταξύ διαδοχικών εκτιμητών ήταν μικρότερη του 10^{-5} . Τα αποτελέσματα από τα 1800 τυχαία δείγματα παρουσιάζονται στον Πίνακα 2.1, στον οποίο δίνονται οι μέσες τιμές ($av(\hat{\theta})$) και οι τυπικές αποκλίσεις ($se(\hat{\theta})$) των 200 εκτιμητών και επίσης, ο μέσος αριθμός κλήσεων της μεθόδου EM ($av(t)$) και η τυπική απόκλισή του ($se(t)$).

Αξίζει να σημειωθεί ότι η σύγκλιση επιτεύχθηκε ακόμα και στις περιπτώσεις που οι αρχικές τιμές που χρησιμοποιήθηκαν απείχαν σημαντικά από τις πραγματικές τιμές των παραμέτρων, γεγονός που ενισχύει την αξία της μεθόδου EM. Επίσης, για το εύρος τιμών των παραμέτρων που εξετάστηκε και αναφορικά με την ταχύτητα σύγκλισης της μεθόδου, φαίνεται ότι καθώς οι τιμές των β και γ αυξάνουν, τόσο ο μέσος αριθμός των



$\theta = (\beta, \gamma)$	$av(\hat{\theta})$	$se(\hat{\theta})$	$av(t)$	$se(t)$
(0.5, 0.25)	(0.536, 0.278)	(0.136, 0.090)	171.76	56.38
(1, 0.5)	(1.051, 0.541)	0.192, 0.144)	151.98	40.58
(1.5, 1)	(1.548, 1.070)	(0.208, 0.245)	138.80	34.25
(2, 1.5)	(2.028, 1.570)	(0.233, 0.340)	144.23	29.08
(2.5, 2)	(2.559, 2.126)	(0.284, 0.486)	149.13	28.83
(3, 2.5)	(3.046, 2.616)	(0.315, 0.554)	157.23	27.23
(3.5, 3)	(3.528, 3.126)	(0.376, 0.796)	163.57	27.99
(4, 3.5)	(4.036, 3.622)	(0.385, 0.864)	163.49	31.54
(5, 4)	(5.007, 4.082)	(0.478, 0.994)	168.71	34.13

Πίνακας 2.1 Μέσες τιμές και τυπικές αποκλίσεις των εκτιμητών και του αριθμού κλήσεων της μεθόδου EM, από τυχαία δείγματα μεγέθους 200 από την κατανομή EEG.

απαιτούμενων κλήσεων t όσο και η διασπορά στις τιμές του, αρχικά μειώνονται και στη συνέχεια αυξάνουν. Η τάση αύξησης παρουσιάζεται καθώς τα β και γ ξεπερνούν αντίστοιχα τις τιμές 1.5 και 1 στην πρώτη περίπτωση και 3 και 2.5 στη δεύτερη. Επιπλέον, ο αριθμός των απαιτούμενων κλήσεων t , δε φαίνεται απαγορευτικός ως προς τη χρήση της μεθόδου EM. Βεβαίως, μπορούν να χρησιμοποιηθούν διάφορες τεχνικές επιτάχυνσής της, ενδεχομένως εις βάρος της απλότητας, της σταθερότητας και της ευκολίας προγραμματισμού της.

2.7.3 Ασυμπτωτική κατανομή των εκτιμητών μέγιστης πιθανοφάνειας

Σύμφωνα με την ασυμπτωτική θεωρία κανονικότητας, η κατανομή των εκτιμητών μέγιστης πιθανοφάνειας $\hat{\theta}$, είναι κατά προσέγγιση η διδιάστατη κανονική με μέση τιμή θ και πίνακα διακυμάνσεων-συνδιακυμάνσεων τον αντίστροφο του πίνακα αναμενόμενης πληροφορίας $J(\theta) = E(I; \theta)$. Στην τελευταία σχέση, $I = I(\theta; y_{obs})$ είναι ο συμμετρικός, δεύτερης τάξης πίνακας παρατηρούμενης πληροφορίας ενώ η αναμενόμενη τιμή λαμβάνεται ως προς την κατανομή της τυχαίας μεταβλητής T . Τα στοιχεία του πίνακα $I(\theta; y_{obs})$, $I_{ij} = -\partial^2 \ell / \partial \theta_i \partial \theta_j$, $i, j = 1, 2$, δίνονται αναλυτικά από τις σχέσεις:



$$\begin{aligned}
 I_{11} &= \frac{n}{\beta^2} - 2(1-\gamma) \sum_{i=1}^n \frac{t_i^2 e^{-\beta t_i}}{\{1-(1-\gamma)e^{-\beta t_i}\}^2}, \\
 I_{22} &= \frac{n}{\gamma^2} - 2 \sum_{i=1}^n \frac{e^{-2\beta t_i}}{\{1-(1-\gamma)e^{-\beta t_i}\}^2}, \\
 I_{12} &= I_{21} = -2 \sum_{i=1}^n \frac{t_i e^{-\beta t_i}}{\{1-(1-\gamma)e^{-\beta t_i}\}^2}.
 \end{aligned} \tag{2.9}$$

Το πρόβλημα υπολογισμού του πίνακα αναμενόμενης πληροφορίας για $0 < \gamma < 1$, αντιμετωπίζεται από τους Adamidis και Loukas (1998). Οι λεπτομέρειες της διαχείρισης του αντίστοιχου προβλήματος για $\gamma > 1$, παρουσιάζονται στο ακόλουθο θεώρημα.

Θεώρημα 2.4

Για τυχαίο δείγμα μεγέθους n από την (2.1) και για $\gamma > 1$, τα στοιχεία του συμμετρικού δεύτερης τάξης πίνακα αναμενόμενης πληροφορίας $J(\theta)$, δίνονται από τις σχέσεις:

$$J_{11} = \frac{n}{3\beta^2(1-\gamma)} \left\{ 1 - \gamma - \frac{\gamma\pi^2}{3} + 2\gamma L\left(\frac{1}{\gamma}, 2\right) - \gamma \log^2(\gamma - 1) + \gamma \log^2\left(1 - \frac{1}{\gamma}\right) \right\},$$

$$J_{22} = \frac{n}{3\gamma^2},$$

$$J_{12} = J_{21} = \frac{n(\gamma - 1 - \gamma^2 \log \gamma)}{3\beta\gamma(\gamma - 1)^2},$$

όπου $Q(p; r) = \sum_{i=1}^{\infty} p^i i^{-r}$ $p \in (0, 1)$ είναι η πολυλογαριθμική συνάρτηση (Erdelyi et al., 1953, σελ.31).

Απόδειξη

Αφού $J_{ij} = E(I_{ij})$ $i, j = 1, 2$, είναι προφανές ότι για τον υπολογισμό του πίνακα $J(\theta)$ απαιτείται ο υπολογισμός των αναμενόμενων τιμών

$$E \left[\frac{T^s e^{-r\beta T}}{\{1-(1-\gamma)e^{-\beta T}\}^2} \right],$$



για $(s,r)=(0,2), (1,1), (2,1)$. Οι ζητούμενες αναμενόμενες τιμές αποκτώνται αξιοποιώντας το γεγονός ότι αν $T \sim EEG$, για $\gamma > 1$ η τυχαία μεταβλητή $Z = \{1 - (1-\gamma)e^{-\beta T}\}^{-1}$ ακολουθεί ομοιόμορφη κατανομή με παραμέτρους γ^{-1} και 1, δηλαδή $Z \sim U(\gamma^{-1}, 1)$. Κατά συνέπεια, οι ροπές πρώτης και δεύτερης τάξης περί το μηδέν της Z , είναι αντίστοιχα

$$E(Z; \gamma) = \frac{\gamma+1}{2\gamma},$$

$$E(Z^2; \gamma) = \frac{\gamma^2 + \gamma + 1}{3\gamma^2}.$$

Για $(s,r) = (0,2)$ η αντίστοιχη αναμενόμενη τιμή υπολογίζεται ως

$$E\left[\frac{e^{-2\beta T}}{\{1 - (1-\gamma)e^{-\beta T}\}^2}\right] = \frac{1}{(1-\gamma)^2} E\{(Z-1)^2\} = \frac{1}{(1-\gamma)^2} E(Z^2 - 2Z + 1) = \frac{1}{3\gamma^2}.$$

Για $(s,r) = (1,1)$, η ζητούμενη αναμενόμενη τιμή προκύπτει σύμφωνα με την ακόλουθη διαδικασία:

$$\begin{aligned} E\left[\frac{T e^{-\beta T}}{\{1 - (1-\gamma)e^{-\beta T}\}^2}\right] &= -\frac{1}{\beta(1-\gamma)} E\left[Z(Z-1) \log\left\{\frac{Z-1}{Z(1-\gamma)}\right\}\right] \\ &= \frac{\gamma}{\beta(\gamma-1)^2} \int_{\gamma^{-1}}^1 z(z-1) \log\left\{\frac{z-1}{z(1-\gamma)}\right\} dz \\ &= \frac{\gamma}{\beta(\gamma-1)^2} \int_{\gamma^{-1}}^1 \log\left\{\frac{z-1}{z(1-\gamma)}\right\} d\left(\frac{z^3}{3} - \frac{z^2}{2}\right) \\ &= \frac{\gamma}{\beta(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2}\right) \log\left\{\frac{z-1}{z(1-\gamma)}\right\} \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \left(\frac{z^3}{3} - \frac{z^2}{2}\right) \frac{1}{z(z-1)} dz \right] \\ &= \frac{\gamma}{\beta(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2}\right) \log\left\{\frac{z-1}{z(1-\gamma)}\right\} \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \left(\frac{z^3}{3} - \frac{z^2}{3} - \frac{z^2}{6}\right) \frac{1}{z(z-1)} dz \right] \\ &= \frac{\gamma}{\beta(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2}\right) \log\left\{\frac{z-1}{z(1-\gamma)}\right\} \Big|_{\gamma^{-1}}^1 - \frac{1}{3} \int_{\gamma^{-1}}^1 z dz + \frac{1}{6} \int_{\gamma^{-1}}^1 \frac{z}{z-1} dz \right] \end{aligned}$$



$$\begin{aligned}
 &= \frac{\gamma}{\beta(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \frac{z^2}{6} \Big|_{\gamma^{-1}}^1 + \frac{1}{6} \int_{\gamma^{-1}}^1 dz + \frac{1}{6} \int_{\gamma^{-1}}^1 \frac{dz}{z-1} \right] \\
 &= \frac{\gamma}{\beta(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \frac{z^2}{6} \Big|_{\gamma^{-1}}^1 + \frac{z}{6} \Big|_{\gamma^{-1}}^1 + \frac{\log(1-z)}{6} \Big|_{\gamma^{-1}}^1 \right] \\
 &= \frac{\gamma}{\beta(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right) \log(1-z) - \left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log\{z(\gamma-1)\} - \frac{z^2}{6} + \frac{z}{6} \right] \Big|_{\gamma^{-1}}^1 .
 \end{aligned}$$

Εφαρμόζοντας τον κανόνα του L'Hospital, αποδεικνύεται ότι η οριακή τιμή της συνάρτησης $\log(1-z) \left\{ 3^{-1} z^3 - 2^{-1} z^2 + 6^{-1} \right\}$ για $z \rightarrow 1$, ισούται με το μηδέν αφού

$$\lim_{z \rightarrow 1} \left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right) \log(1-z) = \lim_{z \rightarrow 1} \frac{\log(1-z)}{\left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right)^{-1}} \quad (\text{μορφή } \frac{\infty}{\infty})$$

$$= \lim_{z \rightarrow 1} \frac{-(1-z)^{-1}}{-(z^2-z) \left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right)^{-2}}$$

$$= \lim_{z \rightarrow 1} \frac{\left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right)^2}{-z^3 + 2z^2 - z} \quad (\text{μορφή } \frac{0}{0})$$

$$= \lim_{z \rightarrow 1} \frac{2 \left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right) (z^2 - z)}{-3z^2 + 4z - 1} \quad (\text{μορφή } \frac{0}{0})$$

$$= \lim_{z \rightarrow 1} \frac{2(z^2 - z)^2 + 2 \left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right) (2z - 1)}{-6z + 4} = 0 .$$

Έτσι,

$$E \left[\frac{Te^{-\beta T}}{\left\{ 1 - (1-\gamma)e^{-\beta T} \right\}^2} \right] = \frac{\gamma}{\beta(\gamma-1)^2} \left[\frac{1}{6} \log(\gamma-1) - \left(\frac{1}{3\gamma^3} - \frac{1}{2\gamma^2} + \frac{1}{6} \right) \log \left(\frac{\gamma-1}{\gamma} \right) \right]$$



$$\begin{aligned}
& + \left(\frac{1}{3\gamma^3} - \frac{1}{2\gamma^2} \right) \log \left(\frac{\gamma-1}{\gamma} \right) + \frac{1}{6\gamma^2} - \frac{1}{6\gamma} \Big] \\
& = \frac{\gamma^2 \log \gamma + 1 - \gamma}{6\beta\gamma(\gamma-1)^2}.
\end{aligned}$$

Στη συνέχεια, παρουσιάζονται τα βήματα υπολογισμού της αναμενόμενης τιμής που προκύπτει για $(s,r) = (2,1)$.

$$\begin{aligned}
E \left[\frac{T^2 e^{-\beta T}}{\{1 - (1-\gamma)e^{-\beta T}\}^2} \right] &= \frac{1}{\beta^2(1-\gamma)} E \left[Z(Z-1) \log^2 \left\{ \frac{Z-1}{Z(1-\gamma)} \right\} \right] \\
&= \frac{-\gamma}{\beta^2(\gamma-1)^2} \int_{\gamma^{-1}}^1 z(z-1) \log^2 \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz \\
&= \frac{-\gamma}{\beta^2(\gamma-1)^2} \int_{\gamma^{-1}}^1 \log^2 \left\{ \frac{z-1}{z(1-\gamma)} \right\} d \left(\frac{z^3}{3} - \frac{z^2}{2} \right) \\
&= \frac{-\gamma}{\beta^2(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log^2 \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - 2 \int_{\gamma^{-1}}^1 \left(\frac{z^3}{3} - \frac{z^2}{2} \right) \frac{1}{z(z-1)} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz \right] \\
&= \frac{-\gamma}{\beta^2(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log^2 \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - 2 \int_{\gamma^{-1}}^1 \left(\frac{z^3}{3} - \frac{z^2}{3} - \frac{z^2}{6} \right) \frac{1}{z(z-1)} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz \right] \\
&= \frac{-\gamma}{\beta^2(\gamma-1)^2} \left[\left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log^2 \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \frac{2}{3} \int_{\gamma^{-1}}^1 z \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz \right. \\
&\quad \left. + \frac{1}{3} \int_{\gamma^{-1}}^1 \frac{z}{z-1} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz \right].
\end{aligned}$$

Όμως,

$$\begin{aligned}
\int_{\gamma^{-1}}^1 z \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz &= \int_{\gamma^{-1}}^1 \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} d \left(\frac{z^2}{2} \right) \\
&= \frac{z^2}{2} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \frac{1}{2} \int_{\gamma^{-1}}^1 \frac{z}{z-1} dz
\end{aligned}$$



$$= \frac{z^2}{2} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \frac{1}{2} \int_{\gamma^{-1}}^1 dz - \frac{1}{2} \int_{\gamma^{-1}}^1 \frac{1}{z-1} dz$$

$$= \frac{z^2}{2} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \frac{z}{2} \Big|_{\gamma^{-1}}^1 - \frac{\log(1-z)}{2} \Big|_{\gamma^{-1}}^1.$$

Επίσης,

$$\int_{\gamma^{-1}}^1 \frac{z}{z-1} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} dz = \int_{\gamma^{-1}}^1 \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} d\{z + \log(1-z)\}$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \frac{\{z + \log(1-z)\}}{z(z-1)} dz$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \frac{1}{z-1} dz - \int_{\gamma^{-1}}^1 \frac{\log(1-z)}{z(z-1)} dz$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \log(1-z) \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \log(1-z) \left(\frac{1}{z-1} - \frac{1}{z} \right) dz$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \log(1-z) \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \frac{\log(1-z)}{z-1} dz + \int_{\gamma^{-1}}^1 \frac{\log(1-z)}{z} dz$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \log(1-z) \Big|_{\gamma^{-1}}^1 - \frac{\log^2(1-z)}{2} \Big|_{\gamma^{-1}}^1 - \int_{\gamma^{-1}}^1 \frac{1}{z} \sum_{v=1}^{\infty} \frac{z^v}{v} dz$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \log(1-z) \Big|_{\gamma^{-1}}^1 - \frac{\log^2(1-z)}{2} \Big|_{\gamma^{-1}}^1 - \sum_{v=1}^{\infty} \frac{z^v}{v^2} \Big|_{\gamma^{-1}}^1$$

$$= \{z + \log(1-z)\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \Big|_{\gamma^{-1}}^1 - \log(1-z) \Big|_{\gamma^{-1}}^1 - \frac{\log^2(1-z)}{2} \Big|_{\gamma^{-1}}^1 - \frac{\pi^2}{6} + Q\left(\frac{1}{\gamma}, 2\right),$$

όπου Q είναι η πολυλογαριθμική συνάρτηση.

Άρα,

$$E \left[\frac{T^2 e^{-\beta T}}{\{1 - (1-\gamma)e^{-\beta T}\}^2} \right] = \frac{-\gamma}{\beta^2 (\gamma-1)^2} \left[\left[\left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log^2 \left(\frac{z-1}{z(1-\gamma)} \right) - \frac{z^2}{3} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \right] \right]$$



$$+ \frac{z}{3} + \frac{1}{3} \left\{ z + \log(1-z) \right\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} - \frac{\log^2(1-z)}{6} \left] \Bigg|_{\gamma^{-1}}^1 - \frac{\pi^2}{18} + \frac{1}{3} \mathcal{Q} \left(\frac{1}{\gamma}, 2 \right) \right].$$

Κατά συνέπεια, το πρόβλημα υπολογισμού της ζητούμενης αναμενόμενης τιμής, ανάγεται ουσιαστικά στον υπολογισμό της οριακής τιμής για $z \rightarrow 1$, της παράστασης A :

$$\begin{aligned} A &= \left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log^2 \left\{ \frac{z-1}{z(1-\gamma)} \right\} - \frac{z^2}{3} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} + \frac{z}{3} + \frac{1}{3} \left\{ z + \log(1-z) \right\} \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} \\ &\quad - \frac{\log^2(1-z)}{6} \\ &= \left(\frac{z^3}{3} - \frac{z^2}{2} \right) \left[\log^2(1-z) - 2 \log(1-z) \log \{ z(\gamma-1) \} + \log^2 \{ z(\gamma-1) \} \right] \\ &\quad + \left(\frac{z}{3} - \frac{z^2}{3} \right) \log \left\{ \frac{z-1}{z(1-\gamma)} \right\} + \frac{z}{3} + \frac{\log^2(1-z)}{3} - \frac{\log(1-z) \log \{ z(\gamma-1) \}}{3} - \frac{\log^2(1-z)}{6} \\ &= \left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right) \log^2(1-z) + \left(-\frac{2z^3}{3} + z^2 - \frac{1}{3} \right) \log(1-z) \log \{ z(\gamma-1) \} \\ &\quad + \left(\frac{z^3}{3} - \frac{z^2}{2} \right) \log^2 \{ z(\gamma-1) \} + \left(\frac{z}{3} - \frac{z^2}{3} \right) \log(1-z) - \left(\frac{z}{3} - \frac{z^2}{3} \right) \log \{ z(\gamma-1) \} + \frac{z}{3}. \end{aligned}$$

Η χρήση του κανόνα του L'Hospital για την άρση των απροσδιοριστιών της παράστασης A , οδηγεί στα ακόλουθα αποτελέσματα:

$$\begin{aligned} 1. \lim_{z \rightarrow 1} \left(\frac{z^3}{3} - \frac{z^2}{2} + \frac{1}{6} \right) \log^2(1-z) &= \lim_{z \rightarrow 1} \left(\frac{z^3}{3} - \frac{z^2}{3} - \frac{z^2}{6} + \frac{1}{6} \right) \log^2(1-z) \\ &= \frac{1}{3} \lim_{z \rightarrow 1} z^2(z-1) \log^2(1-z) + \frac{1}{6} \lim_{z \rightarrow 1} (1-z^2) \log^2(1-z) \\ &= \frac{1}{3} \lim_{z \rightarrow 1} \frac{\log^2(1-z)}{z^{-2}(z-1)^{-1}} + \frac{1}{6} \lim_{z \rightarrow 1} \frac{\log^2(1-z)}{(1-z^2)^{-1}} \quad (\text{μορφές } \frac{\infty}{\infty}) \\ &= \frac{1}{3} \lim_{z \rightarrow 1} \frac{-2(1-z)^{-1} \log(1-z)}{-(3z^2-2z)z^{-4}(z-1)^{-2}} + \frac{1}{6} \lim_{z \rightarrow 1} \frac{-2(1-z)^{-1} \log(1-z)}{2z(1-z^2)^{-2}} \\ &= \frac{2}{3} \lim_{z \rightarrow 1} \frac{\log(1-z)}{(3z-2)z^{-3}(1-z)^{-1}} - \frac{1}{6} \lim_{z \rightarrow 1} \frac{\log(1-z)}{z(1-z)^{-1}(1+z)^{-2}} \quad (\text{μορφές } \frac{\infty}{\infty}) \end{aligned}$$



$$\begin{aligned}
 &= \frac{2}{3} \lim_{z \rightarrow 1} \frac{-(1-z)^{-1}}{\{3z^3(1-z) - (3z-2)(3z^2-4z^3)\} z^{-6}(1-z)^{-2}} \\
 &\quad - \frac{1}{6} \lim_{z \rightarrow 1} \frac{-(1-z)^{-1}}{\{(1-z)(1+z)^2 - z(-3z^2-2z+1)\} (1-z)^{-2}(1+z)^{-4}} \\
 &= -\frac{2}{3} \lim_{z \rightarrow 1} \frac{z^4(1-z)}{3z(1-z) - (3z-2)(3-4z)} + \frac{1}{6} \lim_{z \rightarrow 1} \frac{(1-z)(1+z)^4}{(1-z)(1+z)^2 - z(-3z^2-2z+1)} \\
 &= -\frac{2}{3} \cdot \frac{0}{1} + \frac{1}{6} \cdot \frac{0}{4} = 0.
 \end{aligned}$$

$$\begin{aligned}
 2. \lim_{z \rightarrow 1} \left(-\frac{2z^3}{3} + z^2 - \frac{1}{3} \right) \log(1-z) &= \lim_{z \rightarrow 1} \left(-\frac{2z^3}{3} + \frac{2z^2}{3} + \frac{z^2}{3} - \frac{1}{3} \right) \log(1-z) \\
 &= \frac{2}{3} \lim_{z \rightarrow 1} z^2(1-z) \log(1-z) - \frac{1}{3} \lim_{z \rightarrow 1} (1-z^2) \log(1-z) \\
 &= \frac{2}{3} \lim_{z \rightarrow 1} \frac{\log(1-z)}{z^{-2}(1-z)^{-1}} - \frac{1}{3} \lim_{z \rightarrow 1} \frac{\log(1-z)}{(1-z^2)^{-1}} \quad (\text{μορφές } \frac{\infty}{\infty}) \\
 &= \frac{2}{3} \lim_{z \rightarrow 1} \frac{-(1-z)^{-1}}{-(2z-3z^2)z^{-4}(1-z)^{-2}} - \frac{1}{3} \lim_{z \rightarrow 1} \frac{-(1-z)^{-1}}{2z(1-z^2)^{-2}} \\
 &= \frac{2}{3} \lim_{z \rightarrow 1} \frac{z^3(1-z)}{2-3z} + \frac{1}{6} \lim_{z \rightarrow 1} \frac{(1+z)^2(1-z)}{z} \\
 &= \frac{2}{3} \cdot \frac{0}{-1} + \frac{1}{6} \cdot \frac{0}{1} = 0.
 \end{aligned}$$

$$\begin{aligned}
 3. \lim_{z \rightarrow 1} \left(\frac{z}{3} - \frac{z^2}{3} \right) \log(1-z) &= \frac{1}{3} \lim_{z \rightarrow 1} z(1-z) \log(1-z) \\
 &= \frac{1}{3} \lim_{z \rightarrow 1} \frac{\log(1-z)}{z^{-1}(1-z)^{-1}} \quad (\text{μορφή } \frac{\infty}{\infty}) \\
 &= \frac{1}{3} \lim_{z \rightarrow 1} \frac{-(1-z)^{-1}}{-(1-2z)z^{-2}(1-z)^{-2}} \\
 &= \frac{1}{3} \lim_{z \rightarrow 1} \frac{z^2(1-z)}{1-2z} \\
 &= \frac{1}{3} \cdot \frac{0}{-1} = 0.
 \end{aligned}$$



Άρα,

$$\lim_{z \rightarrow 1} A = -\frac{1}{6} \log^2(\gamma - 1) + \frac{1}{3}$$

ενώ επίσης,

$$A|_{z=\gamma^{-1}} = -\frac{1}{6} \log^2\left(1 - \frac{1}{\gamma}\right) + \frac{1}{3\gamma}.$$

Τελικά,

$$\begin{aligned} E\left[\frac{T^2 e^{-\beta T}}{\{1 - (1-\gamma)e^{-\beta T}\}^2}\right] &= \frac{-\gamma}{\beta^2(\gamma-1)^2} \left\{ \lim_{z \rightarrow 1} A - A|_{z=\gamma^{-1}} - \frac{\pi^2}{18} + \frac{1}{3} Q\left(\frac{1}{\gamma}, 2\right) \right\} \\ &= \frac{-\gamma}{3\beta^2(\gamma-1)^2} \left\{ -\frac{1}{2} \log^2(\gamma-1) + 1 + \frac{1}{2} \log^2\left(1 - \frac{1}{\gamma}\right) - \frac{1}{\gamma} - \frac{\pi^2}{6} + Q\left(\frac{1}{\gamma}, 2\right) \right\}. \end{aligned}$$

Αξιοποιώντας τα παραγόμενα αποτελέσματα στον υπολογισμό των αναμενόμενων τιμών των I_{ij} $i, j = 1, 2$ των σχέσεων (2.9), προκύπτουν τα στοιχεία του πίνακα αναμενόμενης πληροφορίας.

$$\begin{aligned} J_{11} &= E\left[\frac{n}{\beta^2} - 2(1-\gamma) \sum_{i=1}^n \frac{T_i^2 e^{-\beta T_i}}{\{1 - (1-\gamma)e^{-\beta T_i}\}^2}\right] \\ &= \frac{n}{\beta^2} + \frac{2n\gamma}{3\beta^2(1-\gamma)} \left\{ -\frac{1}{2} \log^2(\gamma-1) + 1 + \frac{1}{2} \log^2\left(1 - \frac{1}{\gamma}\right) - \frac{1}{\gamma} - \frac{\pi^2}{6} + Q\left(\frac{1}{\gamma}, 2\right) \right\} \\ &= \frac{n}{3\beta^2(1-\gamma)} \left\{ 1 - \gamma - \gamma \log^2(\gamma-1) + \gamma \log^2\left(1 - \frac{1}{\gamma}\right) - \frac{\gamma\pi^2}{3} + 2\gamma Q\left(\frac{1}{\gamma}, 2\right) \right\}. \end{aligned}$$

$$J_{22} = E\left[\frac{n}{\gamma^2} - 2 \sum_{i=1}^n \frac{e^{-2\beta T_i}}{\{1 - (1-\gamma)e^{-\beta T_i}\}^2}\right] = \frac{n}{3\gamma^2}.$$

$$J_{12} = J_{21} = E\left[-2 \sum_{i=1}^n \frac{T_i e^{-\beta T_i}}{\{1 - (1-\gamma)e^{-\beta T_i}\}^2}\right] = \frac{n(-\gamma^2 \log \gamma - 1 + \gamma)}{3\beta\gamma(\gamma-1)^2}.$$

Εκτιμητές του ασυμπτωτικού πίνακα διακυμάνσεων-συνδιακυμάνσεων μπορούν να προκύψουν αντιστρέφοντας τους πίνακες $J(\theta)$ ή $I(\theta; \gamma_{obs})$ και υπολογίζοντας την τιμή τους στο $\theta = \hat{\theta}$.



2.8 Χαρακτηρισμοί

Στα ακόλουθα δύο θεωρήματα παρουσιάζονται χαρακτηρισμοί της κατανομής EEG για $\gamma \neq 1$. Αποδεικνύεται ότι ο μέσος υπολειπόμενος χρόνος ζωής χαρακτηρίζει την κατανομή και επίσης, ότι στην οικογένεια των μοντέλων ανάλυσης επιβίωσης με πεπερασμένη τιμή συνάρτησης κινδύνου στο μηδέν, η πυκνότητα της κατανομής EEG είναι ανάλογη του στιγμιαίου ρυθμού μεταβολής της συνάρτησης κινδύνου.

Θεώρημα 2.5

Η μη αρνητική τυχαία μεταβλητή T ακολουθεί την κατανομή EEG αν και μόνο αν ο μέσος υπολειπόμενος χρόνος ζωής της, δίνεται από τη σχέση:

$$m(t; \theta) = -\frac{e^{\beta t}}{\beta(1-\gamma)} \left\{ 1 - (1-\gamma)e^{-\beta t} \right\} \log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\}, \quad t \geq 0. \quad (2.10)$$

Απόδειξη

Το ευθύ έχει αποδειχθεί στην ενότητα 2.6. Για την απόδειξη του αντίστροφου, από τη (2.10), προκύπτει η σχέση:

$$\frac{S(t)}{\int_t^\infty S(u) du} = \frac{\left[\log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\} \right]'}{\log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\}}$$

ή η ισοδύναμή της

$$\left\{ \log \int_t^\infty S(u) du \right\}' = \begin{cases} \left[\log \left[-\log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\} \right] \right]' & \text{για } \gamma \in (0,1) \\ \left[\log \left[\log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\} \right] \right]' & \text{για } \gamma > 1. \end{cases} \quad (2.11)$$

Για $\gamma > 1$, από τη (2.11) έπεται ότι

$$\log \int_t^\infty S(u) du = \log \left[\log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\} \right] + c_1$$

$$\Leftrightarrow \int_t^\infty S(u) du = c_2 \log \left\{ 1 - (1-\gamma)e^{-\beta t} \right\}, \quad (c_2 = e^{c_1}).$$

Παραγωγίζοντας την τελευταία σχέση ως προς t , η συνάρτηση επιβίωσης γράφεται στη μορφή:



$$S(t) = \frac{-c_2 \beta (1-\gamma) e^{-\beta t}}{1 - (1-\gamma) e^{-\beta t}}$$

ενώ η τιμή της σταθεράς c_2 υπολογίζεται από τη συνθήκη $S(0) = 1$, ως $-\gamma \{\beta(1-\gamma)\}^{-1}$ και αποδεικνύεται το ζητούμενο. Όμοια αποκτάται η συνάρτηση επιβίωσης της κατανομής *EEG* για $0 < \gamma < 1$.

Θεώρημα 2.6

Η μη αρνητική τυχαία μεταβλητή T με πυκνότητα $f(t)$ και συνάρτηση κατανομής $F(t)$, ακολουθεί την κατανομή *EEG* αν και μόνο αν

$$h'(t) = cf(t), \quad (2.12)$$

όπου $c \in R$, $h(0) < \infty$ και $h'(t)$ είναι η παράγωγος της συνάρτησης κινδύνου στο t .

Απόδειξη

Από τη σχέση (2.5) αποδεικνύεται το ευθύ για $c = \beta(\gamma - 1)\gamma^{-1}$. Για την απόδειξη του αντίστροφου, ξεκινώντας από τη σχέση (2.12) προκύπτουν τα ακόλουθα:

$$\begin{aligned} h'(t) = cf(t) &\Leftrightarrow \int_0^t h'(t) dt = cF(t) \Leftrightarrow h(t) - h(0) = cF(t) \\ &\Leftrightarrow f(t) = cF(t)\{1 - F(t)\} + h(0)\{1 - F(t)\} \\ &\Leftrightarrow -\{1 - F(t)\}' = -c\{1 - F(t) - 1\}\{1 - F(t)\} + h(0)\{1 - F(t)\} \\ &\Leftrightarrow -\{1 - F(t)\}' = -c\{1 - F(t)\}^2 + \{c + h(0)\}\{1 - F(t)\} \\ &\Leftrightarrow \{1 - F(t)\}' + \{c + h(0)\}\{1 - F(t)\} = c\{1 - F(t)\}^2. \end{aligned}$$

Η τελευταία είναι μια διαφορική εξίσωση Bernoulli, η οποία γράφεται επίσης ως

$$\{1 - F(t)\}' \{1 - F(t)\}^{-2} + \{c + h(0)\}\{1 - F(t)\}^{-1} = c$$

και με τις αντικαταστάσεις

$$z(t) = \{1 - F(t)\}^{-1},$$

$$z'(t) = -\{1 - F(t)\}^{-2} \{1 - F(t)\}',$$

ανάγεται στην ακόλουθη γραμμική διαφορική εξίσωση πρώτης τάξης:

$$z'(t) - \{c + h(0)\}z(t) = -c.$$



Οι λύσεις της τελευταίας δίνονται από τον τύπο:

$$z(t) = ke^{\{c+h(0)\}t} + \frac{c}{c+h(0)}, \quad (k \text{ αυθαίρετη σταθερά}).$$

Ειδικά για τη λύση που πληρεί την αρχική συνθήκη $z(0) = 1$, η τιμή της σταθεράς k υπολογίζεται ως $h(0)\{c+h(0)\}^{-1}$. Άρα,

$$z(t) = \frac{h(0)e^{\{c+h(0)\}t} + c}{c+h(0)}$$

ή ισοδύναμα,

$$F(t) = \frac{1 - e^{-\{c+h(0)\}t}}{1 + c\{h(0)\}^{-1}e^{-\{c+h(0)\}t}}.$$

Θέτοντας $c+h(0) = \beta$ και $c\{h(0)\}^{-1} = \gamma - 1$, προκύπτει η σχέση (2.2) και αποδεικνύεται το ζητούμενο.

2.9 Εφαρμογές

Στη συνέχεια, παρουσιάζονται εφαρμογές της κατανομής *EEG* ως μοντέλου με αύξουσα συνάρτηση κινδύνου αφού η αποτίμηση της πρακτικής της αξίας ως μοντέλου με φθίνουσα συνάρτηση κινδύνου έγινε από τους Adamidis και Loukas (1998). Οι εφαρμογές αφορούν σε δύο δείγματα πραγματικών δεδομένων (παρατίθενται στο Παράρτημα). Το πρώτο δείγμα, αποτελείται από 100 παρατηρήσεις στους κύκλους αντοχής δειγμάτων νήματος πολυεστέρα-βισκόζης μήκους 100 εκατοστών, όπως προέκυψαν στα πλαίσια πειραματικής διαδικασίας ελέγχου του ορίου θραύσης τους, με επίπεδο πίεσης 2.3% (Quesenberry και Kent, 1982). Το δεύτερο δείγμα, αποτελείται από 107 χρόνους ζωής των πίσω δεξιών φρένων ερπυστριοφόρων ελκυστήρων τύπου D9G-66A (Barlow και Campo, 1975 και Chang και Rao, 1993). Παράλληλα με την *EEG*, προσαρμόστηκαν στα δεδομένα οι κατανομές Γάμα και Weibull με αντίστοιχες πυκνότητες:

$$f_1(t; \theta_1) = \beta_1^{\gamma_1} t^{\gamma_1 - 1} e^{-\beta_1 t} \{\Gamma(\gamma_1)\}^{-1}, \quad t > 0,$$

$$f_2(t; \theta_2) = \gamma_2 \beta_2^{\gamma_2} t^{\gamma_2 - 1} e^{-(\beta_2 t)^{\gamma_2}}, \quad t > 0,$$

όπου $\theta_j = (\beta_j, \gamma_j)$ $j = 1, 2$. Η προσαρμογή τους έγινε με τη μέθοδο της μέγιστης



πιθανοφάνειας και η καταλληλότητά τους για την περιγραφή των δεδομένων εξετάστηκε τόσο με στατιστικές όσο και με γραφικές μεθόδους. Στον Πίνακα 2.2, παρουσιάζονται οι τιμές των εκτιμητών μέγιστης πιθανοφάνειας, η τιμή του λογαρίθμου της πιθανοφάνειας εκτιμώμενη στο $\hat{\theta}$ ($\log L$) και οι τιμές των στατιστικών χ^2 (χ_4^2) και Kolmogorov-Smirnov (K-S) με τις αντίστοιχες p-τιμές. Για την απόκτηση των εκτιμητών μέγιστης πιθανοφάνειας

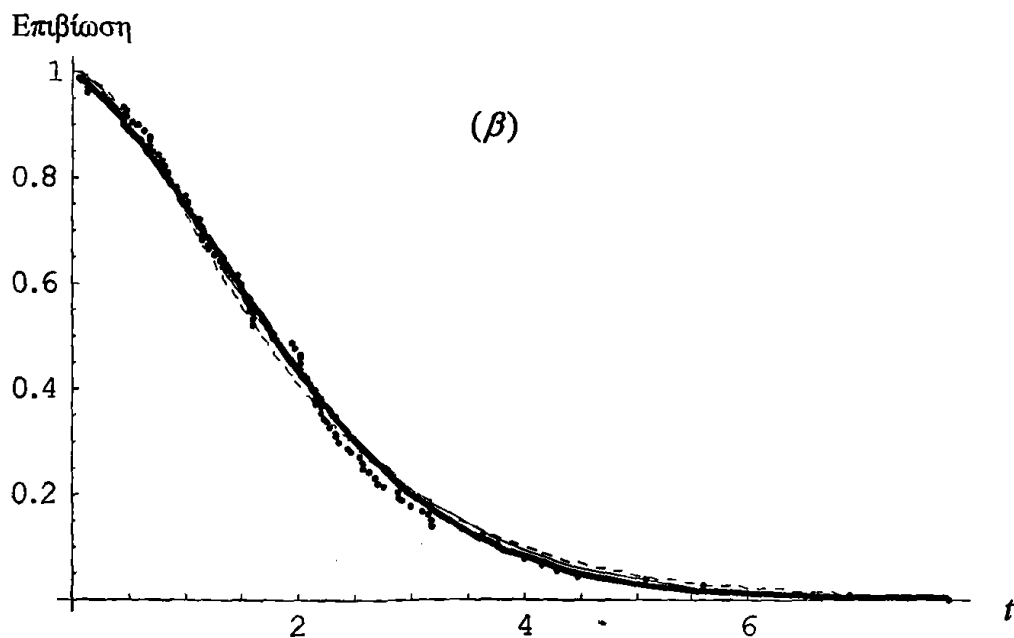
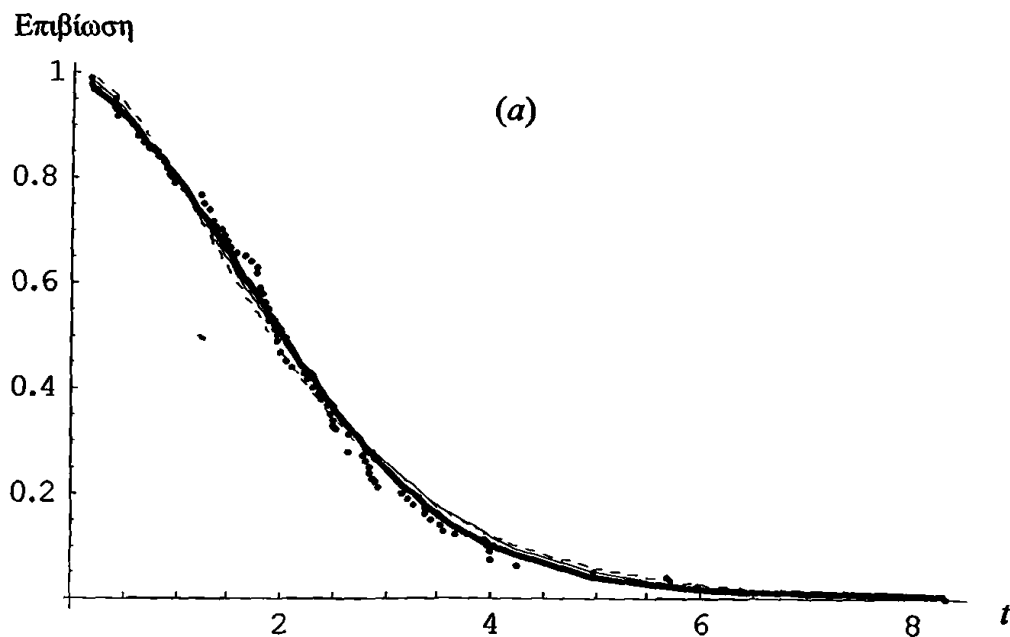
	Κατανομή	Εκτιμητές	$\log L$	χ_4^2 (p-τιμή)	K-S (p-τιμή)
1(n=100)	EEG	$\hat{\theta}=(1.059, 7.449)$	-164.604	7.094 (0.1310)	0.059 (0.9481)
	Γάμα	$\hat{\theta}=(1.008, 2.239)$	-164.727	7.704 (0.1030)	0.095 (0.3118)
	Weibull	$\hat{\theta}=(0.403, 1.604)$	-164.682	7.882 (0.0960)	0.076 (0.6080)
2(n=107)	EEG	$\hat{\theta}=(1.035, 5.395)$	-170.574	4.256 (0.3725)	0.043 (0.9999)
	Γάμα	$\hat{\theta}=(0.943, 1.908)$	-171.476	6.438 (0.1687)	0.068 (0.7343)
	Weibull	$\hat{\theta}=(0.447, 1.486)$	-170.876	5.534 (0.2368)	0.049 (0.9999)

Πίνακας 2.2 Οι εκτιμητές μέγιστης πιθανοφάνειας και οι τιμές του λογαρίθμου πιθανοφάνειας και των στατιστικών χ^2 και Kolmogorov-Smirnov με τις αντίστοιχες p τιμές, από την προσαρμογή των τριών κατανομών στους κύκλους αντοχής δειγμάτων νήματος πολυεστέρα-βισκόζης (1) και στους χρόνους ζωής των πίσω δεξιών φρένων ερπυστριοφόρων ελκυστήρων (2).

της κατανομής *EEG*, απαιτήθηκαν 210 και 181 κλήσεις της μεθόδου EM για το πρώτο και δεύτερο δείγμα αντίστοιχα ενώ και στις δύο περιπτώσεις, για τον υπολογισμό της τιμής του στατιστικού χ^2 , οι παρατηρήσεις ταξινομήθηκαν σε 7 ομάδες. Στο Σχήμα 2.3, απεικονίζονται η εμπειρική συνάρτηση επιβίωσης και οι εκτιμώμενες συναρτήσεις επιβίωσης των τριών κατανομών, για καθένα από τα δύο δείγματα.



Οι στατιστικές όσο και οι γραφικές μέθοδοι που χρησιμοποιήθηκαν, συνηγορούν υπέρ της άποψης ότι η κατανομή *EEG* προσαρμόζεται στα δεδομένα τουλάχιστον το ίδιο ικανοποιητικά όσο οι δημοφιλείς εναλλακτικές της κατανομές, Weibull και Γάμα.



Σχήμα 2.3 Καμπύλες επιβίωσης της εμπειρικής κατανομής (•), της κατανομής *EEG* (—), της κατανομής Weibull (---) και της Γάμα κατανομής (-.-) (α) στους κύκλους αντοχής δειγμάτων νήματος πολυεστέρα-βισκόζης και (β) στους χρόνους ζωής φρένων ερπυστροφόρων ελκυστήρων.



Κεφάλαιο 3

Ένα Μοντέλο Ανάλυσης Επιβίωσης με Μονότονες, Κυπελλοειδούς Μορφής και Μονοκόρυφες Συναρτήσεις Κινδύνου

Αν και οι πλέον δημοφιλείς κατανομές ανάλυσης επιβίωσης είναι εκείνες με μονότονη συνάρτηση κινδύνου (π.χ. Weibull, Γάμα), συχνά, η φύση του μηχανισμού 'θανάτου-αποτυχίας' τις καθιστά ανεπαρκείς (ή και ακατάλληλες) για την περιγραφή της ικανότητας επιβίωσης του υπό εξέταση πληθυσμού. Η συχνότητα διαχείρισης ενός τέτοιου ενδεχομένου γίνεται ιδιαίτερα μεγάλη, όταν μελετάται η συνολική επιβίωση βιολογικών οργανισμών ή εργοστασιακών προϊόντων. Στην περίπτωση αυτή, η συνηθισμένη συμπεριφορά της παρατηρούμενης συνάρτησης κινδύνου χαρακτηρίζεται από αρχικά υψηλές τιμές που όμως φθίνουν σταδιακά μέχρι μιας συγκεκριμένης τιμής όπου προσωρινά σχεδόν σταθεροποιούνται και στη συνέχεια αυξάνουν λόγω γήρατος ή φθοράς χρήσης (Gaver και Acar, 1979). Κατά συνέπεια, οι πλέον ρεαλιστικές κατανομές για τη μοντελοποίηση του τρόπου εξέλιξης ενός τέτοιου φαινομένου, είναι αυτές που μπορούν να δώσουν κυπελλοειδούς μορφής συναρτήσεις κινδύνου. Μια ανασκόπηση τέτοιων κατανομών παρουσιάζεται από τους Rajarshi και Rajarshi (1988) και πιο πρόσφατα από τους Lai et al. (2001). Σε άλλες περιπτώσεις, όπως για παράδειγμα όταν παρατηρείται μεγάλη συχνότητα πρόωρων 'θανάτων-αποτυχιών', καταλληλότερες για την περιγραφή της βιωσιμότητας ή ανθεκτικότητας του υπό μελέτη πληθυσμού, είναι οι κατανομές με μονοκόρυφη συνάρτηση κινδύνου όπως οι Lognormal και Inverse Gaussian (Johnson et al., 1994).



Στη συνέχεια, παρουσιάζεται ένα μοντέλο ανάλυσης επιβίωσης με τρεις παραμέτρους, το οποίο περιλαμβάνει την κατανομή Weibull ως ειδική περίπτωση. Η προτεινόμενη κατανομή μπορεί να δώσει μονότονες, κυπελλοειδούς μορφής και μονοκόρυφες συναρτήσεις κινδύνου και ως εκ τούτου, δυνητικά να μοντελοποιήσει διαδικασίες - φαινόμενα με τα χαρακτηριστικά που προαναφέρθηκαν. Η τελευταία αναφέρεται στη συνέχεια ως η κατανομή *IDBU*, από τα αρχικά των λέξεων της διεθνούς ορολογίας *increasing, decreasing, bathtub-shaped* και *unimodal* που υποδηλώνουν τα διαφορετικά είδη μονοτονίας της συνάρτησης κινδύνου. Μελετώνται διάφορες μαθηματικές και στατιστικές της ιδιότητες ενώ παρουσιάζεται η φυσική διεργασία παραγωγής του μοντέλου μέσα από τη θεωρία ανταγωνιστικών κινδύνων. Το πρόβλημα της εκτίμησης των παραμέτρων, αντιμετωπίζεται με τη μέθοδο της μέγιστης πιθανοφάνειας. Τέλος, παρουσιάζονται εφαρμογές του μοντέλου σε πραγματικά δεδομένα.

3.1 Η κατανομή

Η σ.π.π. της κατανομής *IDBU*, δίνεται από τη σχέση:

$$f(t; \theta) = a\beta\lambda t^{\beta-1} (1 + \lambda t^\beta)^{a-1} \exp\{1 - (1 + \lambda t^\beta)^a\}, \quad t > 0, \quad (3.1)$$

με $\theta = (a, \beta, \lambda)$, όπου $a, \beta > 0$ είναι παράμετροι μορφής (*shape*) και $\lambda > 0$ παράμετρος κλίμακας (*scale*).

Αποδεικνύεται ότι για $\beta < 1$, η σ.π.π. της κατανομής είναι φθίνουσα με $\lim_{t \rightarrow 0^+} f(t; \theta) = \infty$ και $\lim_{t \rightarrow \infty} f(t; \theta) = 0$, για $\beta = 1$ είναι επίσης φθίνουσα με $\lim_{t \rightarrow 0^+} f(t; \theta) = a\lambda$ και $\lim_{t \rightarrow \infty} f(t; \theta) = 0$ ενώ για $\beta > 1$, είναι μονοκόρυφη με $\lim_{t \rightarrow 0^+} f(t; \theta) = 0$, $\lim_{t \rightarrow \infty} f(t; \theta) = 0$ και μέγιστο στο σημείο μηδενισμού της παραγώγου της ως προς t

$$f'(t; \theta) = f(t; \theta)t^{-1}(1 + \lambda t^\beta)^{-1} [\beta - 1 + a\beta\lambda t^\beta \{1 - (1 + \lambda t^\beta)^a\} - \lambda t^\beta].$$

Το είδος μονοτονίας της f διαπιστώνεται από την τελευταία σχέση. Επίσης, για $\beta > 1$, ισχύει ότι

$$f(t; \theta) \leq c a \beta (1 + \lambda t^\beta)^a \exp\{1 - (1 + \lambda t^\beta)^a\},$$

όπου $c = \max(1, \lambda)$ ενώ για $\beta < 1$, είναι

$$f(t; \theta) \leq t^{-1} a \beta (1 + \lambda t^\beta)^a \exp\{1 - (1 + \lambda t^\beta)^a\}.$$

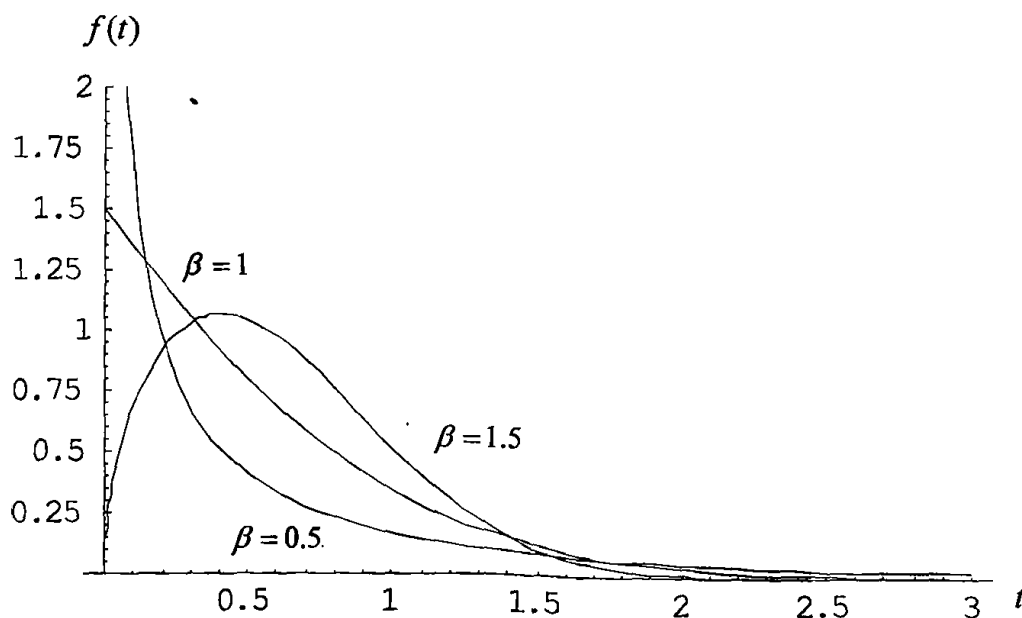


Όμως, με χρήση του κανόνα του L'Hospital αποδεικνύεται ότι

$$\lim_{t \rightarrow \infty} a\beta(1 + \lambda t^\beta)^a \exp\{1 - (1 + \lambda t^\beta)^a\} = \lim_{t \rightarrow \infty} t^{-1} a\beta(1 + \lambda t^\beta)^a \exp\{1 - (1 + \lambda t^\beta)^a\} = 0.$$

Άρα, και $\lim_{t \rightarrow \infty} f(t; \theta) = 0$ για κάθε $\beta > 0$. Οι οριακές τιμές της f στο 0^+ προκύπτουν άμεσα,

Στο Σχήμα 3.1 απεικονίζονται πυκνότητες της οικογένειας κατανομών στη σχέση (3.1), για επιλεγμένες τιμές των παραμέτρων της.



Σχήμα 3.1 Πυκνότητες της κατανομής IDBU για $\lambda=1$, $a=1.5$ και $\beta=0.5, 1, 1.5$.

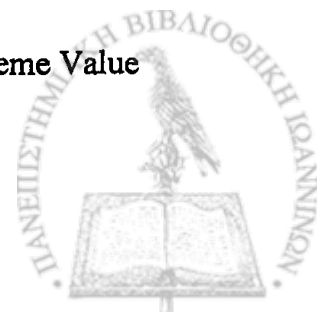
3.2 Σχέσεις με άλλες κατανομές

Για $a=1$, η κατανομή IDBU περιλαμβάνει ως ειδική περίπτωση την κατανομή Weibull με παραμέτρους μορφής και κλίμακας β και λ αντίστοιχα. Επίσης, αν $T \sim IDBU$, αποδεικνύεται ότι:

(i) Η τυχαία μεταβλητή $Y = 1 + \lambda T^\beta$, ακολουθεί κατανομή Weibull με παραμέτρους μορφής και κλίμακας α και 1 αντίστοιχα, περικομμένη στο $(1, \infty)$,

(ii) Η τυχαία μεταβλητή $Y = (1 + \lambda T^\beta)^a - 1$, ακολουθεί Εκθετική κατανομή με μέση τιμή 1,

(iii) Η τυχαία μεταβλητή $Y = \log(1 + \lambda T^\beta)$, ακολουθεί την τροποποιημένη Extreme Value



κατανομή (Johnson et al., 1995) με παραμέτρους μορφής και κλίμακας 1 και a αντίστοιχα.

(iv) Η τυχαία μεταβλητή $Y = \{\log(1 + \lambda T^\beta)\}^{1/\beta}$, ακολουθεί την κατανομή Exponential Power (Smith and Bain, 1975) με παραμέτρους μορφής και κλίμακας β και a αντίστοιχα.

Σχέσεις με άλλες κατανομές μπορούν να βρεθούν μέσω αυτών, των περιπτώσεων (i) – (iv). Για παράδειγμα στην (iv), η τυχαία μεταβλητή $Y_1 = 1 - \exp\{1 - \exp(aY^\beta)\}$ ακολουθεί ομοιόμορφη κατανομή με παραμέτρους 0 και 1.

Τα τελευταία αποτελέσματα προκύπτουν χρησιμοποιώντας κάποια από τις μεθοδολογίες αλλαγής μεταβλητών.

3.3 Πιθανότητες και ροπές

Η συνάρτηση κατανομής της (3.1), η οποία εκφράζει την πιθανότητα ‘θανάτου-αποτυχίας’ μέχρι το χρόνο t , δίνεται από τη σχέση:

$$F(t; \theta) = 1 - \exp\{1 - (1 + \lambda t^\beta)^a\}, \quad t > 0.$$

Το p εκατοστιαίο σημείο της κατανομής δηλαδή ο χρόνος μέχρι τον οποίο η τυχαία πειραματική μονάδα επιβιώνει με πιθανότητα $1 - p$, υπολογίζεται ως

$$t_p = F^{-1}(p) = \left[\lambda^{-1} \left\{ \left[1 - \log(1 - p) \right]^{1/a} - 1 \right\} \right]^{1/\beta}, \quad 0 < p < 1.$$

Κατά συνέπεια, η διάμεσος δηλαδή ο χρόνος μέχρι τον οποίο δυνητικά επιβιώνουν οι μισές πειραματικές μονάδες, είναι

$$\left[\lambda^{-1} \left\{ (1 + \log 2)^{1/a} - 1 \right\} \right]^{1/\beta}.$$

Οι ροπές r τάξης περί το μηδέν της κατανομής *IDBU*, προσδιορίζονται από το ολοκλήρωμα:

$$\mu_r = E(T^r; \theta) = r \int_0^\infty t^{r-1} \exp\{1 - (1 + \lambda t^\beta)^a\} dt.$$

Αν και η απόκτησή τους σε κλειστή μορφή δεν είναι εφικτή, μπορούν εύκολα να υπολογιστούν αριθμητικά. Ο Πίνακας 3.1 δίνει τη μέση τιμή ($E(T)$) και την τυπική απόκλιση ($\sqrt{Var(T)}$) της κατανομής για διάφορες τιμές των παραμέτρων της. Για τον υπολογισμό τους χρησιμοποιήθηκε ρουτίνα της NAG (1991).



(a, β)	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 2$	
	$E(T)$	$\sqrt{Var(T)}$	$E(T)$	$\sqrt{Var(T)}$	$E(T)$	$\sqrt{Var(T)}$
(0.9, 0.5)	11.84	28.808	2.96	7.202	0.74	1.8
(1.5, 2)	0.948	0.453	0.67	0.32	0.474	0.226
(0.25, 2)	6.737	9.089	4.763	6.427	3.368	4.545
(0.25, 4.5)	2.034	1.013	1.743	0.868	1.494	0.744
(2, 0.3)	2.048	6.237	0.203	0.619	0.02	0.061
(2, 0.7)	0.806	0.924	0.299	0.343	0.111	0.128
(0.5, 1)	8	12.649	4	6.325	2	3.162
(1.5, 1)	1.103	0.97	0.552	0.485	0.276	0.243
(1, 0.8)	2.695	3.397	1.133	1.428	0.476	0.6
(1, 1.25)	1.622	1.305	0.931	0.75	0.535	0.431

Πίνακας 3.1 Μέσες τιμές και τυπικές αποκλίσεις της κατανομής *IDBU* για επιλεγμένες τιμές των παραμέτρων της.

3.4 Συναρτήσεις επιβίωσης και κινδύνου

Ένας ισοδύναμος και ενδεχομένως πιο κομψός τρόπος ορισμού της κατανομής *IDBU*, είναι μέσω της συνάρτησης επιβίωσής της

$$S(t; \theta) = \exp\left\{1 - (1 + \lambda t^\beta)^a\right\}, \quad t > 0. \quad (3.2)$$

Κατά συνέπεια, από τις (3.1) και (3.2), η συνάρτηση κινδύνου της κατανομής, δίνεται από τη σχέση:

$$h(t; \theta) = a\beta\lambda t^{\beta-1} (1 + \lambda t^\beta)^{a-1}, \quad t > 0. \quad (3.3)$$

Η τελευταία συνάρτηση μπορεί να είναι μονότονη, κυπελλοειδούς μορφής ή μονοκόρυφη. Το είδος μονοτονίας της διαφοροποιείται μέσα στα τμήματα του παραμετρικού χώρου που δημιουργούνται από τις καμπύλες $a=1$, $\beta=1$ και $a\beta=1$. Πιο συγκεκριμένα εξετάζοντας την παράγωγο της h ως προς t ,

$$h'(t; \theta) = a\beta\lambda t^{\beta-2} (1 + \lambda t^\beta)^{a-2} \left\{ \beta - 1 + \lambda t^\beta (a\beta - 1) \right\},$$

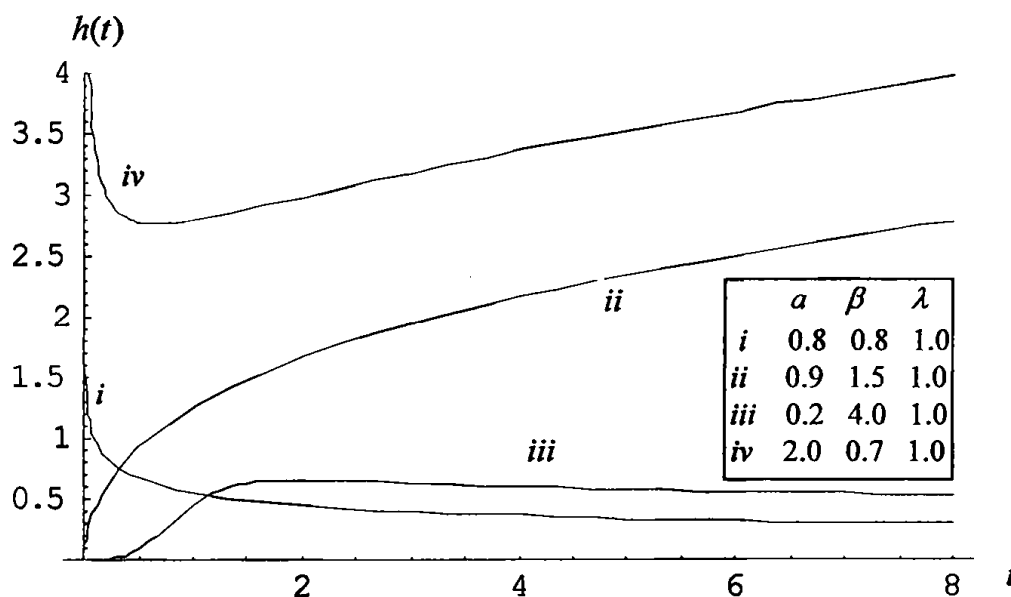
συνάγονται τα ακόλουθα συμπεράσματα:

- για $a=\beta=1$, η $h(t; \theta)$ είναι σταθερή (εκθετική),
- για $a > 1$ και $\beta \geq 1$ ($a < 1$ και $\beta \leq 1$), η $h(t; \theta)$ είναι μονότονα αύξουσα (φθίνουσα),
- για $a \geq 1$ και $\beta < 1$,



- i) αν $a\beta \leq 1$, η $h(t;\theta)$ είναι μονότονα φθίνουσα,
 ii) αν $a\beta > 1$, η $h(t;\theta)$ είναι κυπελλοειδούς μορφής,
 d) για $a \leq 1$ και $\beta > 1$,
 i) αν $a\beta < 1$, η $h(t;\theta)$ είναι μονοκόρυφη,
 ii) αν $a\beta \geq 1$, η $h(t;\theta)$ είναι μονότονα αύξουσα.

Επίσης, κάτω από τις υποθέσεις c(ii) και d(i), η h παρουσιάζει αντίστοιχα ελάχιστο και μέγιστο στο σημείο $\{(1-\beta)\lambda^{-1}(a\beta-1)^{-1}\}^{1/\beta}$. Από όσα προαναφέρθηκαν, είναι επίσης προφανές, ότι η προτεινόμενη κατανομή παρέχει μια ευρεία κλάση μονότονων συναρτήσεων κινδύνου, ευρύτερη εκείνης της κατανομής Weibull ($a=1$). Στο Σχήμα 3.2, παρουσιάζονται οι διαφορετικοί τύποι μονοτονίας της συνάρτησης κινδύνου για κατάλληλα επιλεγμένες τιμές των παραμέτρων της.



Σχήμα 3.2 Συναρτήσεις κινδύνου της κατανομής IDBU: (i) φθίνουσα, (ii) αύξουσα, (iii) μονοκόρυφη και (iv) κυπελλοειδούς μορφής.

3.5 Φυσική ερμηνεία

Η προτεινόμενη κατανομή αποτελεί γενίκευση - υπό την έννοια της επέκτασης του παραμετρικού χώρου - του μοντέλου που περιγράφεται στη συνέχεια.

Υποθέτοντας ότι η παράμετρος a παίρνει μόνο θετικές ακέραιες τιμές, η συνάρτηση κινδύνου της σχέσης (3.3), μπορεί να γραφεί ως άθροισμα a όρων ως



$$h(t; \theta) = a\beta\lambda t^{\beta-1} \sum_{r=0}^{a-1} \binom{a-1}{r} (\lambda t^\beta)^r = \sum_{r=0}^{a-1} \gamma_r \beta_r t^{\beta_r-1}, \quad t > 0, \quad (3.4)$$

$$\text{όπου } \gamma_r = \binom{a}{r+1} \lambda^{r+1} \text{ και } \beta_r = \beta(r+1).$$

Κατά συνέπεια, η σχέση (3.4) δίνει τη συνάρτηση κινδύνου σειριακού συστήματος αποτελούμενου από a συνιστώσες με χρόνους ζωής ανεξάρτητες τυχαίες μεταβλητές Weibull. Ένας ισοδύναμος τρόπος θεώρησης του φυσικού μηχανισμού που παράγει την (3.4) με βάση τη θεωρία ανταγωνιστικών κινδύνων, είναι αυτός που θέλει κάθε πειραματική μονάδα εκτεθειμένη σε ένα από a διαφορετικά αίτια 'αποτυχίας'. Αν T_i ($i = 0, \dots, a-1$) είναι ο χρόνος ζωής που θα παρατηρούνταν υπό την επίδραση μόνο του $i^{\text{οστού}}$ αίτιου αποτυχίας και υποθέτοντας ότι οι T_i είναι ανεξάρτητες τυχαίες μεταβλητές με συναρτήσεις κίνδυνου $h_i(t; \theta_i) = \gamma_i \beta_i t^{\beta_i-1}$ $i = 0, \dots, a-1$ αντίστοιχα, τότε ο παρατηρούμενος χρόνος ζωής $T = \min\{T_i\}_{i=0}^{a-1}$ περιγράφεται από το μοντέλο της σχέσης (3.4). Η πιθανότητα αποτυχίας στο διάστημα $[c, d)$ από το $i^{\text{οστό}}$ αίτιο, παρουσία όλων των υπολοίπων και δοθείσης της επιβίωσης μέχρι το χρόνο c , υπολογίζεται ως εξής:

$$P_i(c, d) = P(c \leq T_i < d, T_i < T_j \quad i \neq j | T \geq c)$$

$$= \frac{\int_{t_i}^{\infty} \int_c^d \dots \int_{t_i}^{\infty} \prod_{j=0}^{a-1} f_j(t_j; \theta_j) dt_j}{S_T(c; \theta)}$$

$$= \frac{\int_c^d f_i(t_i; \theta_i) \left\{ \prod_{\substack{j=0 \\ i \neq j}}^{a-1} S_j(t_i; \theta_j) \right\} dt_i}{S_T(c; \theta)}$$

$$= \frac{\int_c^d h_i(t_i; \theta_i) S_T(t_i; \theta) dt_i}{S_T(c; \theta)}$$

$$= \int_c^d h_i(t_i; \theta_i) \exp \left\{ - \int_0^{t_i} h_T(t; \theta) dt + \int_0^c h_T(t; \theta) dt \right\} dt_i$$



$$\begin{aligned}
&= \int_c^d h_i(t_i; \theta_i) \exp\left\{-\int_c^{t_i} h_T(t; \theta) dt\right\} dt_i \\
&= \int_c^d h_i(t_i; \theta_i) \exp\left\{-\int_c^{t_i} \sum_{i=0}^{a-1} h_i(t; \theta_i) dt\right\} dt_i.
\end{aligned}$$

Ως εκ τούτου, η πιθανότητα αποτυχίας από το $i^{\text{οστό}}$ αίτιο είναι $\pi_i = P_i(0, \infty)$.

Μπορεί να παρατηρήσει κανείς ότι για $\beta > 1$ ($\beta = 1$ και $a > 1$), οι συνιστώσες συναρτήσεις κινδύνου είναι αύξουσες ($a - 1$ αύξουσες και μια σταθερή) και το 'σύστημα' παρουσιάζει φθορά ή γήρας με την πάροδο του χρόνου, εν αντιθέσει με την περίπτωση $\beta < 1$ και $a\beta < 1$ ($\beta < 1$ και $a\beta = 1$), όπου οι επιμέρους συναρτήσεις κινδύνου είναι φθίνουσες ($a - 1$ φθίνουσες και μια σταθερή) και το 'σύστημα' γίνεται προοδευτικά ανθεκτικότερο. Η περίπτωση που $a = \beta = 1$, είναι η ειδική περίπτωση της Εκθετικής κατανομής με σταθερή συνάρτηση κινδύνου ενώ για $\beta < 1$ και $a\beta > 1$, η συνάρτηση h στη σχέση (3.4), αποτελείται από τουλάχιστον μια αύξουσα και μια φθίνουσα συνάρτηση κινδύνου, δίνοντας σύμφωνα με επιχειρήματα της προηγούμενης ενότητας γράφημα κυπελλοειδούς μορφής. Στην τελευταία περίπτωση, το 'σύστημα' αρχικά παρουσιάζει βελτίωση στη λειτουργία του, για να περάσει στη συνέχεια σε κατάσταση σταδιακής φθοράς.

3.6 Συμπερασματολογία

3.6.1 Εκτιμητική με τη μέθοδο της μέγιστης πιθανοφάνειας

Για τυχαίο δείγμα μεγέθους n , $y_{obs} = (t_i, i = 1, \dots, n)$ από την (3.1), η συνάρτηση πιθανοφάνειας και ο λογάριθμός της, δίνονται αντίστοιχα από τις σχέσεις:

$$L = L(\theta; y_{obs}) = (a\beta\lambda)^n \left(\prod_{i=1}^n t_i^{\beta-1} \right) \left\{ \prod_{i=1}^n (1 + \lambda t_i^\beta)^{a-1} \right\} \exp\left\{ n - \sum_{i=1}^n (1 + \lambda t_i^\beta)^a \right\},$$

$$\ell = \ell(\theta; y_{obs}) = \log L(\theta; y_{obs})$$

$$= n \log(a\beta\lambda) + (\beta - 1) \sum_{i=1}^n \log t_i + (a - 1) \sum_{i=1}^n \log(1 + \lambda t_i^\beta) + n - \sum_{i=1}^n (1 + \lambda t_i^\beta)^a.$$

Θέτοντας τις μερικές παραγώγους $\partial \ell / \partial \theta_j$, $j = 1, 2, 3$ ίσες με το μηδέν, προκύπτουν οι εξισώσεις μέγιστης πιθανοφάνειας:



$$\frac{\partial \ell}{\partial a} = \frac{n}{a} + \sum_{i=1}^n \log(1 + \lambda t_i^\beta) \{1 - (1 + \lambda t_i^\beta)^a\} = 0,$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \log t_i + \sum_{i=1}^n \frac{\lambda t_i^\beta \log t_i}{1 + \lambda t_i^\beta} \{a - 1 - a(1 + \lambda t_i^\beta)^a\} = 0,$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^n \frac{t_i^\beta}{1 + \lambda t_i^\beta} \{a - 1 - a(1 + \lambda t_i^\beta)^a\} = 0.$$

Είναι προφανές ότι η επίλυση του συστήματος των παραπάνω μη γραμμικών ως προς τις παραμέτρους εξισώσεων, δεν μπορεί να γίνει με αναλυτικές μεθόδους. Κατά συνέπεια, για την απόκτηση των εκτιμητών μέγιστης πιθανοφάνειας θα πρέπει να χρησιμοποιηθεί κάποια από τις μεθόδους αριθμητικής ανάλυσης, όπως για παράδειγμα η μέθοδος Newton-Raphson.

3.6.2 Ασυμπτωτική κατανομή των εκτιμητών μέγιστης πιθανοφάνειας

Σύμφωνα με γνωστές ιδιότητες των εκτιμητών μέγιστης πιθανοφάνειας, η ασυμπτωτική κατανομή του $\hat{\theta}$ είναι κατά προσέγγιση η τριδιάστατη κανονική με μέση τιμή θ και πίνακα διακυμάνσεων-συνδιακυμάνσεων τον αντίστροφο του αναμενόμενου πίνακα πληροφορίας $J(\theta)$, με στοιχεία

$$J_{ij} = E(-\partial^2 \ell / \partial \theta_i \partial \theta_j), \quad i, j = 1, 2, 3.$$

Η ασυμπτωτική κανονικότητα εξακολουθεί να ισχύει αν για λόγους υπολογιστικής ευκολίας ο αναμενόμενος πίνακας πληροφορίας $J(\theta)$, αντικατασταθεί από τον παρατηρούμενο $I(\theta)$, με στοιχεία

$$I_{ij} = -\partial^2 \ell / \partial \theta_i \partial \theta_j, \quad i, j = 1, 2, 3,$$

εκτιμώμενο στο $\theta = \hat{\theta}$, αφού ο τελευταίος κάτω από ασθενείς συνθήκες αποτελεί συνεπή εκτιμητή του πρώτου. Τα στοιχεία του παρατηρούμενου πίνακα πληροφορίας, δίνονται αναλυτικά από τις ακόλουθες σχέσεις:

$$I_{11} = \frac{n}{a^2} + \sum_{i=1}^n (1 + \lambda t_i^\beta)^a \log^2(1 + \lambda t_i^\beta) \Big|_{\theta=\hat{\theta}},$$

$$I_{22} = \frac{n}{\beta^2} - \sum_{i=1}^n \frac{\lambda t_i^\beta \log^2 t_i}{(1 + \lambda t_i^\beta)^2} \{a - 1 - a(1 + \lambda t_i^\beta)^a (1 + a \lambda t_i^\beta)\} \Big|_{\theta=\hat{\theta}},$$



$$I_{33} = \frac{n}{\lambda^2} + (a-1) \sum_{i=1}^n \frac{t_i^{2\beta}}{(1 + \lambda t_i^\beta)^2} \left\{ 1 + a(1 + \lambda t_i^\beta)^a \right\} \Bigg|_{\theta=\hat{\theta}},$$

$$I_{12} = I_{21} = - \sum_{i=1}^n \frac{\lambda t_i^\beta \log t_i}{1 + \lambda t_i^\beta} \left\{ 1 - (1 + \lambda t_i^\beta)^a - a(1 + \lambda t_i^\beta)^a \log(1 + \lambda t_i^\beta) \right\} \Bigg|_{\theta=\hat{\theta}},$$

$$I_{13} = I_{31} = - \sum_{i=1}^n \frac{t_i^\beta}{1 + \lambda t_i^\beta} \left\{ 1 - (1 + \lambda t_i^\beta)^a - a(1 + \lambda t_i^\beta)^a \log(1 + \lambda t_i^\beta) \right\} \Bigg|_{\theta=\hat{\theta}},$$

$$I_{23} = I_{32} = - \sum_{i=1}^n \frac{t_i^\beta \log t_i}{(1 + \lambda t_i^\beta)^2} \left\{ a - 1 - a(1 + \lambda t_i^\beta)^a (1 + a \lambda t_i^\beta) \right\} \Bigg|_{\theta=\hat{\theta}}.$$

3.7 Εφαρμογές

Στη συνέχεια, παρουσιάζονται δύο εφαρμογές της κατανομής *IDBU* σε πραγματικά δεδομένα. Η πρώτη εφαρμογή, αφορά σε δείγμα 46 τιμών στους χρόνους συντήρησης συστήματος τηλεπικοινωνίας αεροπλάνου (Chhikara και Folks, 1977) και η δεύτερη, σε δείγμα 101 παρατηρήσεων στους χρόνους ζωής μάντων αεροναυπηγικής όταν υποβάλλονται σε επίπεδο πίεσης 90% (Andrews και Herzberg, 1985, σελ. 182). Τα δύο δείγματα παρατίθενται στο Παράρτημα. Η προσαρμογή της κατανομής στα δύο δείγματα έγινε με τη μέθοδο της μέγιστης πιθανοφάνειας. Η ποιότητα προσαρμογής της εξετάστηκε τόσο με στατιστικές όσο και με γραφικές μεθόδους. Στον Πίνακα 3.2, παρουσιάζονται οι εκτιμητές μέγιστης πιθανοφάνειας και οι τιμές του στατιστικού των Kolmogorov-Smirnov (K-S) με τις αντίστοιχες p-τιμές. Από τα στοιχεία του Πίνακα, είναι προφανές ότι η προσαρμοζόμενη συνάρτηση κινδύνου είναι μονοκόρυφη στην πρώτη περίπτωση και κυπελλοειδούς μορφής στη δεύτερη.

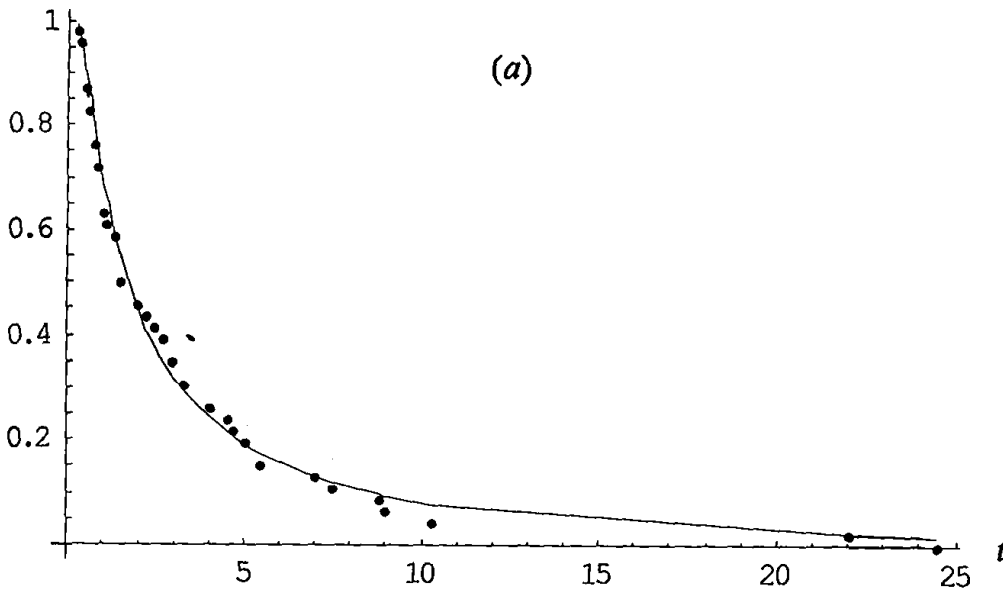
Δείγμα	Εκτιμητές	K-S (p-τιμές)
1 (n = 46)	$\hat{\theta} = (0.1226, 3.3643, 12.4828)$	0.073 (0.999)
2 (n = 101)	$\hat{\theta} = (1.3082, 0.8618, 0.6925)$	0.086 (0.439)

Πίνακας 3.2 Οι εκτιμητές μέγιστης πιθανοφάνειας και οι τιμές του στατιστικού των Kolmogorov-Smirnov με τις αντίστοιχες p-τιμές από την προσαρμογή της κατανομής *IDBU* στους χρόνους συντήρησης συστήματος τηλεπικοινωνίας αεροπλάνου (1) και στους χρόνους ζωής μάντων αεροναυπηγικής (2).

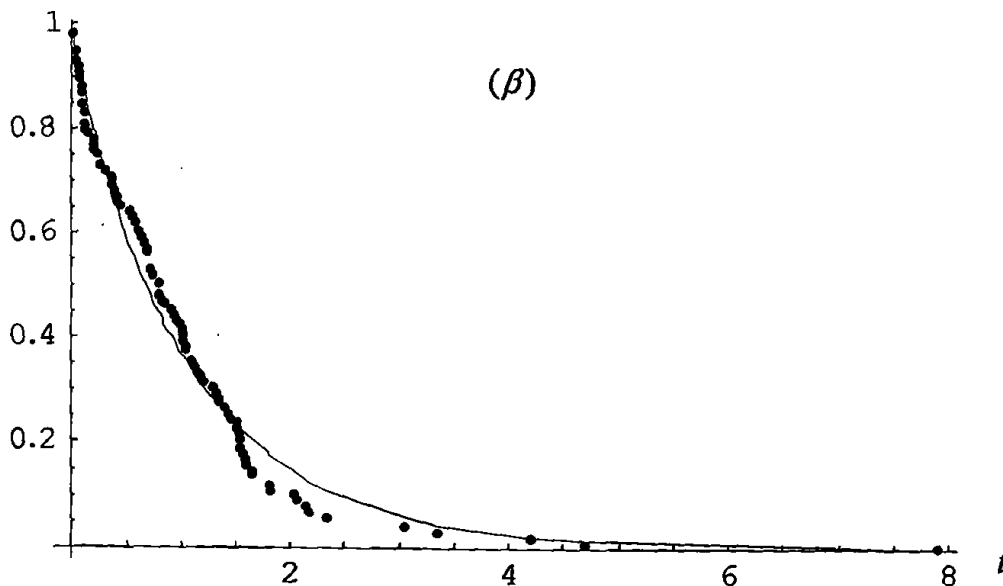


Στο Σχήμα 3.3, απεικονίζονται τα γραφήματα της εμπειρικής και της εκτιμώμενης συνάρτησης επιβίωσης για καθένα από τα δύο δείγματα.

Επιβίωση



Επιβίωση



Σχήμα 3.3 Καμπύλες επιβίωσης της εμπειρικής κατανομής (•) και της κατανομής IDBU (—) (α) στους χρόνους συντήρησης συστήματος τηλεπικοινωνίας αεροπλάνου και (β) στους χρόνους ζωής μάντων αεροναυπηγικής.

Τόσο οι στατιστικές όσο και οι γραφικές μέθοδοι που χρησιμοποιήθηκαν, συγκλίνουν στην άποψη ότι η κατανομή IDBU περιγράφει ικανοποιητικά τα δεδομένα.



Κεφάλαιο 4

Διδιάστατες Γενικευμένες Εκθετικές – Γεωμετρικές Κατανομές

Στο κεφάλαιο αυτό, παρουσιάζονται διδιάστατες κατανομές *EEG* και περιγράφονται οι φυσικές διεργασίες που τις παράγουν. Για κάθε μορφή, δίνονται οι περιθώριες και οι δεσμευμένες κατανομές και υπολογίζονται οι συναρτήσεις επιβίωσης και κινδύνου. Η συσχέτιση των τυχαίων μεταβλητών που περιγράφουν το διδιάστατο πληθυσμό, εξετάζεται μέσω του συντελεστή συσχέτισης του Pearson.

4.1 Διδιάστατη *EEG*-1

Η πρώτη διδιάστατη μορφή της κατανομής *EEG*, περιγράφει την από κοινού συμπεριφορά των χρόνων ζωής ενός πληθυσμού, όταν οι περιθώριοι πληθυσμοί περιγράφονται από τις υποθέσεις του θεωρήματος 2.1 και έχουν κοινή παράμετρο ετερογένειας. Οι λεπτομέρειες στον τρόπο παραγωγής της, δίνονται στο θεώρημα που ακολουθεί.

Θεώρημα 4.1

Έστω ότι οι $X|A$ και $Y|A$ είναι ανεξάρτητες τυχαίες μεταβλητές, οι οποίες ακολουθούν την τροποποιημένη Extreme Value κατανομή, με αντίστοιχες πυκνότητες

$$f(x|a; \beta_1) = a\beta_1 \exp\{\beta_1 x + a(1 - e^{\beta_1 x})\}, \quad x \geq 0,$$
$$f(y|a; \beta_2) = a\beta_2 \exp\{\beta_2 y + a(1 - e^{\beta_2 y})\}, \quad y \geq 0,$$



όπου $a, \beta_1, \beta_2 \in R^+$ και η A είναι εκθετική τυχαία μεταβλητή με σ.π.π.

$$f(a; \gamma) = \gamma e^{-\gamma a}, \quad a \geq 0,$$

με $\gamma \in R^+$. Τότε:

$$f(x, y; \theta) = \frac{2\beta_1\beta_2\gamma e^{\beta_1x + \beta_2y}}{(e^{\beta_1x} + e^{\beta_2y} + \gamma - 2)^3}, \quad x, y \geq 0, \quad (4.1)$$

όπου $\theta = (\beta_1, \beta_2, \gamma)$.

Απόδειξη

Από την υπόθεση της ανεξαρτησίας των $X | A$ και $Y | A$, προκύπτει ότι

$$f(x, y | a; \beta_1, \beta_2) = a^2 \beta_1 \beta_2 \exp\{\beta_1 x + \beta_2 y + a(2 - e^{\beta_1 x} - e^{\beta_2 y})\}, \quad x, y \geq 0$$

και άρα,

$$f(x, y, a; \theta) = a^2 \beta_1 \beta_2 \gamma \exp\{\beta_1 x + \beta_2 y + a(2 - \gamma - e^{\beta_1 x} - e^{\beta_2 y})\}, \quad x, y \geq 0.$$

Κατά συνέπεια, η από κοινού κατανομή των X, Y , προσδιορίζεται ολοκληρώνοντας την τελευταία ως προς A στο διάστημα $[0, \infty)$, οπότε

$$\begin{aligned} f(x, y; \theta) &= \int_0^{\infty} a^2 \beta_1 \beta_2 \gamma \exp\{\beta_1 x + \beta_2 y + a(2 - \gamma - e^{\beta_1 x} - e^{\beta_2 y})\} da \\ &= \beta_1 \beta_2 \gamma e^{\beta_1 x + \beta_2 y} \int_0^{\infty} a^2 \exp\{-a(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)\} da \\ &= \frac{2\beta_1\beta_2\gamma e^{\beta_1x + \beta_2y}}{(e^{\beta_1x} + e^{\beta_2y} + \gamma - 2)^3}, \quad x, y \geq 0. \end{aligned}$$

4.1.1 Περιθώριες και δεσμευμένες κατανομές

Η περιθώρια κατανομή της X , είναι

$$\begin{aligned} f(x; \theta_1) &= \int_0^{\infty} \frac{2\beta_1\beta_2\gamma e^{\beta_1x + \beta_2y}}{(e^{\beta_1x} + e^{\beta_2y} + \gamma - 2)^3} dy \\ &= -\beta_1\gamma e^{\beta_1x} \int_0^{\infty} d\{(e^{\beta_1x} + e^{\beta_2y} + \gamma - 2)^{-2}\} \\ &= \frac{\beta_1\gamma e^{\beta_1x}}{(e^{\beta_1x} + \gamma - 1)^2} \end{aligned}$$



$$= \frac{\beta_1 \gamma e^{-\beta_1 x}}{\{1 - (1 - \gamma)e^{-\beta_1 x}\}^2}, \quad x \geq 0, \quad (4.2)$$

όπου $\theta_1 = (\beta_1, \gamma)$.

Όμοια, η περιθώρια κατανομή της Y , υπολογίζεται ως

$$f(y; \theta_2) = \frac{\beta_2 \gamma e^{-\beta_2 y}}{\{1 - (1 - \gamma)e^{-\beta_2 y}\}^2}, \quad y \geq 0, \quad (4.3)$$

όπου $\theta_2 = (\beta_2, \gamma)$.

Είναι προφανές ότι οι περιθώριες κατανομές EEG, έχουν την ίδια παράμετρο μορφής και διαφορετική παράμετρο κλίμακας.

Οι δεσμευμένες κατανομές των $X|Y$ και $Y|X$ προσδιορίζονται από τις σχέσεις (4.1) - (4.3) και οι αντίστοιχες σ.π.π. δίνονται από τις σχέσεις:

$$f(x|y; \theta) = \frac{2\beta_1 e^{\beta_1 x} (e^{\beta_2 y} + \gamma - 1)^2}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^3}, \quad x \geq 0,$$

$$f(y|x; \theta) = \frac{2\beta_2 e^{\beta_2 y} (e^{\beta_1 x} + \gamma - 1)^2}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^3}, \quad y \geq 0.$$

4.1.2 Οι συναρτήσεις επιβίωσης και κινδύνου

Η συνάρτηση επιβίωσης της διδιάστατης κατανομής EEG-1, υπολογίζεται ως εξής:

$$\begin{aligned} S(x, y; \theta) &= \int_x^\infty \int_y^\infty \frac{2\beta_1 \beta_2 \gamma e^{\beta_1 x + \beta_2 y}}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^3} dy dx \\ &= - \int_x^\infty \beta_1 \gamma e^{\beta_1 x} \int_y^\infty d\{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^{-2}\} dx \\ &= \int_x^\infty \frac{\beta_1 \gamma e^{\beta_1 x}}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^2} dx \\ &= - \int_x^\infty \gamma d\{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^{-1}\} \\ &= \frac{\gamma}{e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2}, \quad x, y \geq 0. \end{aligned}$$



Αξίζει να σημειωθεί ότι οι περιθώριες κατανομές μπορούν να προσδιοριστούν από τη διδιάστατη συνάρτηση επιβίωσης, η οποία για $y=0$ ή $x=0$, δίνει αντίστοιχα τις συναρτήσεις επιβίωσης των X και Y .

Αντικαθιστώντας τις εκφράσεις των συναρτήσεων πυκνότητας πιθανότητας και επιβίωσης στη σχέση (1.4), η από κοινού συνάρτηση κινδύνου των X και Y , υπολογίζεται ως

$$h(x, y; \theta) = \frac{2\beta_1\beta_2 e^{\beta_1 x + \beta_2 y}}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^2}, \quad x, y \geq 0.$$

Οι δεσμευμένες συναρτήσεις επιβίωσης και κινδύνου της $X|Y$, δίνονται αντίστοιχα από τις σχέσεις:

$$S(x|y; \theta) = \frac{(e^{\beta_2 y} + \gamma - 1)^2}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^2}, \quad x \geq 0,$$

$$h(x|y; \theta) = \frac{2\beta_1 e^{\beta_1 x}}{e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2}, \quad x \geq 0.$$

Η παράγωγος της δεσμευμένης συνάρτησης κινδύνου, είναι

$$h'(x|y; \theta) = \frac{2\beta_1^2 e^{\beta_1 x} (e^{\beta_2 y} + \gamma - 2)}{(e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2)^2},$$

από το πρόσημο της οποίας προκύπτει ότι η δεσμευμένη συνάρτηση κινδύνου είναι αύξουσα για $\gamma \geq 1$ ενώ για $0 < \gamma < 1$, αν $y > \beta_2^{-1} \log(2 - \gamma)$, είναι αύξουσα ενώ αν $0 < y < \beta_2^{-1} \log(2 - \gamma)$, είναι φθίνουσα. Με εναλλαγή των x και y και των β_1, β_2 , προκύπτουν τα αντίστοιχα στοιχεία για την κατανομή της $Y|X$.

Η μια από τις δύο δεσμευμένες συναρτήσεις κινδύνου στη σχέση (1.5), είναι

$$\bar{h}(x|y; \theta) = \frac{\beta_1 e^{\beta_1 x}}{e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2}, \quad x \geq 0.$$

Η μονοτονία της, είναι η ίδια με εκείνη της συνάρτησης κινδύνου της $X|Y$. Η δεσμευμένη συνάρτηση κινδύνου $\bar{h}(y|x)$ και η μονοτονία της, προκύπτουν με εναλλαγή των x και y και των β_1, β_2 , στην τελευταία σχέση και στα σχόλια για τη μονοτονία της $h(x|y)$, αντίστοιχα.

Οι περιθώριες συναρτήσεις επιβίωσης των X και Y , προκύπτουν θέτοντας αντίστοιχα



$y = 0$ και $x = 0$ στη διδιάστατη συνάρτηση επιβίωσης. Κατά συνέπεια, η αναλυτική τους έκφραση είναι

$$S(x; \theta_1) = \frac{\gamma e^{-\beta_1 x}}{1 - (1 - \gamma)e^{-\beta_1 x}}, \quad x \geq 0,$$

$$S(y; \theta_2) = \frac{\gamma e^{-\beta_2 y}}{1 - (1 - \gamma)e^{-\beta_2 y}}, \quad y \geq 0.$$

Επίσης, αφού οι περιθώριες κατανομές EEG έχουν την ίδια παράμετρο μορφής, οι αντίστοιχες συναρτήσεις κινδύνου

$$h(x; \theta_1) = \frac{\beta_1}{1 - (1 - \gamma)e^{-\beta_1 x}}, \quad x \geq 0,$$

$$h(y; \theta_2) = \frac{\beta_2}{1 - (1 - \gamma)e^{-\beta_2 y}}, \quad y \geq 0,$$

είναι και οι δύο αύξουσες για $\gamma > 1$ και φθίνουσες για $0 < \gamma < 1$.

4.1.3 Ο συντελεστής συσχέτισης

Στην παράγραφο αυτή, εξετάζεται ο συντελεστής συσχέτισης του Pearson για τη διδιάστατη κατανομή της σχέσης (4.1).

Για τον υπολογισμό της συνδιακύμανσης των X και Y , απαιτείται ο υπολογισμός της αναμενόμενης τιμής του γινομένου XY . Όμως,

$$\begin{aligned} E(XY; \theta) &= \int_0^{\infty} \int_0^{\infty} \frac{\gamma}{e^{\beta_1 x} + e^{\beta_2 y} + \gamma - 2} dx dy \\ &= \int_0^{\infty} \int_0^{\infty} \frac{\gamma e^{-\beta_1 x}}{1 + (e^{\beta_2 y} + \gamma - 2)e^{-\beta_1 x}} dx dy \\ &= -\frac{\gamma}{\beta_1} \int_0^{\infty} \frac{1}{e^{\beta_2 y} + \gamma - 2} \int_0^{\infty} d \left[\log \left\{ 1 + (e^{\beta_2 y} + \gamma - 2)e^{-\beta_1 x} \right\} \right] dy \\ &= \frac{\gamma}{\beta_1} \int_0^{\infty} \frac{\log(e^{\beta_2 y} + \gamma - 1)}{e^{\beta_2 y} + \gamma - 2} dy. \end{aligned}$$

Η επίλυση του τελευταίου ολοκληρώματος με αναλυτικές μεθόδους, δε στάθηκε δυνατή,



ωστόσο αποδεικνύεται ότι συγκλίνει και η τιμή του μπορεί να υπολογιστεί αριθμητικά.

Λαμβάνοντας υπόψιν τις σχέσεις (4.2) και (4.3) και τα αποτελέσματα της ενότητας 2.4, οι μέσες τιμές των X και Y υπολογίζονται αντίστοιχα ως

$$E(X; \theta_1) = \frac{-\gamma \log \gamma}{\beta_1(1-\gamma)},$$

$$E(Y; \theta_2) = \frac{-\gamma \log \gamma}{\beta_2(1-\gamma)}$$

ενώ οι διακυμάνσεις τους, είναι

$$\text{Var}(X; \theta_1) = \frac{\gamma}{\beta_1^2} \left\{ 2\Phi(1-\gamma, 2, 1) - \frac{\gamma \log^2 \gamma}{(1-\gamma)^2} \right\},$$

$$\text{Var}(Y; \theta_2) = \frac{\gamma}{\beta_2^2} \left\{ 2\Phi(1-\gamma, 2, 1) - \frac{\gamma \log^2 \gamma}{(1-\gamma)^2} \right\}.$$

Υπολογίζοντας τις αριθμητικές τιμές του τελευταίου ολοκληρώματος και των $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$ και αντικαθιστώντας τα αποτελέσματα στη σχέση (1.7), υπολογίζεται η αριθμητική τιμή του συντελεστή συσχέτισης του Pearson.

4.2 Διδιάστατη $EEG-2$

Η διδιάστατη κατανομή $EEG-2$, είναι η από κοινού κατανομή των ελάχιστων διατεταγμένων στατιστικών από δύο ισομεγέθη δείγματα εκθετικών τυχαίων μεταβλητών, όταν το πλήθος τους σε κάθε δείγμα ακολουθεί γεωμετρική κατανομή.

Θεώρημα 4.2

Έστω $\{X_i\}_{i=1}^N$ και $\{Y_i\}_{i=1}^N$ τυχαία δείγματα από την Εκθετική κατανομή, με αντίστοιχες πυκνότητες

$$f(x; a) = ae^{-ax}, \quad x \geq 0,$$

$$f(y; \beta) = \beta e^{-\beta y}, \quad y \geq 0,$$

όπου $a, \beta \in R^+$ και N γεωμετρική τυχαία μεταβλητή με συνάρτηση πιθανότητας

$$f(n; p) = (1-p)p^{n-1}, \quad n \in N,$$



με $p \in (0,1)$. Αν $X = \min\{X_i\}_{i=1}^N$ και $Y = \min\{Y_i\}_{i=1}^N$, και η $X|N$ είναι ανεξάρτητη της $Y|N$, τότε:

$$f(x, y; \theta) = \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1 + pe^{-(ax+\beta y)}\}}{\{1 - pe^{-(ax+\beta y)}\}^3}, \quad x, y \geq 0, \quad (4.4)$$

όπου $\theta = (a, \beta, p)$.

Απόδειξη

Αν $X = \min\{X_i\}_{i=1}^N$ και $Y = \min\{Y_i\}_{i=1}^N$, τότε:

$$f(x|n; a) = nae^{-nax}, \quad x \geq 0,$$

$$f(y|n; \beta) = n\beta e^{-n\beta y}, \quad y \geq 0.$$

Αφού οι $X|N$ και $Y|N$ είναι ανεξάρτητες, η από κοινού δεσμευμένη κατανομή τους, είναι

$$f(x, y|n; a, \beta) = n^2 a\beta e^{-n(ax+\beta y)}, \quad x, y \geq 0$$

και άρα,

$$f(x, y, n; \theta) = n^2 a\beta(1-p)p^{n-1} e^{-n(ax+\beta y)}, \quad x, y \geq 0.$$

Η από κοινού κατανομή των X και Y , υπολογίζεται σύμφωνα με τα ακόλουθα:

$$\begin{aligned} f(x, y; \theta) &= \sum_{n=1}^{\infty} n^2 a\beta(1-p)p^{n-1} e^{-n(ax+\beta y)} \\ &= a\beta(1-p)p^{-1} \sum_{n=1}^{\infty} n^2 \{pe^{-(ax+\beta y)}\}^n \\ &= a\beta(1-p)p^{-1} \left[\sum_{n=1}^{\infty} n(n-1) \{pe^{-(ax+\beta y)}\}^n + \sum_{n=1}^{\infty} n \{pe^{-(ax+\beta y)}\}^n \right] \\ &= a\beta(1-p)pe^{-2(ax+\beta y)} \sum_{n=1}^{\infty} n(n-1) \{pe^{-(ax+\beta y)}\}^{n-2} \\ &\quad + a\beta(1-p)e^{-(ax+\beta y)} \sum_{n=1}^{\infty} n \{pe^{-(ax+\beta y)}\}^{n-1} \\ &= a\beta(1-p)pe^{-2(ax+\beta y)} \left(\sum_{n=0}^{\infty} z^n \right)'' \Big|_{z=pe^{-(ax+\beta y)}} \end{aligned}$$



$$\begin{aligned}
& + a\beta(1-p)e^{-(ax+\beta y)} \left(\sum_{n=0}^{\infty} z^n \right)' \Big|_{z=pe^{-(ax+\beta y)}} \\
& = \frac{2a\beta(1-p)pe^{-2(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^3} + \frac{a\beta(1-p)e^{-(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^2} \\
& = \frac{a\beta(1-p)e^{-(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^2} \left[\frac{2pe^{-(ax+\beta y)}}{1-pe^{-(ax+\beta y)}} + 1 \right] \\
& = \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1+pe^{-(ax+\beta y)}\}}{\{1-pe^{-(ax+\beta y)}\}^3}, \quad x, y \geq 0.
\end{aligned}$$

4.2.1 Περιθώριες και δεσμευμένες κατανομές

Η περιθώρια κατανομή της X , αποκτάται με την ακόλουθη διαδικασία:

$$\begin{aligned}
f(x; \theta_1) &= \int_0^{\infty} \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1+pe^{-(ax+\beta y)}\}}{\{1-pe^{-(ax+\beta y)}\}^3} dy \\
&= -\frac{a(1-p)}{2p} \int_0^{\infty} d \left[\frac{1+p^2e^{-2(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^2} \right] \\
&= \frac{a(1-p)}{2p} \left[\frac{1+p^2e^{-2ax}}{(1-pe^{-ax})^2} - 1 \right] \\
&= \frac{a(1-p)e^{-ax}}{(1-pe^{-ax})^2}, \quad x \geq 0, \tag{4.5}
\end{aligned}$$

όπου $\theta_1 = (a, p)$. Επομένως, η X ακολουθεί την κατανομή EEG με παραμέτρους μορφής και κλίμακας $1-p$ και a , αντίστοιχα.

Με ανάλογο τρόπο, αποδεικνύεται ότι η Y ακολουθεί την κατανομή EEG , με παραμέτρους μορφής και κλίμακας $1-p$ και β αντίστοιχα,

$$f(y; \theta_2) = \frac{\beta(1-p)e^{-\beta y}}{(1-pe^{-\beta y})^2}, \quad y \geq 0, \tag{4.6}$$



όπου $\theta_2 = (\beta, p)$.

Από τις σχέσεις (4.4)-(4.6), οι δεσμευμένες κατανομές των $X|Y$ και $Y|X$ υπολογίζονται αντίστοιχα ως

$$f(x|y;\theta) = \frac{ae^{-ax}(1-pe^{-\beta y})^2\{1+pe^{-(ax+\beta y)}\}}{\{1-pe^{-(ax+\beta y)}\}^3}, \quad x \geq 0,$$

$$f(y|x;\theta) = \frac{\beta e^{-\beta y}(1-pe^{-ax})^2\{1+pe^{-(ax+\beta y)}\}}{\{1-pe^{-(ax+\beta y)}\}^3}, \quad y \geq 0.$$

4.2.2 Οι συναρτήσεις επιβίωσης και κινδύνου

Η συνάρτηση επιβίωσης της διδιάστατης κατανομής EEG-2, υπολογίζεται από το διπλό ολοκλήρωμα:

$$\begin{aligned} S(x, y; \theta) &= \int_x^\infty \int_y^\infty \frac{a\beta(1-p)e^{-(ax+\beta y)}\{1+pe^{-(ax+\beta y)}\}}{\{1-pe^{-(ax+\beta y)}\}^3} dy dx \\ &= - \int_x^\infty \frac{a(1-p)}{2p} \int_y^\infty d \left[\frac{1+p^2 e^{-2(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^2} \right] dx \\ &= \frac{a(1-p)}{2p} \int_x^\infty \left[\frac{1+p^2 e^{-2(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^2} - 1 \right] dx \\ &= a(1-p) \int_x^\infty \frac{e^{-(ax+\beta y)}}{\{1-pe^{-(ax+\beta y)}\}^2} dx \\ &= -\frac{1-p}{p} \int_x^\infty d \left[\{1-pe^{-(ax+\beta y)}\}^{-1} \right] \\ &= \frac{1-p}{p} \left\{ \frac{1}{1-pe^{-(ax+\beta y)}} - 1 \right\} \\ &= \frac{(1-p)e^{-(ax+\beta y)}}{1-pe^{-(ax+\beta y)}}, \quad x, y \geq 0 \end{aligned}$$

και κατά συνέπεια, η αντίστοιχη συνάρτηση κινδύνου είναι



$$h(x, y; \theta) = \frac{a\beta \{1 + pe^{-(ax+\beta y)}\}}{\{1 - pe^{-(ax+\beta y)}\}^2}, \quad x, y \geq 0.$$

Η δεσμευμένη συνάρτηση επιβίωσης της $X | Y$, δίνεται από τη σχέση:

$$S(x | y; \theta) = \frac{e^{-ax} (1 - pe^{-\beta y})^2}{\{1 - pe^{-(ax+\beta y)}\}^2}, \quad x \geq 0$$

ενώ η αντίστοιχη συνάρτηση κινδύνου, είναι

$$h(x | y; \theta) = \frac{a \{1 + pe^{-(ax+\beta y)}\}}{1 - pe^{-(ax+\beta y)}}, \quad x \geq 0.$$

Η τελευταία είναι φθίνουσα συνάρτηση του x για κάθε y αφού

$$h'(x | y; \theta) = \frac{-2a^2 pe^{-(ax+\beta y)}}{\{1 - pe^{-(ax+\beta y)}\}^2} < 0.$$

Τα αντίστοιχα αποτελέσματα για την κατανομή της $Y | X$, προκύπτουν με εναλλαγή των x και y και των a και β στις τρεις παραπάνω σχέσεις.

Η δεσμευμένη συνάρτηση κινδύνου $\bar{h}(x | y)$ της σχέσης (1.5), υπολογίζεται ως

$$\bar{h}(x | y; \theta) = \frac{a}{1 - pe^{-(ax+\beta y)}}, \quad x \geq 0$$

και είναι φθίνουσα συνάρτηση του x για κάθε y αφού

$$\bar{h}'(x | y; \theta) = \frac{-a^2 pe^{-(ax+\beta y)}}{\{1 - pe^{-(ax+\beta y)}\}^2} < 0.$$

Η δεύτερη από τις δεσμευμένες συναρτήσεις κινδύνου της σχέσης (1.5), $\bar{h}(y | x)$ και η μονοτονία της, προκύπτουν με εναλλαγή των a και β στις δύο τελευταίες σχέσεις.

Οι περιθώριες συναρτήσεις επιβίωσης των X και Y , είναι

$$S(x; \theta_1) = \frac{(1-p)e^{-ax}}{1 - pe^{-ax}}, \quad x \geq 0,$$

$$S(y; \theta_2) = \frac{(1-p)e^{-\beta y}}{1 - pe^{-\beta y}}, \quad y \geq 0$$

και οι αντίστοιχες συναρτήσεις κινδύνου, δίνονται από τις σχέσεις:



$$h(x; \theta_1) = \frac{a}{1 - pe^{-ax}}, \quad x \geq 0,$$

$$h(y; \theta_2) = \frac{\beta}{1 - pe^{-\beta y}}, \quad y \geq 0.$$

Αφού η παράμετρος μορφής των περιθώριων κατανομών *EEG* είναι $0 < 1 - p < 1$, οι περιθώριες συναρτήσεις κινδύνου είναι και οι δύο φθίνουσες.

4.2.3 Ο συντελεστής συσχέτισης

Στη συνέχεια, υπολογίζεται ο συντελεστής συσχέτισης για τη διδιάστατη κατανομή της σχέσης (4.4). Οι ροπές και οι διακυμάνσεις των X και Y με αντίστοιχες πυκνότητες αυτές των σχέσεων (4.5) και (4.6), υπολογίζονται λαμβάνοντας υπόψιν τα αποτελέσματα της ενότητας 2.4.

Η αναμενόμενη τιμή του γινομένου XY , προκύπτει ως εξής:

$$\begin{aligned} E(XY; \theta) &= \int_0^{\infty} \int_0^{\infty} \frac{(1-p)e^{-(ax+\beta y)}}{1-pe^{-(ax+\beta y)}} dx dy \\ &= \int_0^{\infty} \frac{1-p}{ap} \int_0^{\infty} d[\log\{1-pe^{-(ax+\beta y)}\}] dy \\ &= -\frac{1-p}{ap} \int_0^{\infty} \log(1-pe^{-\beta y}) dy \\ &= \frac{1-p}{ap} \left\{ -y \log(1-pe^{-\beta y}) \Big|_0^{\infty} + \int_0^{\infty} \frac{p\beta y e^{-\beta y}}{1-pe^{-\beta y}} dy \right\} \\ &= \frac{\beta}{a} \int_0^{\infty} \frac{y(1-p)e^{-\beta y}}{1-pe^{-\beta y}} dy \\ &= \frac{\beta}{2a} E\{Y^2 | Y \sim EEG(\beta, 1-p)\} \\ &= \frac{(1-p)}{a\beta} \Phi(p, 2, 1) \end{aligned}$$

ενώ από τη σχέση (2.3), οι αναμενόμενες τιμές των X και Y , υπολογίζονται αντίστοιχα ως



$$E(X; \theta_1) = \frac{-(1-p)\log(1-p)}{ap},$$

$$E(Y; \theta_2) = \frac{-(1-p)\log(1-p)}{\beta p}.$$

Άρα, η συνδιακύμανση των X και Y , είναι

$$\text{Cov}(X, Y; \theta) = \frac{1-p}{a\beta} \left\{ \Phi(p, 2, 1) - \frac{(1-p)\log^2(1-p)}{p^2} \right\}.$$

ενώ από τη σχέση (2.4), προκύπτουν οι διακυμάνσεις των X και Y ,

$$\text{Var}(X; \theta_1) = \frac{1-p}{a^2} \left\{ 2\Phi(p, 2, 1) - \frac{(1-p)\log^2(1-p)}{p^2} \right\},$$

$$\text{Var}(Y; \theta_2) = \frac{1-p}{\beta^2} \left\{ 2\Phi(p, 2, 1) - \frac{(1-p)\log^2(1-p)}{p^2} \right\}.$$

Αντικαθιστώντας τις εκφράσεις της συνδιακύμανσης και των διακυμάνσεων των X και Y στη σχέση (1.7), προκύπτει ότι

$$\rho(X, Y; \theta) = \frac{p^2\Phi(p, 2, 1) - (1-p)\log^2(1-p)}{2p^2\Phi(p, 2, 1) - (1-p)\log^2(1-p)}.$$

4.3 Διδιάστατη $EEG-3$

Κάτω από τις υποθέσεις του Θεωρήματος 4.2 για τις υπεισερχόμενες τυχαίες μεταβλητές και τις κατανομές τους, η από κοινού κατανομή των $\min\{X_i\}_{i=1}^N$ και $\max\{Y_i\}_{i=1}^N$ είναι η διδιάστατη EEG που παρουσιάζεται στη συνέχεια.

Θεώρημα 4.3

Έστω $\{X_i\}_{i=1}^N$ και $\{Y_i\}_{i=1}^N$ τυχαία δείγματα από την Εκθετική κατανομή, με αντίστοιχες πυκνότητες

$$f(x; a) = ae^{-ax}, \quad x \geq 0,$$

$$f(y; \beta) = \beta e^{-\beta y}, \quad y \geq 0,$$

όπου $a, \beta \in R^+$ και N γεωμετρική τυχαία μεταβλητή με συνάρτηση πιθανότητας



$$f(n; p) = (1-p)p^{n-1}, \quad n \in N,$$

με $p \in (0,1)$. Αν $X = \min\{X_i\}_{i=1}^N$, $Y = \max\{Y_i\}_{i=1}^N$ και οι τυχαίες μεταβλητές $X|N$ και $Y|N$ είναι ανεξάρτητες, τότε:

$$f(x, y; \theta) = \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1 + p(1-e^{-\beta y})e^{-ax}\}}{\{1 - p(1-e^{-\beta y})e^{-ax}\}^3}, \quad x, y \geq 0, \quad (4.7)$$

όπου $\theta = (a, \beta, p)$.

Απόδειξη

Οι κατανομές των $X|N$ και $Y|N$, είναι αντίστοιχα:

$$f(x|n; a) = nae^{-nax}, \quad x \geq 0,$$

$$f(y|n; \beta) = n\beta(1-e^{-\beta y})^{n-1}e^{-\beta y}, \quad y \geq 0$$

και λόγω της ανεξαρτησίας τους,

$$f(x, y|n; a, \beta) = n^2 a\beta(1-e^{-\beta y})^{n-1}e^{-(nax+\beta y)}, \quad x, y \geq 0.$$

Άρα,

$$f(x, y, n; \theta) = n^2 a\beta(1-p)p^{n-1}(1-e^{-\beta y})^{n-1}e^{-(nax+\beta y)}, \quad x, y \geq 0$$

και συνεπώς, η από κοινού κατανομή των X και Y υπολογίζεται ως

$$\begin{aligned} f(x, y; \theta) &= \sum_{n=1}^{\infty} n^2 a\beta(1-p)p^{n-1}(1-e^{-\beta y})^{n-1}e^{-(nax+\beta y)} \\ &= \frac{a\beta(1-p)e^{-\beta y}}{p(1-e^{-\beta y})} \sum_{n=1}^{\infty} n^2 \{p(1-e^{-\beta y})e^{-ax}\}^n \\ &= \frac{a\beta(1-p)e^{-\beta y}}{p(1-e^{-\beta y})} \left[\sum_{n=1}^{\infty} n(n-1) \{p(1-e^{-\beta y})e^{-ax}\}^n + \sum_{n=1}^{\infty} n \{p(1-e^{-\beta y})e^{-ax}\}^n \right] \\ &= a\beta(1-p)p(1-e^{-\beta y})e^{-(2ax+\beta y)} \sum_{n=1}^{\infty} n(n-1) \{p(1-e^{-\beta y})e^{-ax}\}^{n-2} \\ &\quad + a\beta(1-p)e^{-(ax+\beta y)} \sum_{n=1}^{\infty} n \{p(1-e^{-\beta y})e^{-ax}\}^{n-1} \\ &= a\beta(1-p)p(1-e^{-\beta y})e^{-(2ax+\beta y)} \left(\sum_{n=0}^{\infty} z^n \right)'' \Big|_{z=p(1-e^{-\beta y})e^{-ax}} \end{aligned}$$



$$\begin{aligned}
& + a\beta(1-p)e^{-(ax+\beta y)} \left(\sum_{n=0}^{\infty} z^n \right)' \Big|_{z=p(1-e^{-\beta y})e^{-ax}} \\
& = \frac{2a\beta(1-p)p(1-e^{-\beta y})e^{-(2ax+\beta y)}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} + \frac{a\beta(1-p)e^{-(ax+\beta y)}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^2} \\
& = \frac{a\beta(1-p)e^{-(ax+\beta y)}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^2} \left\{ \frac{2p(1-e^{-\beta y})e^{-ax}}{1-p(1-e^{-\beta y})e^{-ax}} + 1 \right\} \\
& = \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1+p(1-e^{-\beta y})e^{-ax}\}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3}, \quad x, y \geq 0.
\end{aligned}$$

4.3.1 Περιθώριες και δεσμευμένες κατανομές

Η περιθώρια κατανομή της X , αποκτάται σύμφωνα με την ακόλουθη διαδικασία:

$$\begin{aligned}
f(x; \theta_1) &= \int_0^{\infty} \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1+p(1-e^{-\beta y})e^{-ax}\}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} dy \\
&= a\beta(1-p)e^{-ax}(1+pe^{-ax}) \int_0^{\infty} \frac{e^{-\beta y}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} dy \\
&\quad - a\beta(1-p)pe^{-2ax} \int_0^{\infty} \frac{e^{-2\beta y}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} dy \\
&= \frac{a(1-p)(1+pe^{-ax})}{2p} \int_0^{\infty} d \left[\{1-p(1-e^{-\beta y})e^{-ax}\}^{-2} \right] \\
&\quad + \frac{a(1-p)pe^{-2ax}}{2(1-pe^{-ax})} \int_0^{\infty} d \left[\frac{e^{-2\beta y}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^2} \right] \\
&= \frac{a(1-p)(1+pe^{-ax})}{2p} \left\{ \frac{1}{(1-pe^{-ax})^2} - 1 \right\} - \frac{a(1-p)pe^{-2ax}}{2(1-pe^{-ax})} \\
&= \frac{a(1-p)}{2p} \left\{ \frac{1+pe^{-ax}}{(1-pe^{-ax})^2} - (1+pe^{-ax}) - \frac{p^2 e^{-2ax}}{1-pe^{-ax}} \right\}
\end{aligned}$$



$$\begin{aligned}
 &= \frac{a(1-p)}{2p} \left\{ \frac{1+pe^{-ax}}{(1-pe^{-ax})^2} - \frac{1}{1-pe^{-ax}} \right\} \\
 &= \frac{a(1-p)e^{-ax}}{(1-pe^{-ax})^2}, \quad x \geq 0, \tag{4.8}
 \end{aligned}$$

όπου $\theta_1 = (a, p)$ ενώ η περιθώρια κατανομή της Y , υπολογίζεται ως εξής:

$$\begin{aligned}
 f(y; \theta_2) &= \int_0^{\infty} \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1+p(1-e^{-\beta y})e^{-ax}\}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} dx \\
 &= a\beta(1-p)e^{-\beta y} \int_0^{\infty} \frac{e^{-ax}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} dx \\
 &\quad + a\beta(1-p)pe^{-\beta y}(1-e^{-\beta y}) \int_0^{\infty} \frac{e^{-2ax}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^3} dx \\
 &= -\frac{\beta(1-p)e^{-\beta y}}{2p(1-e^{-\beta y})} \int_0^{\infty} d \left[\{1-p(1-e^{-\beta y})e^{-ax}\}^{-2} \right] \\
 &\quad - \frac{\beta(1-p)pe^{-\beta y}(1-e^{-\beta y})}{2} \int_0^{\infty} d \left[\frac{e^{-2ax}}{\{1-p(1-e^{-\beta y})e^{-ax}\}^2} \right] \\
 &= \frac{\beta(1-p)e^{-\beta y}}{2p(1-e^{-\beta y})} \left[\frac{1}{\{1-p(1-e^{-\beta y})\}^2} - 1 \right] + \frac{\beta(1-p)pe^{-\beta y}(1-e^{-\beta y})}{2\{1-p(1-e^{-\beta y})\}^2} \\
 &= \frac{\beta(1-p)e^{-\beta y}}{2p(1-e^{-\beta y})\{1-p(1-e^{-\beta y})\}^2} \left[1 - \{1-p(1-e^{-\beta y})\}^2 + p^2(1-e^{-\beta y})^2 \right] \\
 &= \frac{\beta(1-p)e^{-\beta y}}{\{1-p(1-e^{-\beta y})\}^2} \\
 &= \frac{\beta(1-p)^{-1}e^{-\beta y}}{\{1+p(1-p)^{-1}e^{-\beta y}\}^2}, \quad y \geq 0, \tag{4.9}
 \end{aligned}$$

όπου $\theta_2 = (\beta, p)$.



Χρησιμοποιώντας τις σχέσεις (4.7)-(4.9), οι αντίστοιχες πυκνότητες των τυχαίων μεταβλητών $X|Y$ και $Y|X$, υπολογίζονται ως

$$f(x|y;\theta) = \frac{ae^{-ax} \{1 + p(1 - e^{-\beta y})e^{-ax}\} \{1 - p(1 - e^{-\beta y})\}^2}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^3}, \quad x \geq 0,$$

$$f(y|x;\theta) = \frac{\beta e^{-\beta y} \{1 + p(1 - e^{-\beta y})e^{-ax}\} (1 - pe^{-ax})^2}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^3}, \quad y \geq 0.$$

4.3.2 Οι συναρτήσεις επιβίωσης και κινδύνου

Η διδιάστατη συνάρτηση επιβίωσης της κατανομής στη σχέση (4.7), υπολογίζεται από το διπλό ολοκλήρωμα:

$$\begin{aligned} S(x, y; \theta) &= \int_x^\infty \int_y^\infty \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1 + p(1 - e^{-\beta y})e^{-ax}\}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^3} dy dx \\ &= \int_y^\infty a\beta(1-p)e^{-\beta y} \left[\int_x^\infty \frac{e^{-ax}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^3} dx \right. \\ &\quad \left. + p(1 - e^{-\beta y}) \int_x^\infty \frac{e^{-2ax}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^3} dx \right] dy \\ &= \int_y^\infty a\beta(1-p)e^{-\beta y} \left[-\frac{1}{2ap(1 - e^{-\beta y})} \int_x^\infty d\left\{1 - p(1 - e^{-\beta y})e^{-ax}\right\}^{-2} \right. \\ &\quad \left. - \frac{p(1 - e^{-\beta y})}{2a} \int_x^\infty d\left[\frac{e^{-2ax}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2} \right] \right] dy \\ &= \int_y^\infty a\beta(1-p)e^{-\beta y} \left[\frac{1}{2ap(1 - e^{-\beta y})} \left[\frac{1}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2} - 1 \right] \right. \\ &\quad \left. + \frac{pe^{-2ax}(1 - e^{-\beta y})}{2a\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2} \right] dy \end{aligned}$$



$$\begin{aligned}
 &= \int_y^{\infty} \frac{\beta(1-p)e^{-\beta y} \left[1 - \left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^2 + p^2(1-e^{-\beta y})^2 e^{-2ax} \right]}{2p(1-e^{-\beta y}) \left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^2} dy \\
 &= \int_y^{\infty} \frac{\beta(1-p)e^{-(ax+\beta y)}}{\left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^2} dy \\
 &= \frac{1-p}{p} \int_y^{\infty} d \left[\left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^{-1} \right] \\
 &= \frac{1-p}{p} \left\{ \frac{1}{1-pe^{-ax}} - \frac{1}{1-p(1-e^{-\beta y})e^{-ax}} \right\} \\
 &= \frac{(1-p)e^{-(ax+\beta y)}}{(1-pe^{-ax}) \left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}}, \quad x, y \geq 0.
 \end{aligned}$$

Κατά συνέπεια, η συνάρτηση κινδύνου της διδιάστατης κατανομής EEG-3, είναι

$$h(x, y; \theta) = \frac{a\beta(1-pe^{-ax}) \left\{ 1 + p(1-e^{-\beta y})e^{-ax} \right\}}{\left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^2}, \quad x, y \geq 0.$$

Οι δεσμευμένες συναρτήσεις επιβίωσης και κινδύνου της $X | Y$, δίνονται από τις σχέσεις:

$$S(x | y; \theta) = \frac{e^{-ax} \left\{ 1 - p(1-e^{-\beta y}) \right\}^2}{\left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^2}, \quad x \geq 0,$$

$$h(x | y; \theta) = \frac{a \left\{ 1 + p(1-e^{-\beta y})e^{-ax} \right\}}{1 - p(1-e^{-\beta y})e^{-ax}}, \quad x \geq 0$$

ενώ οι αντίστοιχες συναρτήσεις της κατανομής της $Y | X$, είναι

$$S(y | x; \theta) = \frac{e^{-\beta y} \left\{ 1 - p^2(1-e^{-\beta y})e^{-2ax} \right\}}{\left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\}^2}, \quad y \geq 0,$$

$$h(y | x; \theta) = \frac{\beta(1-pe^{-ax})^2 \left\{ 1 + p(1-e^{-\beta y})e^{-ax} \right\}}{\left\{ 1 - p(1-e^{-\beta y})e^{-ax} \right\} \left\{ 1 - p^2(1-e^{-\beta y})e^{-2ax} \right\}}, \quad y \geq 0.$$

Επίσης, αφού



$$h'(x|y;\theta) = \frac{-2a^2 p e^{-ax} (1 - e^{-\beta y})}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2} < 0,$$

η δεσμευμένη συνάρτηση κινδύνου $h(x|y)$, είναι φθίνουσα συνάρτηση του x για κάθε τιμή του y ενώ εξετάζοντας το πρόσημο της παραγώγου της $h(y|x)$ ως προς y ,

$$h'(y|x;\theta) = \frac{\beta^2 p(1 - p e^{-ax})^2 e^{-(ax+\beta y)}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2 \{1 - p^2(1 - e^{-\beta y})e^{-2ax}\}^2} \left[2\{1 - p^2(1 - e^{-\beta y})e^{-2ax}\} + \{1 - p^2(1 - e^{-\beta y})^2 e^{-2ax}\} p e^{-ax} \right] > 0,$$

προκύπτει ότι η $h(y|x)$, είναι αύξουσα συνάρτηση του y για κάθε τιμή του x .

Οι δεσμευμένες συναρτήσεις κινδύνου της σχέσης (1.5), υπολογίζονται ως

$$\bar{h}(x|y;\theta) = \frac{a\{1 - p^2(1 - e^{-\beta y})e^{-2ax}\}}{(1 - p e^{-ax})\{1 - p(1 - e^{-\beta y})e^{-ax}\}}, \quad x \geq 0,$$

$$\bar{h}(y|x;\theta) = \frac{\beta(1 - p e^{-ax})}{1 - p(1 - e^{-\beta y})e^{-ax}}, \quad y \geq 0.$$

Η πρώτη είναι φθίνουσα συνάρτηση του x για κάθε y αφού

$$\bar{h}'(x|y;\theta) = -\frac{a^2 p(1 - e^{-\beta y})e^{-ax}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2} - \frac{a^2 p e^{-ax}}{(1 - p e^{-ax})^2} < 0$$

ενώ η δεύτερη, είναι αύξουσα συνάρτηση του y για κάθε τιμή του x αφού

$$\bar{h}'(y|x;\theta) = \frac{\beta^2 p(1 - p e^{-ax})e^{-(ax+\beta y)}}{\{1 - p(1 - e^{-\beta y})e^{-ax}\}^2} > 0.$$

Οι συναρτήσεις επιβίωσης που αντιστοιχούν στις πυκνότητες των σχέσεων (4.8) και (4.9), δίνονται από τις σχέσεις:

$$S(x;\theta_1) = \frac{(1-p)e^{-ax}}{1 - p e^{-ax}}, \quad x \geq 0,$$

$$S(y;\theta_2) = \frac{(1-p)^{-1} e^{-\beta y}}{1 + p(1-p)^{-1} e^{-\beta y}}, \quad y \geq 0$$

και κατά συνέπεια, η αναλυτική έκφραση των περιθωρίων συναρτήσεων κινδύνου των X και



Y , είναι

$$h(x; \theta_1) = \frac{a}{1 - pe^{-ax}}, \quad x \geq 0,$$

$$h(y; \theta_2) = \frac{\beta}{1 + p(1-p)^{-1}e^{-\beta y}}, \quad y \geq 0.$$

Όμως, αφού $1 - p < 1$, η πρώτη από τις περιθώριες συναρτήσεις κινδύνου είναι φθίνουσα συνάρτηση του x ενώ η δεύτερη αύξουσα συνάρτηση του y .

4.3.3 Ο συντελεστής συσχέτισης

Για τον υπολογισμό της αναμενόμενης τιμής του γινομένου XY , είναι

$$\begin{aligned} E(XY; \theta) &= \int_0^\infty \int_0^\infty \frac{(1-p)e^{-(ax+\beta y)}}{(1-pe^{-ax})\{1-p(1-e^{-\beta y})e^{-ax}\}} dy dx \\ &= \int_0^\infty \frac{(1-p)e^{-ax}}{1-pe^{-ax}} \int_0^\infty \frac{e^{-\beta y}}{1-p(1-e^{-\beta y})e^{-ax}} dy dx \\ &= - \int_0^\infty \frac{1-p}{\beta p(1-pe^{-ax})} \int_0^\infty d[\log\{1-p(1-e^{-\beta y})e^{-ax}\}] dx \\ &= - \frac{1-p}{\beta p} \int_0^\infty \frac{\log(1-pe^{-ax})}{1-pe^{-ax}} dx \\ &\stackrel{pe^{-ax}=z}{=} - \frac{1-p}{\alpha\beta p} \int_0^p \frac{\log(1-z)}{(1-z)z} dz \\ &= - \frac{1-p}{\alpha\beta p} \left\{ \int_0^p \frac{\log(1-z)}{1-z} dz + \int_0^p \frac{\log(1-z)}{z} dz \right\} \\ &= - \frac{1-p}{\alpha\beta p} \left\{ - \frac{\log^2(1-z)}{2} \Big|_0^p - \int_0^p \sum_{v=1}^\infty \frac{z^{v-1}}{v} dz \right\} \\ &= - \frac{1-p}{\alpha\beta p} \left\{ - \frac{\log^2(1-p)}{2} - \sum_{v=1}^\infty \frac{p^v}{v^2} \right\} \end{aligned}$$



$$= \frac{1-p}{a\beta p} \left\{ \frac{\log^2(1-p)}{2} + Q(p,2) \right\},$$

όπου $Q(p,2)$ $p \in (0,1)$, είναι η διλογαριθμική συνάρτηση, η οποία συνδέεται με τη συνάρτηση του Lerch, μέσω της σχέσης:

$$Q(p,2) = p\Phi(p,2,1),$$

(Magnus et al., 1966, σελ. 33). Άρα,

$$E(XY; \theta) = \frac{1-p}{a\beta p} \left\{ \frac{\log^2(1-p)}{2} + p\Phi(p,2,1) \right\}.$$

Από τη σχέση (2.3), οι αναμενόμενες τιμές των X και Y με σ.π.π. αυτές των σχέσεων (4.8) και (4.9), υπολογίζονται αντίστοιχα ως

$$E(X; \theta_1) = \frac{-(1-p)\log(1-p)}{ap},$$

$$E(Y; \theta_2) = \frac{-\log(1-p)}{\beta p}.$$

Άρα, η συνδιακύμανση των X και Y , δίνεται από τη σχέση:

$$\text{Cov}(X, Y; \theta) = \frac{1-p}{a\beta p} \left\{ p\Phi(p,2,1) + \frac{(p-2)\log^2(1-p)}{2p} \right\}.$$

Επίσης, από τη σχέση (2.4) προκύπτει ότι

$$\text{Var}(X; \theta_1) = \frac{1-p}{a^2} \left\{ 2\Phi(p,2,1) - \frac{(1-p)\log^2(1-p)}{p^2} \right\},$$

$$\text{Var}(Y; \theta_2) = \frac{1}{(1-p)\beta^2} \left\{ 2\Phi\left(-\frac{p}{1-p}, 2, 1\right) - \frac{(1-p)\log^2(1-p)}{p^2} \right\}.$$

Επομένως, χρησιμοποιώντας τα παραπάνω αποτελέσματα στη σχέση (1.7), ο συντελεστής συσχέτισης του Pearson για τη διδιάστατη κατανομή EEG-3, υπολογίζεται ως

$$\rho(X, Y; \theta)$$

$$= \frac{(1-p) \left\{ 2p^2\Phi(p,2,1) + (p-2)\log^2(1-p) \right\}}{2 \sqrt{\left\{ 2p^2\Phi(p,2,1) - (1-p)\log^2(1-p) \right\} \left\{ 2p^2\Phi\left(-\frac{p}{1-p}, 2, 1\right) - (1-p)\log^2(1-p) \right\}}}$$



4.4 Διδιάστατη EEG - 4

Η κατανομή EEG-4, περιγράφει την από κοινού συμπεριφορά των μέγιστων διατεταγμένων στατιστικών από δύο εκθετικά δείγματα ίσου μεγέθους, όταν το πλήθος των τυχαίων μεταβλητών κάθε δείγματος, είναι γεωμετρική τυχαία μεταβλητή.

Θεώρημα 4.4

Έστω $\{X_i\}_{i=1}^N$ και $\{Y_i\}_{i=1}^N$ τυχαία δείγματα από δύο εκθετικούς πληθυσμούς, με αντίστοιχες πυκνότητες

$$f(x; a) = ae^{-ax}, \quad x \geq 0,$$

$$f(y; \beta) = \beta e^{-\beta y}, \quad y \geq 0,$$

όπου $a, \beta \in \mathbb{R}^+$ και N γεωμετρική τυχαία μεταβλητή με συνάρτηση πιθανότητας

$$f(n; p) = (1-p)p^{n-1}, \quad n \in \mathbb{N},$$

με $p \in (0,1)$. Αν η $X|N$ είναι ανεξάρτητη της $Y|N$, όπου $X = \max\{X_i\}_{i=1}^N$ και

$Y = \max\{Y_i\}_{i=1}^N$, τότε:

$$f(x, y; \theta) = \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1 + p(1-e^{-ax})(1-e^{-\beta y})\}}{\{1 - p(1-e^{-ax})(1-e^{-\beta y})\}^3}, \quad x, y \geq 0, \quad (4.10)$$

όπου $\theta = (a, \beta, p)$.

Απόδειξη

Είναι

$$f(x|n; a) = na(1-e^{-ax})^{n-1}e^{-ax}, \quad x \geq 0,$$

$$f(y|n; \beta) = n\beta(1-e^{-\beta y})^{n-1}e^{-\beta y}, \quad y \geq 0$$

και αφού οι $X|N$ και $Y|N$ είναι ανεξάρτητες,

$$f(x, y|n; a, \beta) = n^2 a\beta \{(1-e^{-ax})(1-e^{-\beta y})\}^{n-1} e^{-(ax+\beta y)}, \quad x, y \geq 0.$$

Άρα, η από κοινού κατανομή των X, Y, N είναι

$$f(x, y, n; \theta) = n^2 a\beta(1-p)p^{n-1} \{(1-e^{-ax})(1-e^{-\beta y})\}^{n-1} e^{-(ax+\beta y)}, \quad x, y \geq 0$$

και επομένως,



$$\begin{aligned}
f(x, y; \theta) &= \sum_{n=1}^{\infty} n^2 a\beta(1-p)p^{n-1} \left\{ (1-e^{-ax})(1-e^{-\beta y}) \right\}^{n-1} e^{-(ax+\beta y)} \\
&= a\beta(1-p)e^{-(ax+\beta y)} \sum_{n=1}^{\infty} n^2 \left\{ p(1-e^{-ax})(1-e^{-\beta y}) \right\}^{n-1} \\
&= a\beta(1-p)e^{-(ax+\beta y)} \sum_{n=1}^{\infty} n(n-1) \left\{ p(1-e^{-ax})(1-e^{-\beta y}) \right\}^{n-1} \\
&\quad + a\beta(1-p)e^{-(ax+\beta y)} \sum_{n=1}^{\infty} n \left\{ p(1-e^{-ax})(1-e^{-\beta y}) \right\}^{n-1} \\
&= a\beta(1-p)pe^{-(ax+\beta y)} (1-e^{-ax})(1-e^{-\beta y}) \sum_{n=1}^{\infty} n(n-1) \left\{ p(1-e^{-ax})(1-e^{-\beta y}) \right\}^{n-2} \\
&\quad + a\beta(1-p)e^{-(ax+\beta y)} \sum_{n=1}^{\infty} n \left\{ p(1-e^{-ax})(1-e^{-\beta y}) \right\}^{n-1} \\
&= a\beta(1-p)pe^{-(ax+\beta y)} (1-e^{-ax})(1-e^{-\beta y}) \left(\sum_{n=0}^{\infty} z^n \right)'' \Big|_{z=p(1-e^{-ax})(1-e^{-\beta y})} \\
&\quad + a\beta(1-p)e^{-(ax+\beta y)} \left(\sum_{n=0}^{\infty} z^n \right)' \Big|_{z=p(1-e^{-ax})(1-e^{-\beta y})} \\
&= \frac{2a\beta(1-p)pe^{-(ax+\beta y)} (1-e^{-ax})(1-e^{-\beta y})}{\left\{ 1-p(1-e^{-ax})(1-e^{-\beta y}) \right\}^3} + \frac{a\beta(1-p)e^{-(ax+\beta y)}}{\left\{ 1-p(1-e^{-ax})(1-e^{-\beta y}) \right\}^2} \\
&= \frac{a\beta(1-p)e^{-(ax+\beta y)}}{\left\{ 1-p(1-e^{-ax})(1-e^{-\beta y}) \right\}^2} \left\{ \frac{2p(1-e^{-ax})(1-e^{-\beta y})}{1-p(1-e^{-ax})(1-e^{-\beta y})} + 1 \right\} \\
&= \frac{a\beta(1-p)e^{-(ax+\beta y)} \left\{ 1+p(1-e^{-ax})(1-e^{-\beta y}) \right\}}{\left\{ 1-p(1-e^{-ax})(1-e^{-\beta y}) \right\}^3}, \quad x, y \geq 0.
\end{aligned}$$

4.4.1 Περιθώριες και δεσμευμένες κατανομές

Για τον υπολογισμό της περιθώριας κατανομής της X , είναι



$$\begin{aligned}
 f(x; \theta_1) &= \int_0^{\infty} \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1 + p(1-e^{-ax})(1-e^{-\beta y})\}}{\{1 - p(1-e^{-ax})(1-e^{-\beta y})\}^3} dy \\
 &= a\beta(1-p)e^{-ax} \{1 + p(1-e^{-ax})\} \int_0^{\infty} \frac{e^{-\beta y}}{\{1 - p(1-e^{-ax})(1-e^{-\beta y})\}^3} dy \\
 &\quad - a\beta(1-p)pe^{-ax}(1-e^{-ax}) \int_0^{\infty} \frac{e^{-2\beta y}}{\{1 - p(1-e^{-ax})(1-e^{-\beta y})\}^3} dy \\
 &= \frac{a(1-p)e^{-ax} \{1 + p(1-e^{-ax})\}}{2p(1-e^{-ax})} \int_0^{\infty} d \left[\{1 - p(1-e^{-ax})(1-e^{-\beta y})\}^{-2} \right] \\
 &\quad + \frac{a(1-p)pe^{-ax}(1-e^{-ax})}{2\{1 - p(1-e^{-ax})\}} \int_0^{\infty} d \left[\frac{e^{-2\beta y}}{\{1 - p(1-e^{-ax})(1-e^{-\beta y})\}^2} \right] \\
 &= \frac{a(1-p)e^{-ax} \{1 + p(1-e^{-ax})\}}{2p(1-e^{-ax})} \left[\frac{1}{\{1 - p(1-e^{-ax})\}^2} - 1 \right] - \frac{a(1-p)pe^{-ax}(1-e^{-ax})}{2\{1 - p(1-e^{-ax})\}} \\
 &= \frac{a(1-p)e^{-ax} \{1 + p(1-e^{-ax})\} \{2 - p(1-e^{-ax})\}}{2\{1 - p(1-e^{-ax})\}^2} - \frac{a(1-p)pe^{-ax}(1-e^{-ax})}{2\{1 - p(1-e^{-ax})\}} \\
 &= \frac{a(1-p)e^{-ax}}{2\{1 - p(1-e^{-ax})\}^2} \left[\{1 + p(1-e^{-ax})\} \{2 - p(1-e^{-ax})\} - p(1-e^{-ax}) \{1 - p(1-e^{-ax})\} \right] \\
 &= \frac{a(1-p)e^{-ax}}{\{1 - p(1-e^{-ax})\}^2} \\
 &= \frac{a(1-p)^{-1}e^{-ax}}{\{1 + p(1-p)^{-1}e^{-ax}\}^2}, \quad x \geq 0. \tag{4.11}
 \end{aligned}$$

Με ανάλογο τρόπο, αποδεικνύεται ότι η κατανομή της Y δίνεται από τη σχέση:

$$f(y; \theta_2) = \frac{\beta(1-p)^{-1}e^{-\beta y}}{\{1 + p(1-p)^{-1}e^{-\beta y}\}^2}, \quad y \geq 0 \tag{4.12}$$

και επομένως, οι περιθώριες κατανομές EEG, έχουν την ίδια παράμετρο μορφής και



διαφορετική παράμετρο κλίμακας.

Από τις σχέσεις (4.10)-(4.12), οι πυκνότητες των $X|Y$ και $Y|X$ προσδιορίζονται αντίστοιχα ως

$$f(x|y;\theta) = \frac{ae^{-ax} \{1 + p(1 - e^{-ax})(1 - e^{-\beta y})\} \{1 - p(1 - e^{-\beta y})\}^2}{\{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^3}, \quad x, y \geq 0,$$

$$f(y|x;\theta) = \frac{\beta e^{-\beta y} \{1 + p(1 - e^{-ax})(1 - e^{-\beta y})\} \{1 - p(1 - e^{-ax})\}^2}{\{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^3}, \quad x, y \geq 0.$$

4.4.2 Οι συναρτήσεις επιβίωσης και κινδύνου

Η συνάρτηση επιβίωσης της κατανομής που ορίζεται από τη σχέση (4.10), υπολογίζεται σύμφωνα με την ακόλουθη διαδικασία:

$$\begin{aligned} S(x, y; \theta) &= \int_x^\infty \int_y^\infty \frac{a\beta(1-p)e^{-(ax+\beta y)} \{1 + p(1 - e^{-ax})(1 - e^{-\beta y})\}}{\{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^3} dy dx \\ &= \int_x^\infty a\beta(1-p)e^{-ax} \left[\{1 + p(1 - e^{-ax})\} \int_y^\infty \frac{e^{-\beta y}}{\{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^3} dy \right. \\ &\quad \left. - p(1 - e^{-ax}) \int_y^\infty \frac{e^{-2\beta y}}{\{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^3} dy \right] \\ &= \frac{a(1-p)}{2p} \int_x^\infty \frac{e^{-ax} \{1 + p(1 - e^{-ax})\}}{x(1 - e^{-ax}) \{1 - p(1 - e^{-ax})\}^2} dx \\ &\quad - \frac{a(1-p)}{2p} \int_x^\infty \frac{e^{-ax} \{1 + p(1 - e^{-ax})\}}{x(1 - e^{-ax}) \{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^2} dx \\ &\quad - \frac{a(1-p)p}{2} \int_x^\infty \frac{e^{-ax}(1 - e^{-ax})}{\{1 - p(1 - e^{-ax})\} \{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^2} dx. \end{aligned}$$

Όμως,



$$\begin{aligned}
 & \int_0^{\infty} \frac{e^{-ax} \{1 + p(1 - e^{-ax})\}}{x(1 - e^{-ax}) \{1 - p(1 - e^{-ax})\}^2} dx \\
 &= \frac{1}{a} \int_{1-e^{-ax}=z}^1 \frac{1 + pz}{z(1 - pz)^2} dz \\
 &= \frac{1}{a} \left\{ \int_{1-e^{-ax}}^1 \frac{1}{z(1 - pz)^2} dz + \int_{1-e^{-ax}}^1 \frac{p}{(1 - pz)^2} dz \right\} \\
 &= \frac{1}{a} \left\{ \int_{1-e^{-ax}}^1 \frac{dz}{z} + \int_{1-e^{-ax}}^1 \frac{p}{1 - pz} dz + \int_{1-e^{-ax}}^1 \frac{2p}{(1 - pz)^2} dz \right\} \\
 &= \frac{1}{a} \left\{ \log z - \log(1 - pz) + \frac{2}{1 - pz} \right\} \Bigg|_{1-e^{-ax}}^1 \\
 &= \frac{1}{a} \left\{ \log \left(\frac{z}{1 - pz} \right) + \frac{2}{1 - pz} \right\} \Bigg|_{1-e^{-ax}}^1 .
 \end{aligned}$$

Επίσης,

$$\begin{aligned}
 & \int_0^{\infty} \frac{e^{-ax} \{1 + p(1 - e^{-ax})\}}{x(1 - e^{-ax}) \{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}^2} dx \\
 &= \frac{1}{a} \int_{1-e^{-ax}=z}^1 \frac{1 + pz}{z \{1 - pz(1 - e^{-\beta y})\}^2} dz \\
 &= \frac{1}{a} \left[\int_{1-e^{-ax}}^1 \frac{1}{z \{1 - pz(1 - e^{-\beta y})\}^2} dz + \int_{1-e^{-ax}}^1 \frac{p}{\{1 - pz(1 - e^{-\beta y})\}^2} dz \right] \\
 &= \frac{1}{a} \left[\int_{1-e^{-ax}}^1 \frac{dz}{z} + \int_{1-e^{-ax}}^1 \frac{p(1 - e^{-\beta y})}{1 - pz(1 - e^{-\beta y})} dz + \int_{1-e^{-ax}}^1 \frac{p(2 - e^{-\beta y})}{\{1 - pz(1 - e^{-\beta y})\}^2} dz \right] \\
 &= \frac{1}{a} \left[\log z - \log \{1 - pz(1 - e^{-\beta y})\} + \frac{2 - e^{-\beta y}}{(1 - e^{-\beta y}) \{1 - pz(1 - e^{-\beta y})\}} \right] \Bigg|_{1-e^{-ax}}^1
 \end{aligned}$$



$$= \frac{1}{a} \left[\log \left\{ \frac{z}{1-pz(1-e^{-\beta y})} \right\} + \frac{2-e^{-\beta y}}{(1-e^{-\beta y})\{1-pz(1-e^{-\beta y})\}} \right] \Bigg|_{1-e^{-ax}}^1$$

ενώ

$$\begin{aligned} & \int_x^\infty \frac{e^{-ax}(1-e^{-ax})}{\{1-p(1-e^{-ax})\}\{1-p(1-e^{-ax})(1-e^{-\beta y})\}^2} dx \\ &= \frac{1-e^{-ax}=z}{a} \int_{1-e^{-ax}}^1 \frac{z}{(1-pz)\{1-pz(1-e^{-\beta y})\}^2} dz \\ &= \frac{1}{ap} \left[- \int_{1-e^{-ax}}^1 \frac{1}{\{1-pz(1-e^{-\beta y})\}^2} dz + \int_{1-e^{-ax}}^1 \frac{1}{(1-pz)\{1-pz(1-e^{-\beta y})\}^2} dz \right] \\ &= \frac{1}{ape^{-2\beta y}} \left[- \int_{1-e^{-ax}}^1 \frac{e^{-2\beta y}}{\{1-pz(1-e^{-\beta y})\}^2} dz + \int_{1-e^{-ax}}^1 \frac{1}{1-pz} dz \right. \\ & \quad \left. - \int_{1-e^{-ax}}^1 \frac{1-e^{-\beta y}}{1-pz(1-e^{-\beta y})} dz - \int_{1-e^{-ax}}^1 \frac{e^{-\beta y}(1-e^{-\beta y})}{\{1-pz(1-e^{-\beta y})\}^2} dz \right] \\ &= \frac{1}{ape^{-2\beta y}} \left[- \frac{e^{-\beta y}}{p(1-e^{-\beta y})\{1-pz(1-e^{-\beta y})\}} - \frac{\log(1-pz)}{p} + \frac{\log\{1-pz(1-e^{-\beta y})\}}{p} \right] \Bigg|_{1-e^{-ax}}^1 \\ &= \frac{1}{ap^2 e^{-2\beta y}} \left[- \frac{e^{-\beta y}}{(1-e^{-\beta y})\{1-pz(1-e^{-\beta y})\}} + \log \left\{ \frac{1-pz(1-e^{-\beta y})}{1-pz} \right\} \right] \Bigg|_{1-e^{-ax}}^1 \end{aligned}$$

Άρα,

$$\begin{aligned} S(x, y; \theta) &= \frac{1-p}{2p} \left\{ \log \left(\frac{z}{1-pz} \right) + \frac{2}{1-pz} \right\} \Bigg|_{1-e^{-ax}}^1 \\ & \quad - \frac{1-p}{2p} \left[\log \left\{ \frac{z}{1-pz(1-e^{-\beta y})} \right\} + \frac{2-e^{-\beta y}}{(1-e^{-\beta y})\{1-pz(1-e^{-\beta y})\}} \right] \Bigg|_{1-e^{-ax}}^1 \\ & \quad - \frac{1-p}{2p} \left[- \frac{e^{-\beta y}}{(1-e^{-\beta y})\{1-pz(1-e^{-\beta y})\}} + \log \left\{ \frac{1-pz(1-e^{-\beta y})}{1-pz} \right\} \right] \Bigg|_{1-e^{-ax}}^1 \end{aligned}$$



$$\begin{aligned}
 &= \frac{1-p}{p} \left\{ \frac{1}{1-pz} - \frac{1}{1-pz(1-e^{-\beta y})} \right\} \Bigg|_{z=1-e^{-\alpha x}}^1 \\
 &= \frac{(1-p)ze^{-\beta y}}{(1-pz)\{1-pz(1-e^{-\beta y})\}} \Bigg|_{z=1-e^{-\alpha x}}^1 \\
 &= \frac{(1-p)e^{-\beta y}}{(1-p)\{1-p(1-e^{-\beta y})\}} - \frac{(1-p)ze^{-\beta y}}{(1-pz)\{1-pz(1-e^{-\beta y})\}} \Bigg|_{z=1-e^{-\alpha x}} \\
 &= (1-p)e^{-\beta y} \left\{ \frac{1}{(1-p)(1-pw)} - \frac{z}{(1-pz)(1-pzw)} \right\} \Bigg|_{z=1-e^{-\alpha x}}^{w=1-e^{-\beta y}} \\
 &= e^{-\beta y} \left\{ \frac{(1-pz)(1-pzw) - z(1-p)(1-pw)}{(1-pz)(1-pw)(1-pzw)} \right\} \Bigg|_{z=1-e^{-\alpha x}}^{w=1-e^{-\beta y}} \\
 &= e^{-\beta y} \left\{ \frac{1-z-p^2zw(1-z)}{(1-pz)(1-pw)(1-pzw)} \right\} \Bigg|_{z=1-e^{-\alpha x}}^{w=1-e^{-\beta y}} \\
 &= \frac{e^{-\beta y}(1-z)(1-p^2zw)}{(1-pz)(1-pw)(1-pzw)} \Bigg|_{z=1-e^{-\alpha x}}^{w=1-e^{-\beta y}} \\
 &= \frac{e^{-(\alpha x + \beta y)} \{1 - p^2(1 - e^{-\alpha x})(1 - e^{-\beta y})\}}{\{1 - p(1 - e^{-\alpha x})\} \{1 - p(1 - e^{-\beta y})\} \{1 - p(1 - e^{-\alpha x})(1 - e^{-\beta y})\}}, \quad x, y \geq 0.
 \end{aligned}$$

Η αντίστοιχη συνάρτηση κινδύνου, δίνεται από τη σχέση:

$$\begin{aligned}
 &h(x, y; \theta) \\
 &= \frac{a\beta(1-p)\{1-p(1-e^{-\alpha x})\}\{1-p(1-e^{-\beta y})\}\{1+p(1-e^{-\alpha x})(1-e^{-\beta y})\}}{\{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\}^2 \{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})\}}, \quad x, y \geq 0.
 \end{aligned}$$

Η δεσμευμένη συνάρτηση επιβίωσης της $X|Y$, υπολογίζεται ως

$$S(x|y; \theta) = \frac{e^{-\alpha x} \{1 - p^2(1 - e^{-\alpha x})(1 - e^{-\beta y})^2\}}{\{1 - p(1 - e^{-\alpha x})(1 - e^{-\beta y})\}^2}, \quad x \geq 0$$



ενώ η αντίστοιχη συνάρτηση κινδύνου είναι

$$h(x|y;\theta) = \frac{a\{1-p(1-e^{-\beta y})\}^2\{1+p(1-e^{-\alpha x})(1-e^{-\beta y})\}}{\{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\}\{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})^2\}}, \quad x \geq 0.$$

Η δεσμευμένη συνάρτηση κινδύνου της τελευταίας σχέσης, είναι αύξουσα συνάρτηση του x για κάθε τιμή του y αφού

$$h'(x|y;\theta) = \frac{a^2 p e^{-\alpha x} (1-e^{-\beta y}) \{1-p(1-e^{-\beta y})\}^2}{\{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\}^2 \{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})^2\}^2} \\ \times \left[2\{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})^2\} + p(1-e^{-\beta y}) \{1-p^2(1-e^{-\alpha x})^2(1-e^{-\beta y})^2\} \right] > 0.$$

Οι συναρτήσεις επιβίωσης και κινδύνου και η μονοτονία της τελευταίας για την κατανομή της $Y|X$, προκύπτουν από τις τρεις τελευταίες σχέσεις με εναλλαγή των x και y και των a και β .

Επίσης, είναι

$$\bar{h}(x|y;\theta) = \frac{a(1-p)\{1-p(1-e^{-\beta y})\}\{1-p^2(1-e^{-\alpha x})^2(1-e^{-\beta y})\}}{\{1-p(1-e^{-\alpha x})\}\{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\}\{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})\}}, \quad x \geq 0.$$

Για την παράγωγο της τελευταίας ως προς x , ισχύει ότι

$$\frac{\bar{h}'(x|y;\theta)}{a(1-p)\{1-p(1-e^{-\beta y})\}} \\ = \frac{a p e^{-\alpha x} \left[(1-e^{-\beta y}) \{1-p(1-e^{-\alpha x})\}^2 + e^{-\beta y} \right]}{\{1-p(1-e^{-\alpha x})\}^2 \{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\} \{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})\}} \\ + \frac{a p e^{-\alpha x} (1-e^{-\beta y}) \{1-p^2(1-e^{-\alpha x})^2(1-e^{-\beta y})\}}{\{1-p(1-e^{-\alpha x})\} \{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\}^2 \{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})\}} \\ + \frac{a p^2 e^{-\alpha x} (1-e^{-\beta y}) \{1-p^2(1-e^{-\alpha x})^2(1-e^{-\beta y})\}}{\{1-p(1-e^{-\alpha x})\} \{1-p(1-e^{-\alpha x})(1-e^{-\beta y})\} \{1-p^2(1-e^{-\alpha x})(1-e^{-\beta y})\}^2} > 0.$$

Συνεπώς, η $\bar{h}(x|y)$ είναι αύξουσα συνάρτηση του x . Η δεύτερη από τις δεσμευμένες συναρτήσεις κινδύνου της σχέσης (1.5) και η μονοτονία της, προκύπτουν από τις δύο



τελευταίες σχέσεις με εναλλαγή των x και y και των a και β .

Θέτοντας στην από κοινού συνάρτηση επιβίωσης $y = 0$ και $x = 0$, προκύπτουν αντίστοιχα οι συναρτήσεις επιβίωσης των X και Y ,

$$S(x; \theta_1) = \frac{(1-p)^{-1} e^{-ax}}{1 + p(1-p)^{-1} e^{-ax}}, \quad x \geq 0,$$

$$S(y; \theta_2) = \frac{(1-p)^{-1} e^{-\beta y}}{1 + p(1-p)^{-1} e^{-\beta y}}, \quad y \geq 0.$$

Από τις δύο τελευταίες σχέσεις και από τις (4.11) και (4.12), οι αντίστοιχες συναρτήσεις κινδύνου υπολογίζονται ως

$$h(x; \theta_1) = \frac{a}{1 + p(1-p)^{-1} e^{-ax}}, \quad x \geq 0,$$

$$h(y; \theta_2) = \frac{\beta}{1 + p(1-p)^{-1} e^{-\beta y}}, \quad y \geq 0$$

και είναι αύξουσες αφού $(1-p)^{-1} > 1$.

4.4.3 Ο συντελεστής συσχέτισης

Για τον υπολογισμό της αναμενόμενης τιμής του γινομένου XY , ισχύουν τα ακόλουθα:

$$\begin{aligned} E(XY; \theta) &= \int_0^{\infty} \int_0^{\infty} \frac{e^{-(ax+\beta y)} \{1 - p^2(1 - e^{-ax})(1 - e^{-\beta y})\}}{\{1 - p(1 - e^{-ax})\} \{1 - p(1 - e^{-\beta y})\} \{1 - p(1 - e^{-ax})(1 - e^{-\beta y})\}} dx dy \\ &\stackrel{1-e^{-ax}=z}{=} \int_0^1 \int_0^1 \frac{1 - p^2 zw}{a\beta (1-pw)(1-pz)(1-pzw)} dz dw \\ &= \int_0^1 \frac{1}{a\beta(1-pw)} \left\{ \int_0^1 \frac{1-p}{(1-pz)(1-pzw)} dz + \int_0^1 \frac{p}{1-pz} dz \right\} dw \\ &= \int_0^1 \frac{1}{a\beta(1-pw)} \left[\frac{1-p}{1-w} \left\{ \int_0^1 \frac{1}{1-pz} dz - \int_0^1 \frac{w}{1-pzw} dz \right\} + \int_0^1 \frac{p}{1-pz} dz \right] dw \\ &= \int_0^1 \frac{1}{a\beta(1-pw)} \left[\frac{1-p}{p(1-w)} \left\{ -\log(1-pz) + \log(1-pzw) \right\} \right]_0^1 - \log(1-pz) \Big|_0^1 dw \end{aligned}$$



$$\begin{aligned}
&= \frac{1-p}{a\beta p} \int_0^1 \frac{\log(1-pw)}{(1-pw)(1-w)} dw - \frac{\log(1-p)}{a\beta p} \int_0^1 \frac{1}{1-w} dw \\
&= \frac{1}{a\beta p} \left\{ - \int_0^1 \frac{p \log(1-pw)}{1-pw} dw + \int_0^1 \frac{\log(1-pw)}{1-w} dw \right\} - \frac{\log(1-p)}{a\beta p} \int_0^1 \frac{1}{1-w} dw \\
&= \frac{1}{a\beta p} \left\{ \frac{\log^2(1-pw)}{2} \Big|_0^1 - \log(1-pw) \log(1-w) \Big|_0^1 - \int_0^1 \frac{p \log(1-w)}{1-pw} dw \right. \\
&\quad \left. + \log(1-p) \log(1-w) \Big|_0^1 \right\}.
\end{aligned}$$

Όμως, εφαρμόζοντας τον κανόνα του L'Hospital προκύπτει ότι

$$\begin{aligned}
&\lim_{w \rightarrow 1} \left\{ -\log(1-pw) \log(1-w) + \log(1-p) \log(1-w) \right\} \\
&= -\lim_{w \rightarrow 1} \log(1-w) \log\left(\frac{1-pw}{1-p}\right) \quad (\text{μορφή } 0 \cdot \infty) \\
&= -\lim_{w \rightarrow 1} \frac{\log(1-w)}{\log^{-1}\left(\frac{1-pw}{1-p}\right)} \quad (\text{μορφή } \frac{\infty}{\infty}) \\
&= -\lim_{w \rightarrow 1} \frac{-(1-w)^{-1}}{p(1-pw)^{-1} \log^{-2}\left(\frac{1-pw}{1-p}\right)} \\
&= \lim_{w \rightarrow 1} \frac{(1-pw) \log^2\left(\frac{1-pw}{1-p}\right)}{p(1-w)} \quad (\text{μορφή } \frac{0}{0}) \\
&= \lim_{w \rightarrow 1} \frac{-p \log^2\left(\frac{1-pw}{1-p}\right) - 2p \log\left(\frac{1-pw}{1-p}\right)}{-p} = 0.
\end{aligned}$$

Άρα,

$$E(XY; \theta) = \frac{\log^2(1-p)}{2a\beta p} - \frac{1}{a\beta} \int_0^1 \frac{\log(1-w)}{1-pw} dw$$



$$\begin{aligned}
&= \frac{1-w=e^{-x} \log^2(1-p)}{2a\beta p} + \frac{1}{a\beta} \int_0^{\infty} \frac{x e^{-x}}{1-p(1-e^{-x})} dx \\
&= \frac{\log^2(1-p)}{2a\beta p} + \frac{1}{a\beta} \int_0^{\infty} \frac{x(1-p)^{-1} e^{-x}}{1+p(1-p)^{-1} e^{-x}} dx \\
&= \frac{\log^2(1-p)}{2a\beta p} + \frac{1}{2a\beta} E\left\{X^2 \mid X \sim EEG\left(1, \frac{1}{1-p}\right)\right\} \\
&= \frac{\log^2(1-p)}{2a\beta p} + \frac{1}{a\beta(1-p)} \Phi\left(\frac{-p}{1-p}, 2, 1\right),
\end{aligned}$$

όπου για τον υπολογισμό της ζητούμενης ροπής και ακολούθως των μέσων τιμών και των διακυμάνσεων των X και Y , με σ.π.π. αυτές των σχέσεων (4.11) και (4.12), χρησιμοποιούνται τα αποτελέσματα της ενότητας 2.4. Οι αναμενόμενες τιμές των X και Y , υπολογίζονται ως

$$E(X; \theta_1) = \frac{-\log(1-p)}{ap},$$

$$E(Y; \theta_2) = \frac{-\log(1-p)}{\beta p}$$

και επομένως, η συνδιακύμανσή τους, είναι

$$Cov(X, Y; \theta) = \frac{1}{a\beta p} \left\{ \frac{p}{1-p} \Phi\left(\frac{-p}{1-p}, 2, 1\right) + \frac{(p-2)\log^2(1-p)}{2p} \right\}.$$

Για τις διακυμάνσεις των X και Y , προκύπτει ότι

$$Var(X; \theta_1) = \frac{1}{a^2(1-p)} \left\{ 2\Phi\left(\frac{-p}{1-p}, 2, 1\right) - \frac{(1-p)\log^2(1-p)}{p^2} \right\},$$

$$Var(Y; \theta_2) = \frac{1}{\beta^2(1-p)} \left\{ 2\Phi\left(-\frac{p}{1-p}, 2, 1\right) - \frac{(1-p)\log^2(1-p)}{p^2} \right\}.$$

Αντικαθιστώντας τις εκφράσεις της συνδιακύμανσης και των διακυμάνσεων των X και Y στη σχέση (1.7), προκύπτει ο συντελεστής συσχέτισης του Pearson για την κατανομή της σχέσης (4.10),



$$\rho(X, Y; \theta) = \frac{2p^2 \Phi\left(\frac{-P}{1-p}, 2, 1\right) + (p-2)(1-p) \log^2(1-p)}{2 \left\{ 2p^2 \Phi\left(-\frac{P}{1-p}, 2, 1\right) - (1-p) \log^2(1-p) \right\}}.$$

4.5 Φυσική ερμηνεία

Η κατανομή *EEG*-1, περιγράφει τη συμπεριφορά ενός διδιάστατου πληθυσμού, όταν οι περιθώριοι πληθυσμοί περιγράφονται από την τροποποιημένη Extreme Value κατανομή και είναι ανομοιογενείς ως προς την ίδια συνεχή παράμετρο, η οποία συμπεριφέρεται ως εκθετική τυχαία μεταβλητή.

Οι κατανομές *EEG*-2,3,4, περιγράφουν την από κοινού συμπεριφορά των χρόνων ζωής ενός 'συστήματος' που αποτελείται από δύο βασικές συνιστώσες, όταν ο χρόνος ζωής κάθε συνιστώσας συμπεριφέρεται όπως μια από τις τυχαίες μεταβλητές Z ή W του Θεωρήματος 2.2. Επομένως, ο τρόπος λειτουργίας κάθε βασικής συνιστώσας, είναι αυτός που σκιαγραφείται στα σχόλια του ίδιου θεωρήματος ενώ σε κάθε περίπτωση, ο αριθμός των αρχικών ελαττωμάτων είναι ο ίδιος για καθεμία από αυτές. Η κατανομή *EEG*-2, αναφέρεται σε συστήματα στα οποία οι επιμέρους συνιστώσες των δύο βασικών συνιστωσών έχουν συνδεσμολογία σειράς ενώ η *EEG*-4, σε συστήματα στα οποία οι δευτερεύουσες συνιστώσες έχουν παράλληλη σύνδεση. Όταν μεταξύ των δύο βασικών συνιστωσών, οι επιμέρους συνιστώσες της πρώτης έχουν συνδεσμολογία σειράς και της δεύτερης παράλληλη σύνδεση, προκύπτει η κατανομή *EEG*-3.



Κεφάλαιο 5

Ιδιότητες Κατανομών Ανάλυσης Επιβίωσης - Μέθοδοι Παραγωγής και Γενίκευσής τους

Η υπόθεση της ομοιογένειας ενός πληθυσμού ως προς την ικανότητα επιβίωσης των μελών του, δεν είναι πάντα ρεαλιστική. Αν και ένα μέρος της ετερογένειας στην επιβίωση εξηγείται με την εισαγωγή συμμεταβλητών, συχνά ένα σημαντικό υπόλοιπο ανομοιογένειας παραμένει ανεξήγητο, ειδικότερα στο προσδόκιμο ζωής βιολογικών πληθυσμών. Η επίδρασή του στην επιβίωση αποτελεί αντικείμενο της θεωρίας της ευπάθειας, με αρκετά εκτεταμένη βιβλιογραφία (π.χ. Hougaard, 1984, Manton et al., 1986, Vaupel, 1990, Aalen, 1994 και Hougaard, 1995), στα πλαίσια της οποίας γίνεται αξιωματικά η παραδοχή ότι η σχετική ανομοιογένεια του πληθυσμού δεν παρατηρείται αλλά εκδηλώνεται έμμεσα. Η βασική της ιδέα είναι ότι οι πιο ευπαθείς πληθυσμιακές μονάδες έχουν μικρότερη διάρκεια ζωής και ως εκ τούτου, η παρατηρούμενη συνάρτηση επιβίωσης δεν είναι αντιπροσωπευτική για καθεμία από αυτές.

Θεωρώντας ότι στο μοντέλο των Vaupel et al. της ενότητας 1.2.2, η μεταβλητή της ευπάθειας ακολουθεί κατανομή Γάμα, προκύπτει το Γάμα μοντέλο ευπάθειας. Στο κεφάλαιο αυτό, εισάγεται η πρόσθετη υπόθεση ότι η αναμενόμενη τιμή της μεταβλητής της ευπάθειας είναι ίση με τη μονάδα δηλαδή ότι η τελευταία ακολουθεί κατανομή Γάμα με ίσες παραμέτρους μορφής (shape) και κλίμακας (scale). Έτσι, η δεσμευμένη συνάρτηση κινδύνου, h_1 , - η οποία θα αναφέρεται στη συνέχεια ως η αρχική συνάρτηση κινδύνου - γίνεται κατά μία έννοια μια μέση συνάρτηση κινδύνου για τις πληθυσμιακές μονάδες. Στο κεφάλαιο αυτό, αποδεικνύεται ότι όλες οι κατανομές ανάλυσης επιβίωσης μπορούν να προκύψουν από το



Γάμα μοντέλο ευπάθειας με κατάλληλη επιλογή της αρχικής συνάρτησης κινδύνου και παρουσιάζονται παραδείγματα χρήσης του, για την παραγωγή ή γενίκευση υπαρχόντων κατανομών με την προσθήκη μιας επιπλέον παραμέτρου. Επίσης, μελετώνται διάφορες ιδιότητές του, δίνονται ικανές και αναγκαίες συνθήκες χαρακτηρισμού του συναρτήσεως της μονοτονίας της συνάρτησης κινδύνου και τέλος, παράγονται φράγματα για τη συνάρτηση επιβίωσης. Έτσι, μελετώνται με ενιαίο τρόπο οι ιδιότητες των κατανομών ανάλυσης επιβίωσης και εξάγονται συμπεράσματα για τις συναρτήσεις κινδύνου και επιβίωσης του πληθυσμού, σε σχέση με τη συμπεριφορά της 'μέσης συνάρτησης κινδύνου' των πληθυσμιακών μονάδων και του βαθμού ανομοιογένειας στην ικανότητα επιβίωσής τους.

5.1 Το Γάμα μοντέλο ευπάθειας

Υποθέτοντας ότι η μεταβλητή της ευπάθειας ακολουθεί κατανομή Γάμα με παραμέτρους μορφής και κλίμακας ίσες με $c(\theta)^{-1}$ και χρησιμοποιώντας τη σχέση (1.3), η συνάρτηση κινδύνου του πληθυσμού δίνεται από τη σχέση:

$$h(t; \mathcal{G}) = \frac{h_1(t; \theta_1)}{1 + c(\theta)H_1(t; \theta_1)}, \quad t \geq 0, \quad (5.1)$$

όπου \mathcal{G} είναι το διάνυσμα όλων των παραμέτρων, θ_1 είναι το διάνυσμα των παραμέτρων της h_1 και $c(\theta) > 0$ είναι μία σταθερά ή μία συνάρτηση παραμέτρων και μόνο, αυτών της κατανομής που αντιστοιχεί στην h_1 ή /και μιας ακόμα που θέλουμε να εισάγουμε. Για $c(\theta) = 0$, οι h και h_1 ταυτίζονται και ως εκ τούτου, η περίπτωση αυτή δεν εξετάζεται στη συνέχεια. Από άποψη φυσικής ερμηνείας, η παραμετρική συνάρτηση $c(\theta)$ είναι ένα μέτρο της ανομοιογένειας του πληθυσμού, αφού ισούται με τη διακύμανση της μεταβλητής της ευπάθειας ή διαφορετικά, με το τετράγωνο του συντελεστή μεταβλητότητας του πληθυσμού της.

Μπορεί να παρατηρήσει κανείς ότι $h(0; \mathcal{G}) = h_1(0; \theta_1)$ ενώ όταν η h_1 έχει πεπερασμένη τιμή για $t \rightarrow \infty$, η h τείνει ασυμπτωτικά στο μηδέν. Επίσης, για κάθε $t > 0$, είναι $h(t; \mathcal{G}) < h_1(t; \theta_1)$.

Οι συναρτήσεις πυκνότητας πιθανότητας και επιβίωσης που αντιστοιχούν στη συνάρτηση κινδύνου της σχέσης (5.1), είναι



$$f(t; \mathcal{G}) = \frac{h_1(t; \theta_1)}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}+1}}, \quad t \geq 0, \quad (5.2)$$

$$S(t; \mathcal{G}) = \frac{1}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}}}, \quad t \geq 0. \quad (5.3)$$

Επίσης, αν ξ_p και ξ_{1p} είναι τα p εκατοστιαία σημεία των κατανομών που αντιστοιχούν στις h και h_1 , τότε $\xi_p = \xi_{1p'}$, όπου

$$p' = 1 - \exp\left[-\frac{1}{c(\theta)} \left\{ (1-p)^{-c(\theta)} - 1 \right\}\right].$$

Η οικογένεια κατανομών της σχέσης (5.2), μπορεί να χρησιμοποιηθεί για γενίκευση κατανομών με προσθήκη παραμέτρου, με σκοπό την αύξηση της ευελιξίας τους στην περιγραφή χρόνων ζωής. Για παράδειγμα, θέτοντας $h_1(t; \theta_1) = \lambda t^{a-1}$ και $c(\theta) = (1-k)^2$ με $t, \lambda, a, k > 0$ στην (5.1), προκύπτει η κατανομή με συνάρτηση κινδύνου

$$h(t; \mathcal{G}) = \frac{\lambda a t^{a-1}}{1 + (1-k)^2 \lambda t^a}, \quad t \geq 0$$

και αντίστοιχη σ.π.π.

$$f(t; \mathcal{G}) = \frac{\lambda a t^{a-1}}{\{1 + (1-k)^2 \lambda t^a\}^{(1-k)^{-2}+1}}, \quad t \geq 0,$$

όπου $\mathcal{G} = (a, \lambda, k)$. Η τελευταία οικογένεια κατανομών περιλαμβάνει ως ειδικές περιπτώσεις την κατανομή Weibull για $k = 1$ και τη γενικευμένη compound Rayleigh κατανομή (Bekker et al., 2000) για $k \neq 1$.

Στο θεώρημα που ακολουθεί, αποδεικνύεται ότι η συνάρτηση κινδύνου οποιασδήποτε κατανομής ανάλυσης επιβίωσης μπορεί να γραφεί στη μορφή (5.1) με κατάλληλη επιλογή της αρχικής συνάρτησης κινδύνου ενώ στο αντίστοιχο πόρισμα εξετάζεται μια ιδιότητα κλειστότητας της οικογένειας κατανομών της σχέσης (5.2).

Θεώρημα 5.1

Η παραμετρική οικογένεια κατανομών της σχέσης (5.2), περιλαμβάνει όλες τις κατανομές ανάλυσης επιβίωσης.



Απόδειξη

Αρκεί να δειχθεί ότι υπάρχει πάντα μια συνάρτηση κινδύνου h_1 , μέσω της οποίας η εκάστοτε συνάρτηση κινδύνου h μπορεί να γραφεί στη μορφή (5.1). Η σχέση (5.1) γράφεται διαδοχικά στις ακόλουθες ισοδύναμες μορφές:

$$h(t; \mathcal{G}) = \frac{h_1(t; \theta_1)}{1 + c(\theta)H_1(t; \theta_1)}$$

$$\Leftrightarrow h_1(t; \theta_1) = h(t; \mathcal{G}) + c(\theta)h(t; \mathcal{G})H_1(t; \theta_1)$$

$$\Leftrightarrow H_1'(t; \theta_1) - c(\theta)h(t; \mathcal{G})H_1(t; \theta_1) = h(t; \mathcal{G}).$$

Η τελευταία είναι μια μη ομογενής γραμμική διαφορική εξίσωση πρώτης τάξης ως προς H_1 , η λύση της οποίας δίνει

$$H_1(t; \theta_1) = e^{c(\theta)H(t; \mathcal{G})} \left\{ H_1(0; \theta_1) + \int_0^t h(t; \mathcal{G}) e^{-c(\theta)H(t; \mathcal{G})} dt \right\}$$

$$= e^{c(\theta)H(t; \mathcal{G})} \left\{ H_1(0; \theta_1) - \frac{1}{c(\theta)} e^{-c(\theta)H(t; \mathcal{G})} \Big|_0^t \right\}$$

$$= e^{c(\theta)H(t; \mathcal{G})} \left[H_1(0; \theta_1) - \frac{1}{c(\theta)} \left\{ e^{-c(\theta)H(t; \mathcal{G})} - 1 \right\} \right].$$

Όμως, $H_1(0; \theta_1) = 0$. Άρα,

$$H_1(t; \theta_1) = \frac{1}{c(\theta)} \left\{ e^{c(\theta)H(t; \mathcal{G})} - 1 \right\}$$

ή ισοδύναμα,

$$h_1(t; \theta_1) = h(t; \mathcal{G}) e^{c(\theta)H(t; \mathcal{G})}. \quad (5.4)$$

Από τις δύο τελευταίες σχέσεις, είναι προφανές ότι η h_1 ικανοποιεί τις ιδιότητες μιας συνάρτησης κινδύνου, αποδεικνύοντας τον ισχυρισμό του θεωρήματος.

Στον Πίνακα 5.1, παρουσιάζονται παραδείγματα στην παραγωγή των πιο ευρέως χρησιμοποιούμενων κατανομών ανάλυσης επιβίωσης, χρησιμοποιώντας την (5.1). Για κάθε κατανομή δίνονται οι συναρτήσεις πυκνότητας πιθανότητας f και κινδύνου h , όπως επίσης και η αρχική συνάρτηση κινδύνου h_1 μέσω της οποίας η σχέση (5.1) παράγει την h . Για παράδειγμα, αντικαθιστώντας στη σχέση (5.1) όπου h_1 την $h_1(t; \theta_1) = \lambda t^{a-1} \exp\{c(\theta)\lambda t^a\}$ με



$f(t; \mathcal{G})$	$h(t; \mathcal{G})$	$h_1(t; \theta_1)$	
Exponential $\lambda e^{-\lambda t}$	λ	$\lambda e^{c(\theta)\lambda t}$	$\lambda > 0, t \geq 0$
Gamma $\frac{1}{\Gamma(a)} \lambda (\lambda t)^{a-1} \exp(-\lambda t)$	$\frac{f(t; \mathcal{G})}{1 - {}^*I(a, \lambda t)}$	$\frac{f(t; \mathcal{G})}{\{1 - {}^*I(a, \lambda t)\}^{c(\theta)+1}}$	$\lambda, a > 0,$ $t > 0$
Weibull $\lambda a t^{a-1} \exp(-\lambda t^a)$	$\lambda a t^{a-1}$	$\lambda a t^{a-1} \exp\{c(\theta)\lambda t^a\}$	$\lambda, a > 0,$ $t > 0$
Modified Extreme Value $a \lambda \exp\{\lambda t + a(1 - e^{\lambda t})\}$	$a \lambda \exp(\lambda t)$	$a \lambda \exp\{\lambda t + a c(\theta)(e^{\lambda t} - 1)\}$	$\lambda, a > 0,$ $t \geq 0$
Burr $\frac{\lambda a t^{a-1}}{(1+t^a)^{\lambda+1}}$	$\frac{\lambda a t^{a-1}}{1+t^a}$	$\lambda a t^{a-1} (1+t^a)^{c(\theta)\lambda-1}$	$\lambda, a > 0,$ $t \geq 0$
Lognormal $\frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log t - \mu)^2}{2\sigma^2}\right\}$	$\frac{f(t; \mathcal{G})}{1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right)}$	$\frac{f(t; \mathcal{G})}{\left\{1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right)\right\}^{c(\theta)+1}}$	$-\infty < \mu < \infty,$ $\sigma > 0, t > 0$ $\Phi(t) = \text{α.σ.κ.της}$ τυπικής κανονικής κατανομής
Inverse Gaussian $\sqrt{\frac{\lambda}{2\pi t^3}} \exp\left\{-\frac{\lambda(t - \mu)^2}{2\mu^2 t}\right\}$	$\frac{f(t; \mathcal{G})}{{}^{**}S(t; \mathcal{G})}$	$\frac{f(t; \mathcal{G})}{\{{}^{**}S(t; \mathcal{G})\}^{c(\theta)+1}}$	$\mu, \lambda > 0,$ $t > 0$
Lomax (Pareto 2 nd kind) $\frac{a\lambda}{(1 + \lambda t)^{a+1}}$	$\frac{a\lambda}{1 + \lambda t}$	$a\lambda(1 + \lambda t)^{ac(\theta)-1}$	$\lambda, a > 0,$ $t \geq 0$

$${}^*I(a, t) = \{\Gamma(a)\}^{-1} \int_0^t u^{a-1} e^{-u} du \quad {}^{**}S(t; \mathcal{G}) = \Phi\left\{\sqrt{\lambda/t}(1-t/\mu)\right\} - e^{2\lambda/\mu} \Phi\left\{-\sqrt{\lambda/t}(1+t/\mu)\right\}$$

Πίνακας 5.1 Κατανομές ανάλυσης επιβίωσης με τις αντίστοιχες συναρτήσεις πυκνότητας πιθανότητας f και κινδύνου h και τις αρχικές συναρτήσεις κινδύνου h_1 που χρησιμοποιούνται για την παραγωγή τους.



$t, \lambda, a > 0$, δηλαδή τη συνάρτηση κινδύνου της κατανομής Exponential Power (Smith και Bain, 1975) για $c(\theta)=1$ ή $c(\theta)^{-1}$ πολλαπλάσιο αυτής για $c(\theta) \neq 1$, προκύπτει η συνάρτηση κινδύνου της κατανομής Weibull $h(t; \theta) = \lambda a t^{a-1}$. Η μορφή της παραμετρικής συνάρτησης $c(\theta)$ δεν προσδιορίζεται ώστε ο πίνακας να περιλαμβάνει όλες τις εναλλακτικές εκδοχές της h_1 .

Πόρισμα 5.1

Αν οι $\{X_i\}_{i=1}^n$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με κατανομή στην οικογένεια της σχέσης (5.2), τότε οι κατανομές των $\min\{X_i\}_{i=1}^n$ και $\max\{X_i\}_{i=1}^n$ ανήκουν επίσης στην ίδια οικογένεια.

Απόδειξη

Η εν λόγω ιδιότητα, είναι άμεση συνέπεια του προηγούμενου θεωρήματος αφού οι κατανομές των $\min\{X_i\}_{i=1}^n$ και $\max\{X_i\}_{i=1}^n$ είναι κατανομές ανάλυσης επιβίωσης. Ειδικότερα, η κατανομή του $\min\{X_i\}_{i=1}^n$ ανήκει στην οικογένεια της σχέσης (5.2), με $h_1^* = nh_1$ και $c(\theta)^* = c(\theta)/n$.

5.2 Χαρακτηρισμοί

Στη συνέχεια, δίνονται ικανές και αναγκαίες συνθήκες ώστε η οικογένεια κατανομών στην (5.2) να έχει αύξουσα (IFR), φθίνουσα (DFR), κυπελλοειδούς μορφής (BTFR) ή μονοκόρυφη (UBTFR) συνάρτηση κινδύνου. Επίσης, παράγονται αποτελέσματα που συνδέουν τις μονοτονίες των h και h_1 .

Θεώρημα 5.2

Οι παρακάτω προτάσεις είναι ισοδύναμες.

- (a) Η οικογένεια κατανομών της σχέσης (5.2) είναι IFR (DFR).
- (b) Η συνάρτηση $\log\{1 + c(\theta)H_1(t; \theta_1)\}$ είναι κυρτή (κοίλη) για κάθε $t \geq 0$.
- (c) Η συνάρτηση $\{1 + c(\theta)H_1(t; \theta_1)\}^{-1}$ είναι μια Pólya συνάρτηση κατανομής τάξης 2 (PF_2) (Karlin, 1968) (η συνάρτηση $\{1 + c(\theta)H_1(t+x; \theta_1)\}^{-1}$ είναι ολικά θετική τάξης 2 (totally



positive of order 2 ή TP_2) (Karlin, 1968) στα t και x , για $t+x \geq 0$).

Απόδειξη

(a) \Leftrightarrow (b) Η παράγωγος της συνάρτησης $\log\{1+c(\theta)H_1(t;\theta_1)\}$ ως προς t , είναι

$$[\log\{1+c(\theta)H_1(t;\theta_1)\}]' = \frac{c(\theta)h_1(t;\theta_1)}{1+c(\theta)H_1(t;\theta_1)} = c(\theta)h(t;\mathcal{G}).$$

Άρα, η οικογένεια κατανομών της σχέσης (5.2), είναι IFR (DFR) αν και μόνο αν η συνάρτηση $\log\{1+c(\theta)H_1(t;\theta_1)\}$ είναι κυρτή (κοίλη) για κάθε $t \geq 0$.

(a) \Leftrightarrow (c) Η ισοδυναμία αποδεικνύεται από το γεγονός ότι μια κατανομή ανάλυσης επιβίωσης είναι IFR (DFR) αν και μόνο αν η συνάρτηση επιβίωσής της είναι PF_2 (η $S(t+x)$ είναι TP_2 στα t και x , για $t+x \geq 0$) (Barlow και Proschan, 1965, σελ. 25). Όμως, αν η S είναι PF_2 , για $t_1 \leq t_2$ και $x \geq 0$, προκύπτουν οι ακόλουθες ισοδυναμίες:

$$\begin{aligned} & \left| \begin{array}{cc} S(t_1+x;\mathcal{G}) & S(t_2+x;\mathcal{G}) \\ S(t_1;\mathcal{G}) & S(t_2;\mathcal{G}) \end{array} \right| \geq 0 \\ \Leftrightarrow & \left| \begin{array}{cc} \frac{1}{\{1+c(\theta)H_1(t_1+x;\theta_1)\}^{c(\theta)^{-1}}} & \frac{1}{\{1+c(\theta)H_1(t_2+x;\theta_1)\}^{c(\theta)^{-1}}} \\ \frac{1}{\{1+c(\theta)H_1(t_1;\theta_1)\}^{c(\theta)^{-1}}} & \frac{1}{\{1+c(\theta)H_1(t_2;\theta_1)\}^{c(\theta)^{-1}}} \end{array} \right| \geq 0 \\ \Leftrightarrow & \frac{1}{[\{1+c(\theta)H_1(t_1+x;\theta_1)\}\{1+c(\theta)H_1(t_2;\theta_1)\}]^{c(\theta)^{-1}}} \\ & \geq \frac{1}{[\{1+c(\theta)H_1(t_2+x;\theta_1)\}\{1+c(\theta)H_1(t_1;\theta_1)\}]^{c(\theta)^{-1}}} \\ \Leftrightarrow & \left| \begin{array}{cc} \frac{1}{1+c(\theta)H_1(t_1+x;\theta_1)} & \frac{1}{1+c(\theta)H_1(t_2+x;\theta_1)} \\ \frac{1}{1+c(\theta)H_1(t_1;\theta_1)} & \frac{1}{1+c(\theta)H_1(t_2;\theta_1)} \end{array} \right| \geq 0. \end{aligned}$$

Άρα, ισοδύναμα η συνάρτηση $\{1+c(\theta)H_1(t;\theta_1)\}^{-1}$ είναι PF_2 . Όμοια, αν η $S(t+x)$ είναι TP_2 στα t και x , για $t_1 \leq t_2$ και $x_1 \leq x_2$ ισχύει ότι:



$$\left| \begin{array}{cc} S(t_1 + x_1; \mathcal{G}) & S(t_1 + x_2; \mathcal{G}) \\ S(t_2 + x_1; \mathcal{G}) & S(t_2 + x_2; \mathcal{G}) \end{array} \right| \geq 0$$

ή ισοδύναμα,

$$\Leftrightarrow \left| \begin{array}{cc} \frac{1}{1 + c(\theta)H_1(t_1 + x_1; \theta_1)} & \frac{1}{1 + c(\theta)H_1(t_1 + x_2; \theta_1)} \\ \frac{1}{1 + c(\theta)H_1(t_2 + x_1; \theta_1)} & \frac{1}{1 + c(\theta)H_1(t_2 + x_2; \theta_1)} \end{array} \right| \geq 0$$

δηλαδή η $\{1 + c(\theta)H_1(t + x; \theta_1)\}^{-1}$ είναι TP_2 .

Θεώρημα 5.3

Έστω η συνάρτηση

$$g(t; \mathcal{G}) = h_1'(t; \theta_1)\{1 + c(\theta)H_1(t; \theta_1)\} - c(\theta)h_1^2(t; \theta_1).$$

Η οικογένεια κατανομών της σχέσης (5.2), είναι

i) IFR (DFR) αν και μόνο αν $g(t; \mathcal{G}) > 0$ για κάθε $t \geq 0$ ($g(t; \mathcal{G}) < 0$ για κάθε $t \geq 0$).

ii) BTFR (UBTFR) αν και μόνο αν υπάρχει $t_0 > 0$ τέτοιο ώστε $g(t; \mathcal{G}) < 0$ για κάθε $t \in (0, t_0)$ και $g(t; \mathcal{G}) > 0$ για κάθε $t \in (t_0, \infty)$, με $g(t_0; \mathcal{G}) = 0$ (υπάρχει $t_0 > 0$ τέτοιο ώστε $g(t; \mathcal{G}) > 0$ για κάθε $t \in (0, t_0)$ και $g(t; \mathcal{G}) < 0$ για κάθε $t \in (t_0, \infty)$, με $g(t_0; \mathcal{G}) = 0$).

Απόδειξη

Η παράγωγος της h ως προς t , δίνεται από τη σχέση:

$$h'(t; \mathcal{G}) = \frac{h_1'(t; \theta_1)\{1 + c(\theta)H_1(t; \theta_1)\} - c(\theta)h_1^2(t; \theta_1)}{\{1 + c(\theta)H_1(t; \theta_1)\}^2}.$$

Η συνάρτηση g είναι ο αριθμητής της παραπάνω σχέσης, οπότε τα αποτελέσματα προκύπτουν εφαρμόζοντας τους ορισμούς των IFR, DFR, BTFR και UBTFR ιδιοτήτων, που δίνονται στην ενότητα 1.1.1.

Στη συνέχεια, εξετάζονται περιπτώσεις όπου το είδος μονοτονίας της h_1 καθορίζει εκείνο της h ή αντίστροφα.

Πρόταση 5.1

Αν η h_1 είναι φθίνουσα ή σταθερή, η οικογένεια κατανομών στην (5.2) είναι DFR.



Απόδειξη

Αν η h_1 είναι φθίνουσα συνάρτηση του t , τότε:

$$g(t; \mathcal{G}) = h_1'(t; \theta_1) \{1 + c(\theta) H_1(t; \theta_1)\} - c(\theta) h_1^2(t; \theta_1) < 0$$

ενώ αν είναι σταθερή

$$g(t; \mathcal{G}) = -c(\theta) h_1^2(t; \theta_1) < 0,$$

οπότε το αποτέλεσμα προκύπτει με εφαρμογή του προηγούμενου θεωρήματος.

Το αντίστροφο της Πρότασης 5.1, γενικά δεν ισχύει. Αντιπαράδειγμα, αποτελεί η περίπτωση της κατανομής *EEG* (Adamidis et al., 2005) με συνάρτηση κινδύνου

$$h(t; \mathcal{G}) = \lambda \{1 - (1 - a)e^{-\lambda t}\}^{-1}, \quad t \geq 0,$$

όπου $\lambda, a > 0$, η οποία μπορεί να παραχθεί από την (5.1), με

$$h_1(t; \theta_1) = \lambda a^{-1} e^{\lambda t}, \quad t \geq 0$$

και $c(\theta) = 1$. Αν και για $0 < a \leq 1$ η πρώτη συνάρτηση κινδύνου είναι φθίνουσα, η δεύτερη είναι αύξουσα συνάρτηση του t .

Πρόταση 5.2

Αν η οικογένεια κατανομών της σχέσης (5.2) είναι IFR, η h_1 είναι αύξουσα συνάρτηση του t .

Απόδειξη

Παραγωγίζοντας τη σχέση (5.4) ως προς t και αφού η h είναι αύξουσα, προκύπτει ότι

$$h_1'(t; \theta_1) = e^{c(\theta)H(t; \mathcal{G})} \{h'(t; \mathcal{G}) + c(\theta)h^2(t; \mathcal{G})\} > 0,$$

που αποδεικνύει τον ισχυρισμό για τη μονοτονία της h_1 .

Βεβαίως, αν η h_1 είναι αύξουσα συνάρτηση του t , δεν έπεται ότι και η h έχει το ίδιο είδος μονοτονίας όπως φαίνεται από το παράδειγμα της κατανομής *EEG*.

5.3 Φράγματα της συνάρτησης επιβίωσης

Στην παράγραφο αυτή, εξετάζεται η σχέση διάταξης που διέπει τις συναρτήσεις επιβίωσης και μέσου υπολειπόμενου χρόνου ζωής των κατανομών που αντιστοιχούν στις h και h_1 .



Επίσης, παράγονται φράγματα για τη συνάρτηση επιβίωσης της σχέσης (5.3), μέσω των οποίων φαίνεται η επίδραση της 'μέσης συνάρτησης κινδύνου' των πληθυσμιακών μονάδων και του βαθμού ανομοιογένειας του πληθυσμού - μέσω του $c(\theta)$ - στο εύρος τιμών της παρατηρούμενης συνάρτησης επιβίωσης.

Θεώρημα 5.4

Αν S και S_1 είναι οι συναρτήσεις επιβίωσης που αντιστοιχούν στις h και h_1 , τότε $S(t; \mathcal{G}) \geq S_1(t; \theta_1)$ για κάθε $t \geq 0$.

Απόδειξη

Ολοκληρώνοντας τα δύο μέλη της ανισότητας $h(t; \mathcal{G}) \leq h_1(t; \theta_1)$ στο διάστημα $[0, t]$, έπεται ότι

$$H(t; \mathcal{G}) \leq H_1(t; \theta_1)$$

και ισοδύναμα,

$$e^{-H(t; \mathcal{G})} \geq e^{-H_1(t; \theta_1)}.$$

Επομένως, όπως ενδεχομένως αναμενόταν, $S(t; \mathcal{G}) \geq S_1(t; \theta_1)$ για κάθε $t \geq 0$.

Η ίδια διάταξη διατηρείται και σε ότι αφορά τους μέσους υπολειπόμενους χρόνους ζωής όπως φαίνεται στο θεώρημα που ακολουθεί.

Θεώρημα 5.5

Αν m και m_1 είναι οι συναρτήσεις του μέσου υπολειπόμενου χρόνου ζωής που αντιστοιχούν στις h και h_1 , τότε $m(t_0; \mathcal{G}) \geq m_1(t_0; \theta_1)$ για κάθε $t_0 \geq 0$.

Απόδειξη

Η συνάρτηση $w(t; \mathcal{G}) = S(t; \mathcal{G})\{S_1(t; \theta_1)\}^{-1}$ είναι αύξουσα συνάρτηση του t , αφού

$$\begin{aligned} w'(t; \mathcal{G}) &= \frac{S'(t; \mathcal{G})S_1(t; \theta_1) - S(t; \mathcal{G})S_1'(t; \theta_1)}{S_1^2(t; \theta_1)} \\ &= \frac{1}{S_1^2(t; \theta_1)} \{-f(t; \mathcal{G})S_1(t; \theta_1) + S(t; \mathcal{G})f_1(t; \theta_1)\} \\ &= \frac{1}{S_1^2(t; \theta_1)} \left[-\frac{h_1(t; \theta_1)S_1(t; \theta_1)}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)-1}} + \frac{f_1(t; \theta_1)}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)-1}} \right] \end{aligned}$$



$$= \frac{1}{S_1^2(t; \theta_1)} \left[-\frac{f_1(t; \theta_1)}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}+1}} + \frac{f_1(t; \theta_1)}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}}} \right]$$

$$= \frac{c(\theta)f_1(t; \theta_1)H_1(t; \theta_1)}{S_1^2(t; \theta_1)\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}+1}} > 0.$$

Άρα, για $t > t_0$,

$$\frac{S(t; \mathcal{G})}{S_1(t; \theta_1)} > \frac{S(t_0; \mathcal{G})}{S_1(t_0; \theta_1)} \Rightarrow \int_{t_0}^{\infty} \frac{S(t; \mathcal{G})}{S(t_0; \mathcal{G})} dt > \int_{t_0}^{\infty} \frac{S_1(t; \theta_1)}{S_1(t_0; \theta_1)} dt,$$

που αποδεικνύει το ζητούμενο.

Θεώρημα 5.6

i) Αν η οικογένεια κατανομών της σχέσης (5.2) είναι IFR, τότε:

$$\left\{ \frac{h_1(0; \theta_1)}{h_1(t; \theta_1)} \right\}^{c(\theta)^{-1}} < S(t; \mathcal{G}) < e^{-h_1(0; \theta_1)t} \text{ για κάθε } t \geq 0.$$

ii) Αν η οικογένεια κατανομών σχέσης (5.2) είναι DFR, τότε:

$$e^{-h_1(0; \theta_1)t} < S(t; \mathcal{G}) < \min \left[1, \left\{ \frac{h_1(0; \theta_1)}{h_1(t; \theta_1)} \right\}^{c(\theta)^{-1}} \right] \text{ για κάθε } t \geq 0.$$

Απόδειξη

i) Αφού από υπόθεση η h είναι αύξουσα συνάρτηση του t , για $t > 0$ ισχύει ότι $h(t; \mathcal{G}) > h(0; \mathcal{G}) = h_1(0; \theta_1)$. Ισοδύναμα,

$$\frac{h_1(t; \theta_1)}{1 + c(\theta)H_1(t; \theta_1)} > h_1(0; \theta_1)$$

$$\Leftrightarrow \frac{1}{\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}}} > \left\{ \frac{h_1(0; \theta_1)}{h_1(t; \theta_1)} \right\}^{c(\theta)^{-1}}$$

$$\Leftrightarrow S(t; \mathcal{G}) > \left\{ \frac{h_1(0; \theta_1)}{h_1(t; \theta_1)} \right\}^{c(\theta)^{-1}}.$$

Το άνω φράγμα της συνάρτησης επιβίωσης προκύπτει ολοκληρώνοντας τα δύο μέλη της



ανισότητας $h(t; \mathcal{G}) > h_1(0; \theta_1)$ στο διάστημα $[0, t]$, οπότε προκύπτει ότι

$$H(t; \mathcal{G}) > h_1(0; \theta_1)t$$

και ισοδύναμα,

$$e^{-H(t; \mathcal{G})} < e^{-h_1(0; \theta_1)t}$$

$$\Leftrightarrow S(t; \mathcal{G}) < e^{-h_1(0; \theta_1)t}.$$

ii) Με ανάλογα επιχειρήματα αποκτώνται τα άνω και κάτω φράγματα της συνάρτησης επιβίωσης όταν η h είναι φθίνουσα συνάρτηση του t , λαμβάνοντας υπόψιν ότι η διάταξη των h και $h_1(0)$ αντιστρέφεται.

Στο επόμενο θεώρημα, δίνονται συμπληρωματικά φράγματα της συνάρτησης επιβίωσης, που προκύπτουν θέτοντας περιορισμούς στις τιμές της παραμετρικής συνάρτησης $c(\theta)$.

Θεώρημα 5.7

i) Αν $c(\theta) \leq 1$, τότε:

$$\frac{1}{\{1 + H_1(t; \theta_1)\}^{c(\theta)^{-1}}} \leq S(t; \mathcal{G}) \leq \frac{1}{1 + H_1(t; \theta_1)}.$$

ii) Αν $c(\theta) \geq 1$, τότε:

$$\frac{1}{1 + H_1(t; \theta_1)} \leq S(t; \mathcal{G}) \leq \frac{1}{\{1 + H_1(t; \theta_1)\}^{c(\theta)^{-1}}}.$$

Απόδειξη

i) Για $c(\theta) \leq 1$, το άνω φράγμα της συνάρτησης επιβίωσης προκύπτει χρησιμοποιώντας την ανισότητα $(1+t)^a \geq 1+at$ όπου $a \geq 1$ ή $a \leq 0$ και $t > -1$. Κατά συνέπεια,

$$\frac{1}{S(t; \mathcal{G})} = \{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}} \geq 1 + H_1(t; \theta_1)$$

$$\Leftrightarrow S(t; \mathcal{G}) \leq \frac{1}{1 + H_1(t; \theta_1)}.$$

Επίσης,

$$\{1 + c(\theta)H_1(t; \theta)\}^{c(\theta)^{-1}} \leq \{1 + H_1(t)\}^{c(\theta)^{-1}}$$

$$\Leftrightarrow S(t; \mathcal{G}) \geq \frac{1}{\{1 + H_1(t; \theta_1)\}^{c(\theta)^{-1}}}.$$



ii) Για $c(\theta) \geq 1$, χρησιμοποιώντας την ανισότητα $(1+t)^a \leq 1+at$ που ισχύει για $0 < a < 1$ και $t \geq 0$ ή $-1 < t < 0$, έπεται ότι

$$\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}} \leq 1 + H_1(t; \theta_1)$$

$$\Leftrightarrow S(t; \mathcal{D}) \geq \frac{1}{1 + H_1(t; \theta_1)}.$$

Επίσης,

$$\{1 + c(\theta)H_1(t; \theta_1)\}^{c(\theta)^{-1}} \geq \{1 + H_1(t; \theta_1)\}^{c(\theta)^{-1}}$$

$$\Leftrightarrow S(t; \mathcal{D}) \leq \frac{1}{\{1 + H_1(t; \theta_1)\}^{c(\theta)^{-1}}}.$$

Στο θεώρημα που ακολουθεί παρουσιάζεται μια γενική μέθοδος κατασκευής φραγμάτων μιας συνάρτησης κινδύνου.

Θεώρημα 5.8

Αν μια οικογένεια κατανομών είναι IFR (DFR), τότε:

$$S(t; \mathcal{D}) \begin{cases} \geq (\leq) & e^{\frac{t}{t_0} \ln t_0} & t \leq t_0 \\ \leq (\geq) & e^{\frac{t}{t_0} \ln t_0} & t \geq t_0, \end{cases}$$

όπου t_0 είναι η μοναδική λύση της εξίσωσης $S(t; \mathcal{D}) = t$.

Απόδειξη

Σύμφωνα με το Θεώρημα 4.3 των Barlow και Proschan (1965, σελ. 27), αν μια οικογένεια κατανομών είναι IFR (DFR) και $F(\xi_p) = p$, τότε:

$$S(t; \mathcal{D}) \begin{cases} \geq (\leq) & e^{-at} & t \leq \xi_p \\ \leq (\geq) & e^{-at} & t \geq \xi_p, \end{cases}$$

όπου $a = -\xi_p^{-1} \log(1-p)$, οπότε για $\xi_p = t_0$ και $p = 1-t_0$, προκύπτει το αποτέλεσμα. Για την οικογένεια κατανομών της σχέσης (5.2), t_0 είναι η μοναδική λύση της εξίσωσης

$$\{1 + c(\theta)H_1(t; \theta_1)\}^{-c(\theta)^{-1}} = t.$$



Αξίζει επίσης να σημειωθεί ότι αν t_0 είναι η μοναδική λύση της εξίσωσης $S(t; \mathcal{G}) = t$,

τότε

$$S(t; \mathcal{G}) \begin{cases} \geq \{1 + c(\theta)H_1(t_0; \theta_1)\}^{-c(\theta)^{-1}} & 0 \leq t \leq t_0 \\ < \{1 + c(\theta)H_1(t_0; \theta_1)\}^{-c(\theta)^{-1}} & t > t_0. \end{cases}$$



Παράρτημα

Δεδομένα που χρησιμοποιήθηκαν στο δεύτερο κεφάλαιο

Πίνακας 1. Παρατηρήσεις στους κύκλους αντοχής δειγμάτων νήματος πολυεστέρα-βισκόζης μήκους 100 εκατοστών, όπως προέκυψαν στα πλαίσια πειραματικής διαδικασίας ελέγχου του ορίου θραύσης τους, με επίπεδο πίεσης 2.3% (Quesenberry και Kent, 1982).

86, 146, 251, 653, 98, 249, 400, 292, 131, 169, 175, 176, 76, 264, 15,
364, 195, 262, 88, 264, 157, 220, 42, 321, 180, 198, 38, 20, 61, 121,
282, 224, 149, 180, 325, 250, 196, 90, 229, 166, 38, 337, 65, 151, 341,
40, 40, 135, 597, 246, 211, 180, 93, 315, 353, 571, 124, 279, 81, 186,
497, 182, 423, 185, 229, 400, 338, 290, 398, 71, 246, 185, 188, 568, 55,
55, 61, 244, 20, 284, 393, 396, 203, 829, 239, 236, 286, 194, 277, 143,
198, 264, 105, 203, 124, 137, 135, 350, 193, 188.

n=100

Πίνακας 2. Χρόνοι ζωής (σε ώρες) των πίσω δεξιών φρένων ερπυστριοφόρων ελκυστήρων τύπου D9G-66A (Barlow και Campo, 1975 και Chang και Rao, 1993).

56, 83, 104, 116, 244, 305, 429, 452, 453, 503, 552, 614, 661, 673, 683,
685, 753, 763, 806, 834, 838, 862, 897, 904, 981, 1007, 1008, 1049, 1069,
1107, 1125, 1141, 1153, 1154, 1193, 1201, 1253, 1313, 1329, 1347, 1454,
1464, 1490, 1491, 1532, 1549, 1568, 1574, 1586, 1599, 1608, 1723, 1769,
1795, 1927, 1957, 2005, 2010, 2016, 2022, 2037, 2065, 2096, 2139, 2150,
2156, 2160, 2190, 2210, 2220, 2248, 2285, 2325, 2337, 2351, 2437, 2454,
2546, 2565, 2584, 2624, 2675, 2701, 2755, 2877, 2879, 2922, 2986, 3092,
3160, 3185, 3191, 3439, 3617, 3685, 3756, 3826, 3995, 4007, 4159, 4300,
4487, 5074, 5579, 5623, 6869, 7739.

n=107



Δεδομένα που χρησιμοποιήθηκαν στο τρίτο κεφάλαιο

Πίνακας 3. Χρόνοι συντήρησης συστήματος τηλεπικοινωνίας αεροπλάνου (Chhikara και Folks, 1977).

0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0,
1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0,
4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0, 24.5.

n=46

Πίνακας 4. Παρατηρήσεις στους χρόνους ζωής μάντων αεροναυπηγικής όταν υποβάλλονται σε επίπεδο πίεσης 90% (Andrews και Herzberg, 1985, σελ. 182).

0.01, 0.01, 0.02, 0.02, 0.02, 0.03, 0.03, 0.04, 0.05, 0.06, 0.07, 0.07, 0.08,
0.09, 0.09, 0.10, 0.10, 0.11, 0.11, 0.12, 0.13, 0.18, 0.19, 0.20, 0.23, 0.24,
0.24, 0.29, 0.34, 0.35, 0.36, 0.38, 0.40, 0.42, 0.43, 0.52, 0.54, 0.56, 0.60,
0.60, 0.63, 0.65, 0.67, 0.68, 0.72, 0.72, 0.72, 0.73, 0.79, 0.79, 0.80, 0.80,
0.83, 0.85, 0.90, 0.92, 0.95, 0.99, 1.00, 1.01, 1.02, 1.03, 1.05, 1.10, 1.10,
1.11, 1.15, 1.18, 1.20, 1.29, 1.31, 1.33, 1.34, 1.40, 1.43, 1.45, 1.50, 1.51,
1.52, 1.53, 1.54, 1.54, 1.55, 1.58, 1.60, 1.63, 1.64, 1.80, 1.80, 1.81, 2.02,
2.05, 2.14, 2.17, 2.33, 3.03, 3.03, 3.34, 4.20, 4.69, 7.89.

n=101



Abstract of the Dissertation

Title: Survival Analysis Models

Author: Theodora Dimitrakopoulou

Supervisor: Sotirios Loukas

The study of events involving an element of time has a long and important history in statistical research and practice. Survival analysis is a collection of statistical procedures for the analysis of data, where the response of interest is the time until an event occurs. Though such events may refer to any designated experience of interest, they are generally referred to as 'failures', whereas the time to their occurrences is referred to as 'lifetime' or 'failure time'. Examples of failure times include the lifetimes of machine components in industrial reliability, the durations of strikes or periods of unemployment in economics, the times taken by subjects to complete specified tasks in psychological experimentation and the survival or remission times of patients in clinical trials.

Generally speaking, the estimation, prediction or optimization of survival probabilities or life expectancies has become an issue of considerable interest in many different fields of human life and activity. Therefore, survival analysis has developed into an important tool for researchers in many areas, particularly, those involving biomedical studies and industrial life testing.

This dissertation is occupied with continuous lifetime models. In this context, the first chapter, provides a short overview on the basic concepts of survival analysis. Distribution representations of the time to failure are given when the life lengths are measured by a continuous nonnegative random variable and special emphasis is placed on the hazard function due to its intuitive appeal. In the sequel, several univariate popular lifetime



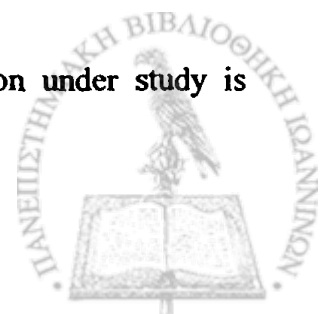
distributions are presented and two specialized models designed to describe more complicated failure patterns (competing risks and frailty models) are briefly examined. The basic concepts of survival analysis for bivariate populations are considered next and the most popular bivariate lifetime distributions are reported.

In the second chapter, various statistical properties and reliability aspects of a two parameter distribution with decreasing and increasing failure rates are explored. The model includes the Exponential-Geometric distribution (Adamidis and Loukas, 1988) as a special case. Characterizations are given and the estimation of parameters is studied by the method of maximum likelihood. An EM algorithm (Dempster et al., 1977) is proposed for computing the estimates and expressions for their asymptotic variances and covariances are derived. Numerical examples based on real data are shown, to illustrate the applicability of the new model. The results of this chapter are included in Adamidis et al. (2005).

Though the most popular lifetime models are those with monotone hazard rates, when the entire life span of a biological entity or a manufactured item is under consideration, high initial and eventual failure rates are frequently observed, indicating a bathtub shaped failure rate (Gaver and Acar, 1979). Also, situations involving a high occurrence of early 'failures' are best modeled by distributions with upturned bathtub shaped hazard rates (Chhikara and Folks, 1977). In the third chapter, a three parameter lifetime distribution with increasing, decreasing, bathtub and upside down bathtub shaped failure rates is introduced. The new model includes the Weibull distribution as a special case. A motivation for its derivation is given using a competing risks interpretation when restricting its parametric space. Several of its statistical properties and reliability aspects are explored and the estimation of the parameters is studied using the standard maximum likelihood procedures. Applications of the model to real data are also included. The results of this chapter are included in Dimitrakopoulou et al. (2006 b).

In the fourth chapter, bivariate extensions of the model introduced in the second chapter are presented, along with the physical considerations leading to their derivation. Marginal and conditional distributions are obtained and their corresponding survival and hazard functions are calculated. The dependence in the proposed bivariate distributions is evaluated by means of the Pearson correlation coefficient.

The models presented so far, implicitly assume that the population under study is



homogeneous, an assumption which is often unrealistic in practice. However, heterogeneity is not only of interest in its own right but actually distorts what is observed. One of the ways of assessing the impact of heterogeneity in mortality studies is via the concept of frailty introduced by Vaupel et al. (1979). When the multiplicative frailty model is under consideration (e.g. Hougaard, 1984), the assumption of a gamma distributed frailty leads to the so called gamma frailty model. Chapter five, is devoted to exploiting some aspects of its relevant distribution theory. Failure rate characterizations are obtained and bounds on the survival function are constructed. Moreover, it is shown that the model can serve as a method of constructing lifetime models or extending existing ones (by adding a parameter in the sense of Marsall and Olkin, (1997)). Therefore, the investigation of its reliability aspects, provides a unified approach in studying lifetime distributions in a reliability context and a way of assessing the impact of the 'average' individual survival capacity - in the presence of heterogeneity - on what is actually observed. The results of this chapter are included in Dimitrakopoulou et al. (2006 a).

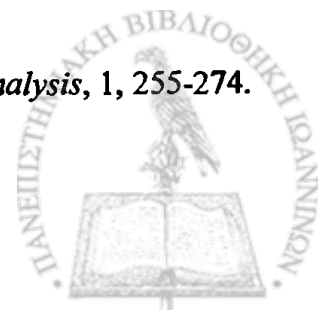


Βιβλιογραφία

- Aalen, O. O. (1988). Heterogeneity in Survival Analysis. *Statistics in Medicine*, 7, 1121-1137.
- Aalen, O. O. (1994). Effects of Frailty in Survival Analysis. *Statistical Methods in Medical Research*, 3, 227-243.
- Adamidis, K., Dimitrakopoulou, T. and Loukas, S. (2005). On an Extension of the Exponential-Geometric Distribution. *Statistics and Probability Letters*, 73, 259-269.
- Adamidis, K. and Loukas, S. (1998). A Lifetime Distribution with Decreasing Failure Rate. *Statistics and Probability Letters*, 39, 35-42.
- Andrews, D. F. and Herzberg, A. M. (1985). *Data. A Collection of Problems from Many Fields for the Student and Research Worker*. Springer-Verlag, New York.
- Barlow, R. E. and Campo, R., (1975). Total Time on Test Processes and Applications to Failure Data Analysis. Reliability and Fault Tree Analysis, Society for Industrial and Applied Mathematics, Philadelphia, 451-481.
- Barlow, R. E. and Proschan, F. (1965). *Mathematical Theory of Reliability*. John Wiley and Sons, Inc., New York.
- Bekker, A., Roux, J. J. J. and Mostert, P. J. (2000). A Generalization of the Compound Rayleigh Distribution: Using a Bayesian Method on Cancer Survival Times. *Communications in Statistics –Theory and Methods*, 29(7), 1419-1433.
- Chang, M. N. and Rao, P. V. (1993). Improved Estimation of Survival Functions in the New-Better-than-Used Class. *Technometrics*, 35, 192-203.
- Chen, Z. (2000). A New Two-Parameter Lifetime Distribution with Bathtub Shape or Increasing Failure Rate Function. *Statistics and Probability Letters*, 49, 155-161.
- Chhikara, R. S. and Folks, J. L. (1977). The Inverse Gaussian Distribution as a Lifetime Model. *Technometrics*, 19(4), 461-468.



- Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm (with discussion). *Journal of the Royal Statistical Society B*, 39, 1-38.
- Dimitrakopoulou, T., Adamidis, K. and Loukas, S. (2006 a). On the Gamma Frailty Model. *Advances and Applications in Statistics*, 6(1), 41-52.
- Dimitrakopoulou, T., Adamidis, K. and Loukas, S. (2006 b). A Lifetime Distribution with an Upside-Down Bathtub-Shaped Hazard Function, 1-4 (Submitted).
- Erdelyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F. G. (1953). *Higher Transcendental Functions*. McGraw-Hill Book Company Inc., New York.
- Finkelstein, M. S. (2003). On One Class of Bivariate Distributions. *Statistics and Probability Letters*, 65, 1-6.
- Freund, J. E. (1961). A Bivariate Extension of the Exponential Distribution. *Journal of the American Statistical Association*, 56(296), 971-977.
- Gaver, D. P. and Acar, M. (1979). Analytical Hazard Representations for Use in Reliability, Mortality and Simulation Studies. *Communications in Statistics – Simulation and Computation*, B8(2), 91-111.
- Ghitany, M. E., Kotz, S. and Xie, M. (2005). On Some Reliability Measures and their Stochastic Orderings for the Topp-Leone Distribution. *Journal of Applied Statistics*, 32(7), 715-722.
- Glaser, R. E. (1980). Bathtub and Related Failure Rate Characterizations. *Journal of the American Statistical Association*, 75, 667-672.
- Gupta, R. D., Gupta, R. C. and Sankaran, P. G. (2004). Some Characterization Results Based on Factorization of the (Reversed) Hazard Rate. *Communications in Statistics - Theory and Methods*, 33(12), 3009-3031.
- Haupt, E. and Schabe, H. (1997). The TTT Transformation and a New Bathtub Distribution Model. *Journal of Statistical Planning and Inference*, 60, 229-240.
- Hjorth, U. (1980). A Reliability Distribution with Increasing, Decreasing, Constant and Bathtub-Shaped Failure Rates. *Technometrics*, 22, 99-107.
- Hougaard, P. (1984). Life Table Methods for Heterogeneous Populations: Distributions Describing the Heterogeneity. *Biometrika*, 71(1), 75-83.
- Hougaard, P. (1995). Frailty Models for Survival Data. *Lifetime Data Analysis*, 1, 255-274.



- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer-Verlag, New York.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Vol. 1, 2nd ed., John Wiley & Sons, Inc., New York.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Vol. 2, 2nd ed., John Wiley & Sons, Inc., New York.
- Jordan, C. W. (1967). *Life Contingencies*. Society of Actuaries, Chicago.
- Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. 2nd ed., John Wiley & Sons, Inc., New Jersey.
- Karlin, S. (1968). *Total Positivity*. Vol. 1, Stanford University Press, Stanford, California.
- Lai, C. D., Xie, M. and Murthy, D. N. P. (2001). Bathtub Shaped Failure Rate Distributions. *Handbook in Reliability*, N. Balakrishnan and C. R. Rao, eds., 20, 69-104.
- Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc., New York.
- Lee, E. T. (1992). *Statistical Methods for Survival Data Analysis*. 2nd ed., John Wiley & Sons, Inc., New York.
- Leemis, L. M. (1995). *Reliability. Probabilistic Models and Statistical Methods*. Prentice-Hall, Inc., New Jersey.
- Magnus, W., Oberhettinger, F. and Soni, R. P. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics*. Springer-Verlang, New York.
- Manton, K. G., Stallard, E. and Vaupel, J. W. (1986). Alternative Models for Heterogeneity of Mortality Risks Among the Aged. *Journal of the American Statistical Association*, 41, 635-644.
- Marshall, A. W. and Olkin, I. (1967). A Multivariate Exponential Distribution. *Journal of the American Statistical Association*, 62, 30-44.
- Marshall, A. W. and Olkin, I. (1997). A New Method for Adding a Parameter to a Family of Distributions with Application to the Exponential and Weibull Families. *Biometrika*, 84, 641-652.
- Mostert P. J., Roux, J. J. J. and Bekker, A. (1999). Bayes Estimators of the Lifetime Parameters Using the Compound Rayleigh Model. *South African Statistical Journal*, 33, 117-138.
- Muth, E. J. (1977). Reliability Models with Positive Memory Derived from the Mean



Residual Life Function. *The Theory and Applications of Reliability*. C. P. Tsokos and I. Shimi, eds., Academic Press, Inc., New York, 401-435.

- NAG Fortran Library (1991). The Numerical Algorithms Group Limited, Mark 15, 1st ed., NAG Ltd, United Kingdom.
- Nelson, W. (1982). *Applied Life Data Analysis*. John Wiley & Sons, Inc., New York.
- Quesenberry, C. P. and Kent, J. (1982). Selecting Among Probability Distributions Used in Reliability. *Technometrics*, 24, 59-65.
- Rajarshi, S. and Rajarshi, M. B. (1988). Bathtub Distributions: A Review. *Communications in Statistics -Theory and Methods*, 17(8), 2597-2621.
- Smith, R. M. and Bain, L. J. (1975). An Exponential Power Life-Testing Distribution. *Communications in Statistics -Theory and Methods*, 4, 469-481.
- Stacy, E. W. (1962). A Generalization of the Gamma Distribution. *Annals of Mathematical Statistics*, 33(3), 1187-1192.
- Vaupel, J. W. (1990). Kindred Lifetimes: Frailty Models in Population Genetics. *Convergent Issues in Genetics and Demography*, J. Adams, D. A. Lam, A. I. Herman and P. E. Smouse, eds., London: Oxford University Press, 156-170.
- Vaupel, J. W., Manton, K. G. and Stallard, E. (1979). The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography*, 16, 439-454.





ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΕΑΕΚ



ΕΥΡΩΠΑΪΚΗ ΕΝΩΣΗ
ΣΥΓΧΡΗΜΑΤΟΔΟΤΗΣΗ
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



Η ΠΑΙΔΕΙΑ ΣΤΗΝ ΚΟΡΥΦΗ
Επιχειρησιακό Πρόγραμμα
Εκπαίδευσης και Αρχικής
Επαγγελματικής Κατάρτισης

This research was co-funded by the European Union in the framework of the program “ΗΡΑΚΛΕΙΤΟΣ” of the “Operational Program for Education and Initial Vocational Training” of the 3rd Community Support Framework of the Hellenic Ministry of Education, funded by 25% from national sources and by 75% from the European Social Fund (ESF).

