

Διπλωματική Εργασία

Μελέτη διαδικασιών Poisson και
Διωνυμικών με εφαρμογή σε
δεδομένα Αστροφυσικής

Κοκή Κωνσταντίνα
Πανεπιστήμιο Αθηνών
Τμήμα Μαθηματικών

Οκτώβριος 2012

Περίληψη

Η παρούσα εργασία ασχολείται με τη μελέτη των διαδικασιών Poisson και τη σύνδεση τους με τις διωνυμικές διαδικασίες. Οι προαναφερθείσες διαδικασίες είναι από τις πιο διαδεδομένες και πιο συχνά χρησιμοποιημένες όταν έχουμε προβλήματα με διακριτές τυχαιές μεταβλητές. Η ορθή μοντελοποίηση ενός προβλήματος αποτελεί τη βάση για την σωστή συμπερασματολογία και ενδεχομένως για την επίλυση του προβλήματος που μας απασχολεί. Αφού διαλέξουμε ένα μοντέλο που να μπορεί να εφαρμοστεί στο πρόβλημα προχωρούμε στη στατιστική ανάλυση των δεδομένων για να πάρουμε στο τέλος τα επιθυμητά αποτελέσματα. Σε δύσκολα προβλήματα θεωρούμε ότι η μπεϋζιανή προσέγγιση υπερτερεί έναντι των άλλων μεθόδων στατιστικής συμπερασματολογίας καθώς μπορεί να είναι ευέλικτη και να εφαρμοστεί σε περίπλοκα προβλήματα τα οποία αριθμητικά δεν μπορούν εύκολα να αντιμετωπιστούν. Σύμφωνα με την μπεϋζιανή στατιστική, η συμπερασματολογία γίνεται με βάση την εκ των υστέρων κατανομή των παραμέτρων, η οποία σε πολύπλοκα προβλήματα πορσεγγίζεται μέσω μεθόδων προσομοίωσης. Στη μελέτη αυτή παρουσιάζουμε μια εφαρμογή της μπεϋζιανής συμπερασματολογίας σε ένα περίπλοκο πρόβλημα της αστροφυσικής. Συγκεκριμένα ασχολούμαστε με την στην εκτίμηση της συνάρτησης φωτεινότητας των ενεργών πυρήνων των γαλαξιών, όπου η μοντελοποίηση του αριθμού των παρατηρήσιμων γαλαξιών γίνεται με διαδικασίες Poisson και διωνυμικές.

Abstract

In this master thesis we study the Poisson and binomial stochastic processes. These processes are widely used in the scientific community, to deal with problems where the random variables are distinct. In order to draw correct inferences we must choose an appropriate model. After choosing the model we believe is correct we wish to make inferences about the parameters of the model. In this study we follow the Bayesian approach to inference, that is all inferences are based on the posterior distribution of the model parameters. In multi-dimensional problems it is rarely possible to calculate the posterior distribution and compute analytically, statistics such as posterior means, posterior variances and posterior probabilities. Instead we may use simulation algorithms, the well known Markov Chain Monte Carlo (MCMC) algorithms in order to approximate the posterior distribution of interest. This master thesis is organized in two parts. In the first part we present the theory of Poisson processes and we make the connection with the binomial processes. Furthermore, we briefly describe the basic theory of Bayesian inference, Markov chains and MCMC. In the second part we present an application of the Bayesian inference in astrophysics. Specifically we present an estimation method of the luminosity function of the active galactic nuclei, where we use models, for the number of the observed galaxies, Poisson and Binomial processes.

Θα ήθελα να ευχαριστήσω την Δρ. Λουκία Μελιγκοτσίδου για τη βοήθεια της και το χρόνο που αφιέρωσε σε μένα. Ευχαριστώ επίσης τους καθηγητές του μεταπτυχιακού προγράμματος, οι οποίοι ήταν πάντα πρόθυμοι να λύσουν κάθε απορία μου και να μεταδώσουν τις γνώσεις τους. Τέλος θα ήθελα να ευχαριστήσω τον Δρ. Θανάση Ακύλα για την βοήθεια που προσέφερε για το αστροφυσικό κομμάτι της εργασίας μου.

Περιεχόμενα

1	Εισαγωγή	1
2	Η στοχαστική διαδικασία Poisson	3
2.1	Η κατανομή Poisson	4
2.1.1	Χώροι πιθανότητας για τη διαδικασία Poisson	7
2.1.2	Το “αναπόφευκτο” της διαδικασίας Poisson	9
2.2	Η διαδικασία Poisson στους γενικούς χώρους πιθανότητας	12
2.2.1	Ορισμοί και ιδιότητες της διαδικασίας Poisson	12
2.2.2	Το θεώρημα της Υπέρθεσης	15
2.2.3	Το θεώρημα της Απεικονίσεως	18
2.2.4	Η Διωνυμική-Πολυωνυμική διαδικασία	22
2.2.5	Ανεξάρτητες εκλεπτύνσεις της διαδικασίας Poisson	24
2.2.6	Το θεώρημα Ύπαρξης	25
2.3	Αθροίσματα Poisson	26
2.4	Η διαδικασία Poisson στην ευθεία των πραγματικών αριθμών	30
2.4.1	Ο νόμος των μεγάλων αριθμών	35
2.4.2	Οι μη ομογενείς διαδικασίες Poisson στην ευθεία των πραγματικών αριθμών	37
2.5	Μαρκαρισμένες διαδικασίες Poisson	38
2.5.1	Το θεώρημα Χρωματισμού	38
2.5.2	Η διαδικασία Poisson στο χώρο γινόμενο	39
3	Μπεϋζιανή Συμπερασματολογία και Μαρκοβιανές αλυσίδες	41
3.1	Μπεϋζιανή στατιστική	41
3.1.1	Καθορισμός των εκ των προτέρων κατανομών	44
3.1.2	Η σημειακή εκτίμηση στη μπεϋζιανή στατιστική	48
3.1.3	Διαστήματα και περιοχές αξιοπιστίας	49
3.1.4	Έλεγχος Υποθέσεων και σύγκριση μοντέλων	50
3.2	Μαρκοβιανές Αλυσίδες - Η μέθοδος MCMC	52
3.2.1	Βασικοί ορισμοί μαρκοβιανών αλυσίδων και το Εργοδι- κό θεώρημα	52

3.2.2	Ο αλγόριθμος Metropolis Hastings	54
3.2.3	Ο αλγόριθμος Gibbs	56
4	Το πρόβλημα της Αστροφυσικής	58
4.1	Μια εισαγωγή στο πρόβλημα	58
4.2	Το πρόβλημα της αστροφυσικής	60
4.2.1	Βασικές έννοιες της αστροφυσικής	60
4.2.2	Poisson δεδομένα	62
4.2.3	Διωνυμικά δεδομένα	68
4.2.4	Από κοινού μοντελοποίηση φωτεινότητας και ερυθρομε- τατότητας με μίξεις κανονικών κατανομών	73
5	Εφαρμογή	77
5.1	Schechter συνάρτηση φωτεινότητας για Poisson δεδομένα	78
5.2	Schechter συνάρτηση φωτεινότητας για διωνυμικά δεδομένα . .	84
5.3	Συμπεράσματα	89
Παράρτημα I		
	Χρήσιμες κατανομές	90
Παράρτημα II		
	Κώδικες Matlab για Schechter συνάρτηση φωτεινότητας	93
Παράρτημα III		
	Αλγόριθμος MCMC για τη συνάρτηση φωτεινότητας που ακολουθεί μίξη κανονικών κατανομών και κώδικας Matlab	99
Βιβλιογραφία		105

Κατάλογος Σχημάτων

5.1	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και Poisson δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$	80
5.2	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και Poisson δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$	81
5.3	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και Poisson δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$	82
5.4	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και Poisson δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$	83
5.5	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$	85
5.6	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$	86
5.7	Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$	87

- 5.8 Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{43}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$ 88

Κεφάλαιο 1

Εισαγωγή

Οι διαδικασίες Poisson είναι από τις συχνά χρησιμοποιημένες διαδικασίες στις θετικές επιστήμες όταν έχουμε διακριτά δεδομένα. Μοντέλα Poisson μπορούν να χρησιμοποιηθούν για να μοντελοποιήσουν φαινόμενα στην επιδημιολογία, στη βιοστατιστική, προβλήματα σε συστήματα τηλεπικοινωνιών, στη θεωρία κινδύνου, στις ουρές αναμονής, στην μηχανική, στην αστρονομία και σε πολλούς άλλους τομείς της επιστήμης. Οι πολύ καλές ιδιότητες της διαδικασίας όπως για παράδειγμα το ότι δεν χάνει τις καλές αλγεβρικές ιδιότητες που έχει μετά από μετασχηματισμούς, έχει κάνει τη διαδικασία Poisson ευρέως γνωστή και πολύ εύχρηστη. Εμείς, επιχειρούμε να εφαρμόσουμε τις Poisson και διωνυμικές διαδικασίες σε αστροφυσικά δεδομένα για την εκτίμηση της συνάρτησης φωτεινότητας των ενεργών πυρήνων των γαλαξιών.

Η εργασία ουσιαστικά χωρίζεται σε δύο μέρη. Στο πρώτο μέρος αναπτύσσεται η θεωρία για τις διαδικασίες Poisson, βασικά στοιχεία για τη μπεϋζιανή συμπερασματολογία, τις μαρκοβιανές αλυσίδες και τον αλγόριθμο MCMC. Στο δεύτερο μέρος της εργασίας παρουσιάζουμε την εφαρμογή σε αστρονομικά δεδομένα, όπως αναφέραμε πιο πάνω και αναλύουμε τα αποτελέσματά μας. Συγκεκριμένα χρησιμοποιούμε τη μπεϋζιανή συμπερασματολογία σε ένα πρόβλημα αστροφυσικής όπου τα δεδομένα μοντελοποιούνται με διαδικασίες Poisson ή διωνυμικές.

Στο κεφάλαιο 2 της εργασίας αυτής αναπτύσσουμε τη θεωρία γύρω από τη διαδικασία Poisson. Στην πρώτη παράγραφο του κεφαλαίου ορίζουμε τους χώρους πιθανότητας στους οποίους μπορεί να οριστεί μια διαδικασία Poisson. Η επόμενη παράγραφος περιέχει χρήσιμους ορισμούς, ιδιότητες και θεωρήματα που ισχύουν για τη διαδικασία Poisson αποδεικνύοντας πόσο εύχρηστη είναι. Σε αυτή την παράγραφο γίνεται σύνδεση με τη διαδικασία Bernoulli. Αναφέρουμε στην παράγραφο 3 ιδιότητες των αθροισμάτων διαδι-

κασιών Poisson, ορίζουμε τη χαρακτηριστική συνάρτηση και τη χαρακτηριστική συναρτησοειδή. Η παράγραφος 4 περιέχει την πολύ σημαντική περίπτωση της διαδικασίας Poisson όταν αυτή ορίζεται στην ευθεία των πραγματικών αριθμών. Τέλος στην παράγραφο 5 ορίζουμε τη διαδικασία στο χώρο γινόμενο και αναφέρουμε το θεώρημα του χρωματισμού.

Στο κεφάλαιο 3 κάνουμε μια σύντομη αναφορά στη μπεϋζιανή συμπερασματολογία καθώς επίσης και στις μαρκοβιανές αλυσίδες με επίκεντρο τη μέθοδο MCMC. Εδώ αναφέρουμε μόνο τη βασική θεωρία της μπεϋζιανής στατιστικής και μαρκοβιανών αλυσίδων έτσι ώστε ο αναγνώστης να μπορεί να παρακολουθήσει την εφαρμογή που κάνουμε στα κεφάλαια 4 και 5.

Το τέταρτο κεφάλαιο αρχίζει με ορισμούς της αστροφυσικής, τους οποίους θα χρησιμοποιήσουμε στην εφαρμογή που ακολουθεί. Η εφαρμογή αφορά την εκτίμηση της συνάρτησης φωτεινότητας των ενεργών πυρήνων των γαλαξιών. Για την εφαρμογή που ακολουθεί αντλήσαμε πληροφορίες και δεδομένα κυρίως από τα άρθρα των B.C.Kelly, X.Fan, M.Vestergaard (2008) και των H.L.Marshall, Y.Anvi, H.Tananbaum, G.Zamorani (1983). Παρουσιάζουμε τον υπολογισμό της συνάρτησης πιθανοφάνειας για Poisson δεδομένα και διωνυμικά δεδομένα όταν η συνάρτηση φωτεινότητας των γαλαξιών είναι Schechter και όταν ακολουθεί κάποια μίξη κανονικών κατανομών. Επίσης, δείχνουμε πως, με μπεϋζιανή ανάλυση προσομοιώνουμε την εκ των υστέρων κατανομή των παραμέτρων του μοντέλου που επιλέξαμε. Λόγω της περιπλοκότητας των εκ των υστέρων κατανομών των παραμέτρων χρησιμοποιούμε με τη μέθοδο MCMC για τη συμπερασματολογία. Στο κεφάλαιο 5 παρουσιάζουμε αναλυτικά τα αποτελέσματα της εφαρμογής μας.

Κεφάλαιο 2

Η στοχαστική διαδικασία Poisson

Στο κεφάλαιο αυτό μελετάμε τη στοχαστική διαδικασία Poisson. Ορίζουμε τη διαδικασία σε διάφορους χώρους πιθανότητας, αναφέρουμε κάποιες ιδιότητες της και μελετάμε τα βασικότερα θεωρήματα της διαδικασίας Poisson. Αντλήσαμε πληροφορίες κυρίως από δύο βιβλία, των J.F.C Kingman (1993) , R.L.Streit (2010). Ο αναγνώστης που επιθυμεί να βρει περισσότερες πληροφορίες και μερικές αποδείξεις τις οποίες παραλείπουμε, θεωρώντας ότι είναι έξω από τους σκοπούς της μελέτης μας, μπορεί να ανατρέξει στα προαναφερθέντα βιβλία.

Η κατανομή Poisson πήρε το όνομα της από το μαθηματικό Simeon-Denis Poisson (1781-1840). Ο Poisson δημιούργησε την κατανομή Poisson στην προσπάθειά του να προσεγγίσει τη διωνυμική κατανομή. Έστω ότι έχουμε ένα νόμισμα το οποίο ρίχνουμε n φορές. Με πιθανότητα p έρχεται κεφάλι και με $1 - p$ έρχεται γράμμα. Η τυχαία μεταβλητή που δηλώνει τον αριθμό των κεφαλών στο πείραμα ακολουθεί τη διωνυμική κατανομή

$$f(r; n, p) = \binom{n}{r} p^r (1 - p)^{n-r}$$

Αν θεωρήσουμε $\mu = np$ το μέσο αριθμό που έρχεται κεφάλι, όταν το n γίνεται μεγάλο και η πιθανότητα p μικρή τότε έχουμε

$$\pi_r(\mu) = \lim_{n \rightarrow \infty} f(r; n, p) = \lim_{n \rightarrow \infty} \binom{n}{r} p^r (1 - p)^{n-r} = \mu^r e^{-\mu} / r!$$

Η διαδικασία Poisson, δηλαδή, προσεγγίζει τη διωνυμική κατανομή, όταν το n τείνει στο άπειρο.

2.1 Η κατανομή Poisson

Ορισμός 2.1.1. Η κατανομή Poisson

Μια διακριτή μεταβλητή X ακολουθεί τη κατανομή Poisson με παράμετρο μ και συμβολίζεται με $\mathcal{P}(\mu)$ αν έχει συνάρτηση πιθανότητας:

$$P(X = n) = \pi_n(\mu) = e^{-\mu} \frac{\mu^n}{n!} \quad \text{όπου } n = 0, 1, 2, \dots \quad \mu > 0$$

Η μέση τιμή της τυχαίας μεταβλητής X συμπίπτει με την παράμετρο. Έχουμε δηλαδή ότι αν $X \sim \mathcal{P}(\mu) \Rightarrow \mathbb{E}(X) = \sum_{n=0}^{\infty} n \pi_n(\mu) = \mu$

Ορισμός 2.1.2. Πιθανογεννήτρια

Για κάθε αριθμό $|z| \leq 1$ ο τυχαίος αριθμός z^X είναι φραγμένος. Ορίζουμε την πιθανογεννήτρια της τυχαίας μεταβλητής X τη μέση τιμή $\mathbb{E}(z^X)$.

Στην περίπτωση της Poisson μεταβλητής θα έχουμε ότι

$$\mathbb{E}(z^X) = \sum_{n=0}^{\infty} \pi_n(\mu) z^n = e^{-\mu} \sum_{n=0}^{\infty} \frac{(\mu z)^n}{n!} = e^{-\mu(1-z)}, \quad 0 \leq \mu < \infty$$

Αν παραγωγίσουμε την πιθανογεννήτρια και θέσουμε το $z = 1$ μπορούμε να υπολογίσουμε τις μέσες τιμές:

$$\begin{aligned} \mathbb{E}(X) &= \mu \\ \mathbb{E}\{X(X-1)\} &= \mu^2 \quad \text{κ.ο.κ} \end{aligned}$$

Μια ενδιαφέρουσα και πολύ χρήσιμη ιδιότητα της κατανομής Poisson είναι η προσθετικότητα της. Αν έχουμε δύο ανεξάρτητες τυχαίες μεταβλητές X, Y τέτοιες ώστε $X \sim \mathcal{P}(\lambda)$, $Y \sim \mathcal{P}(\mu)$ τότε η κατανομή του αθροίσματος τους $X+Y$ θα ακολουθεί επίσης την κατανομή Poisson με παραμετρο το άθροισμα των παραμέτρων τους.

$$\begin{aligned}\mathbb{P}\{X + Y = n\} &= \sum_{r=0}^n \mathbb{P}\{X = r, Y = n - r\} = \sum_{r=0}^n \frac{e^{-\lambda} \lambda^r}{r!} \frac{e^{-\mu} \mu^{n-r}}{(n-r)!} \\ &= \frac{e^{-(\lambda+\mu)}}{n!} \sum_{r=0}^n \binom{n}{r} \lambda^r \mu^{n-r} = \frac{(\lambda + \mu)^n e^{-(\lambda+\mu)}}{n!}\end{aligned}$$

Όμοια μπορεί να αποδειχθεί ότι κάθε πεπερασμένο άθροισμα ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την κατανομή Poisson θα έχει επίσης την κατανομή Poisson με παράμετρο το άθροισμα των επιμέρους παραμέτρων.

Είμαστε τώρα σε θέση να αποδείξουμε το θεώρημα της προσθετικότητας για αριθμήσιμες τυχαίες μεταβλητές

Θεώρημα 2.1.1. Αριθμήσιμη Προσθετικότητα

Countable Additivity Theorem

Έστω X_j , $j = 1, 2, \dots$ ανεξάρτητες τυχαίες μεταβλητές οι οποίες ακολουθούν την $Poisson(\mu_j)$ ($X_j \sim \mathcal{P}(\mu_j)$, $j = 1, 2, \dots$).

Εάν το $\mu = \sum_{j=1}^{\infty} \mu_j$ συγκλίνει, τότε το άθροισμα των τυχαίων μεταβλητών $S = \sum_{j=1}^{\infty} X_j$ συγκλίνει με πιθανότητα 1 και έχει την κατανομή $\mathcal{P}(\mu)$. Στην αντίθετη περίπτωση που το μ αποκλίνει τότε και το S αποκλίνει.

Απόδειξη

Η απόδειξη γίνεται με επαγωγή στο n .

Συμβολισμοί:

$$S_n = \sum_{j=1}^n X_j, \quad \mu_n = \sum_{j=1}^n \mu_j$$

Το άθροισμα των τυχαίων μεταβλητών θα συγκλίνει στην κατανομή Poisson με παράμετρο μ_n ($S_n \sim \mathcal{P}(\mu_n)$)

Επομένως για κάθε r ισχύει ότι:

$$\mathbb{P}\{S_n \leq r\} = \sum_{k=0}^r \pi_k(\mu_n)$$

Παρατηρούμε ότι τα γεγονότα $\{S_n \leq r\}$ φθίνουν καθώς το n αυξάνει, για σταθερό r .

Επίσης έχουμε ότι:

$$\mathbb{P}\{S \leq r\} = \lim_{n \rightarrow \infty} \mathbb{P}\{S_n \leq r\} = \lim_{n \rightarrow \infty} \sum_{k=0}^r \pi_k(\mu_n)$$

Καθώς οι συναρτήσεις π_k είναι συνεχείς και από τη συνέχεια των μ_n έπεται ότι:

$$\mathbb{P}\{S \leq r\} = \sum_{k=0}^r \pi_k(\mu) \text{ και } \mathbb{P}\{S = r\} = \pi_r(\mu)$$

Συνεπώς το S είναι πεπερασμένο και συγκλίνει με πιθανότητα 1 στην $\mathcal{P}(\mu)$.

Αν όμως $\mu_n \rightarrow \infty$ τότε:

$$\sum_{k=0}^r \pi_k(\mu_n) = e^{-\mu_n} \sum_{k=0}^r \frac{\mu_n^k}{k!} \rightarrow 0 \Rightarrow \mathbb{P}\{S > r\} = 1$$

Δηλαδή το S αποκλίνει με πιθανότητα 1. □

Η πολωνυμική κατανομή συνδέεται άμεσα με την κατανομή Poisson μέσω του πεπερασμένου αθροίσματος ανεξάρτητων τυχαίων μεταβλητών. Ειδικότερα θα έχουμε ότι:

Αν $S = X_1 + \dots + X_n$, $X_j \sim \mathcal{P}(\mu_j)$, $j = 1, \dots, n$ και $S \sim \mathcal{P}(\mu)$ τότε:

$$\begin{aligned} \mathbb{P}\{X_1 = r_1 + \dots + X_n = r_n | S = s\} &= \prod_{j=1}^n \frac{\mu_j^{r_j} e^{-\mu_j}}{r_j!} / \frac{\mu^s e^{-\mu}}{s!} = \\ &= \frac{s!}{r_1! \dots r_n!} \left(\frac{\mu_1}{\mu}\right)^{r_1} \left(\frac{\mu_2}{\mu}\right)^{r_2} \dots \left(\frac{\mu_n}{\mu}\right)^{r_n} \end{aligned}$$

Αυτή η κατανομή είναι η πολωνυμική $M(\mu; p_1, \dots, p_n)$ όπου $p_j = \frac{\mu_j}{\mu}$

Αν $n = 2$ τότε η κατανομή θα είναι διωνυμική.

Άρα μπορούμε να κατασκευάσουμε τυχαίες μεταβλητές που ακολουθούν την πολυωνυμική (ή τη διωνυμική κατανομή), δεσμεύοντας κατάλληλα κάποια τυχαία μεταβλητή που ακολουθεί την κατανομή Poisson.

Έστω τώρα δύο τυχαίες μεταβλητές N, M . Θεωρούμε ότι η N ακολουθεί κατανομή Poisson με μέσο μ ($N \sim \mathcal{P}(\mu)$) και η M δεδομένης της N ακολουθεί διωνυμική κατανομή με παραμέτρους N, p ($M|N \sim \text{Binomial}(N, p)$), όπου p σταθερός θετικός αριθμός, μικρότερος ή ίσος της μονάδας.

Για $m, k \geq 0$ έχουμε:

$$\begin{aligned} \mathbb{P}\{M = m, N - M = k\} &= \mathbb{P}\{N = m + k\} \mathbb{P}\{M = m | N = m + k\} \\ &= \frac{e^{-\mu} \mu^{m+k}}{(m+k)!} \binom{m+k}{m} p^m (1-p)^k \\ &= \frac{e^{-\mu p} (\mu p)^m}{m!} \frac{e^{-\mu(1-p)} (\mu(1-p))^k}{k!} \end{aligned}$$

Άρα βλέπουμε ότι οι τυχαίες μεταβλητές $M, N - M$ είναι ανεξάρτητες μεταβλητές Poisson με μέσους $\mu p, \mu(1-p)$ αντίστοιχα.

2.1.1 Χώροι πιθανότητας για τη διαδικασία Poisson

Για να ορίσθει αυστηρά η διαδικασία Poisson πρέπει να καθοριστεί ο χώρος πιθανότητας, δηλαδή μια τριάδα $(\Omega, \mathcal{F}, \mathbb{P})$. Με Ω συμβολίζουμε το σύνολο που περιέχει όλα τα δυνατά εξαγόμενα (δειγματικός χώρος), με $\mathcal{F}$ μια οικογένεια υποσυνόλων του Ω που είναι επιπλέον σ-άλγεβρα και τέλος με $\mathbb{P}$ συμβολίζουμε το μέτρο πιθανότητας, $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$.

Πριν ορίσουμε λοιπόν με αυστηρό τρόπο τη διαδικασία Poisson παραθέτουμε, για λόγους πληρότητας, ορισμένους βασικούς ορισμούς της θεωρίας πιθανοτήτων.¹

¹Ο αναγνώστης μπορεί να αντρέξει σε βιβλία Θεωρίας Πιθανοτήτων για περισσότερες πληροφορίες. Προτείνουμε τα βιβλία: Ν. Παπαδάτου (2006) και το P. Billingsley (1996)

Ορισμός 2.1.3. Τυχαία μεταβλητή

Έστω $(\Omega, \mathcal{F}, \mathbb{P})$ ένας χώρος πιθανότητας και $X: \Omega \rightarrow \mathbb{R}$ τυχούσα συνάρτηση. Η X θα καλείται τυχαία μεταβλητή αν ισχύει ότι $\{\omega \in \Omega : X(\omega) \leq b\} \in \mathcal{F}$ για κάθε $b \in \mathbb{R}$.²

Σε αντιστοιχία με τον πιο πάνω ορισμό, ορίζουμε το τυχαίο διάνυσμα

Ορισμός 2.1.4. Τυχαίο διάνυσμα

Μια συνάρτηση $\underline{X} = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ καλείται τυχαίο διάνυσμα στον μετρήσιμο χώρο $(\Omega, \mathcal{F})$ όταν οι συναρτήσεις $X_j, j = 1, \dots, n : \Omega \rightarrow \mathbb{R}$ είναι τυχαίες μεταβλητές.

Ορισμός 2.1.5. Στοχαστική διαδικασία

Στοχαστική διαδικασία μια οικογένεια τυχαίων μεταβλητών $\{X(\alpha), \alpha \in T\}$ ορισμένων σε κοινό χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mathbb{P})$. Οι τυχαίες μεταβλητές $X(\alpha)$ παίρνουν τιμές σε ένα σύνολο S το οποίο ονομάζεται σύνολο ή χώρος καταστάσεων, ενώ το α ανήκει στο σύνολο T , το οποίο ονομάζουμε παραμετρικό σύνολο ή παραμετρικό χώρο.

Προχωράμε λοιπόν στον ορισμό της διαδικασίας Poisson

Στοχαστική διαδικασία Poisson

Η στοχαστική διαδικασία Poisson με χώρο καταστάσεων S , ορισμένη σε ένα χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$ είναι μια συνάρτηση Π από τον Ω στο σύνολο S^∞ των μετρήσιμων υποσυνόλων του S .

Ορίζουμε με $N(A)$ μια τυχαία μεταβλητή ως:

²Με $\mathbb{R}$ συμβολίζουμε το σύνολο των πραγματικών αριθμών

$$N(A) = \# \{ \Pi(\omega) \cap A \},$$

όπου με $\#$ συμβολίζουμε το πλήθος των σημείων στο σύνολο. Το A είναι το σύνολο δοκιμών το οποίο είναι υποσύνολο του $\mathcal{F}$. Μπορεί να είναι οποιοδήποτε σύνολο με κάποιες καλές ιδιότητες. Απαιτούμε να είναι φραγμένο και χωρίς κενά ή “ τρύπες ”. Για παράδειγμα μπορεί να είναι κάποιο κυρτό σύνολο, ένωση ξένων σφαιρών κτλ.

Ισχύει ότι η τυχαία μεταβλητή ορίζεται ως

$$N(A) : \Omega \rightarrow \{0, 1, 2, \dots, \infty\}$$

και απαιτούμε η συνάρτηση N να είναι μετρήσιμη για κάθε σύνολο A . Επομένως έπεται ότι για κάθε σύνολο A και για κάθε n , το σύνολο των ω με την ιδιότητα $\{\omega; N(A) = n\}$ θα ανήκουν στο $\mathcal{F}$ και άρα τα $N(A)$ είναι τυχαίες μεταβλητές.

Για παράδειγμα για αν ο χώρος καταστάσεων είναι η ευθεία των πραγματικών αριθμών, για να ισχύει η αρκεί $\{\omega; N(A) = n\} \in \mathcal{F}$ το σύνολο A να είναι οποιοδήποτε ανοικτό διάστημα, ενώ αν ο χώρος καταστάσεων είναι το $\mathbb{R}^2$ αρκεί το σύνολο καταστάσεων να είναι οποιοδήποτε ανοικτό ορθογώνιο παραλληλόγραμμο.

2.1.2 Το “αναπόφευκτο” της διαδικασίας Poisson

Ιδιότητα 2.1.1. Η ιδιότητα των ανεξάρτητων διασκορπίσεων

The independent scattering property

Έστω μία στοχαστική διαδικασία Z που ορίζεται στο χώρο $S \subseteq \mathbb{R}^d$ και έστω $A \subseteq S$. Συμβολίζουμε με $Z_A = Z \cap A$ τον περιορισμό της Z στο A .

Λέμε ότι μία στοχαστική διαδικασία Z έχει την ιδιότητα των ανεξάρτητων διασκορπίσεων αν:

Για ξένα υποσύνολα $A_1, A_2, \dots$ του S οι περιορισμοί των στοχαστικών διαδικασιών $Z_{A_1}, Z_{A_2}, \dots$ είναι ανεξάρτητες στοχαστικές διαδικασίες.

Θεώρημα 2.1.2. Το “ αναπόφευκτο” της στοχαστικής διαδικασίας Poisson

The inevitable of the Poisson process

Αν μια στοχαστική διαδικασία ικανοποιεί την ιδιότητα των ανεξάρτητων διασκορπίσεων και ο αριθμός των σημείων οποιοδήποτε φραγμένου συνόλου A είναι πεπερασμένος και μεγαλύτερος του μηδενός με πιθανότητα 1, τότε ο αριθμός των σημείων της διαδικασίας δεδομένου του συνόλου A είναι οπωσδήποτε διαδικασία Poisson.

Απόδειξη

Έστω κάποιο σύνολο $A \neq \emptyset$ χωρίς κενά ή " τρύπες " και ορίζουμε μία οικογένεια συνόλων A_t , $t \geq 0$ ως :

$$A_t = \bigcup_{\alpha \in A} \{x \in \mathbb{R}^m : \|x - \alpha\| \leq t\}$$

όπου με $\|\cdot\|$ συμβολίζουμε την συνήθη ευκλείδεια απόσταση. Δηλαδή το σύνολο A_t ορίζει κύκλο ακτίνας t . Καθώς το A δεν παρουσιάζει κενά το σύνορο του A_s περιέχεται στο A_t για $t > s$.

Θεωρούμε τις συναρτήσεις $p_n(t) = \mathbb{P}\{N(A_t) = n\}$ και $q_n(t) = \mathbb{P}\{N(A_t) \leq n\}$.

Η τυχαία μεταβλητή $N(A_t)$ είναι αύξουσα συνάρτηση του t και συνεπώς η συνάρτηση $q_n(t)$ είναι φθίνουσα, ενώ η συνάρτηση της μέσης τιμής που $(\mu(t) \equiv E[N(A_t)])$ είναι αύξουσα.

Οι p_n, q_n είναι διαφορίσιμες σχεδόν παντού. Για μικρό $h > 0$ η τυχαία μεταβλητή $N(A_t)$ πηδάει από τον n -οστό δακτύλιο στο $(n+1)$ -οστό δακτύλιο που ορίζουν οι ακτίνες $t, t+h$ με πιθανότητα $q_n(t) - q_n(t+h) \geq 0$.

Αν δεν λάβουμε υπόψη μας την πιθανότητα να πέσουν δύο ή περισσότερα σημεία στο δακτύλιο τότε η πιθανότητα αυτή δίνεται και από τη σχέση $\mu(t+h) - \mu(t)$.

Αν η πιθανότητα αυτή είναι ανεξάρτητη του αριθμού των σημείων στο σύνολο A_t τότε ισχύει ότι :

$$q_n(t) - q_n(t+h) = p_n(t) \{\mu(t+h) - \mu(t)\}$$

Καθώς $h \rightarrow 0$ παίρνοντας μερική παράγωγο ως προς t έχουμε:

$$-\frac{dq_n}{dt} = p_n \frac{d\mu}{dt}$$

Εξ' ορισμού θα ισχύει ότι: $p_n = q_n - q_{n-1}$.

Αντικαθιστώντας τη σχέση αυτή στη $-\frac{dq_n}{dt} = p_n \frac{d\mu}{dt}$ παίρνουμε ότι:

Για $n \geq 1$

$$\frac{dp_n(t)}{dt} = (p_{n-1} - p_n) \frac{d\mu(t)}{dt}$$

Για $n = 0$ ισχύει ότι $q_0(t) = p_0(t)$

$$\frac{dp_0}{dt} = -p_0 \frac{d\mu}{dt} \Leftrightarrow \frac{d}{dt}(\mu(t) + \log[p_0(t)]) = 0$$

Αφού $p_0(0) = 1$ και $\mu(0) = 0$ έπεται ότι

$$\log(p_0) + \mu = 0 \text{ και άρα } p_0(t) = e^{-\mu(t)} \quad p_{n-1}(t) \frac{d\mu(t)}{dt} = \frac{dp_n(t)}{dt} + p_n(t) \frac{d\mu(t)}{dt}$$

Πολλαπλασιάζοντας και τα δύο μέλη με $e^{\mu(t)}$ παίρνουμε ότι:

$$\frac{d}{dt} (p_n(t) e^{\mu(t)}) = p_{n-1}(t) e^{\mu(t)} \frac{d\mu(t)}{dt}$$

Ολοκληρώνοντας,

$$p_n(t) = e^{-\mu(t)} \int_0^t p_{n-1}(s) e^{\mu(s)} \frac{d\mu(s)}{ds} ds$$

Επαγωγικά, αρχίζοντας από την ισότητα $p_0(t) = e^{-\mu(t)}$ μπορεί να αποδειχθεί ότι $p_n(t) = e^{-\mu(t)} \frac{\mu^n(t)}{n!}$

Άρα δείξαμε ότι η τυχαία μεταβλητή $N(A_t)$ ακολουθεί την $\mathcal{P}(\mu(t))$ κατανομή. $\square$

Η προαναφερθήσα οικογένεια των συνόλων A_t τα οποία καλέσαμε σύνολα δοκιμών είναι πολύ μεγάλη. Αυτός είναι και ένας από τους λόγους που μπορούμε να μοντελοποιήσουμε πολλά προβλήματα που αντιμετωπίζουμε καθημερινά χρησιμοποιώντας τη διαδικασία Poisson.

2.2 Η διαδικασία Poisson στους γενικούς χώρους πιθανότητας

2.2.1 Ορισμοί και ιδιότητες της διαδικασίας Poisson

Η διαδικασία Poisson συνήθως έχει χώρο καταστάσεων S κάποιον ευκλείδιο χώρο διάστασης d (συνήθως $\mathbb{R}^d$). Ο χώρος καταστάσεων S δεν είναι απαραίτητο να είναι ο ευκλείδιος αλλά αρκεί τα υποσύνολα του να είναι μετρήσιμα και να έχουν μορφή σ-άλγεβρας με την εννοια ότι:

1. Το κενό σύνολο ($\emptyset$) είναι μετρήσιμο.
2. Το συμπλήρωμα μετρήσιμου συνόλου (S^c) είναι επίσης μετρήσιμο.
3. Η αριθμήσιμη ένωση μετρήσιμων συνόλων ($S_{A_1} \cup S_{A_2} \cup \dots$) είναι επίσης μετρήσιμο σύνολο.

Τέλος είναι απαραίτητο να εξασφαλίσουμε ότι υπάρχουν αρκετά μετρήσιμα σύνολα για να διακρίνουν όλα τα σημεία. Κάνουμε λοιπόν την υπόθεση ότι το $D = \{(x, \psi); x = \psi\}$ είναι μετρήσιμο στο χώρο γινόμενο $S \times S$. Η υπόθεση αυτή μας εξασφαλίζει ότι κάθε μονοσύνολο $\{x\}$ είναι μετρήσιμο στον S .

Θεωρούμε τη συνάρτηση $N(A) = \#\{\Pi \cap A\}$ η οποία είναι καλά ορισμένη και ονομάζεται απαριθμήτρια συνάρτηση (ορίζει απαριθμήτρια διαδικασία).

Ορισμός 2.2.1. Διαδικασία Poisson

Κάτω από τις υποθέσεις που αναφέρθηκαν αμέσως πιο πάνω η στοχαστική διαδικασία Poisson στο S είναι ένα τυχαίο αριθμήσιμο υποσύνολο Π του S τέτοιο ώστε:

1. Για κάθε συλλογή από ξένα αριθμήσιμα υποσύνολα $A_1, A_2, \dots, A_n$ του S , οι τυχαίες μεταβλητές $N(A_1), N(A_2), \dots, N(A_n)$ είναι ανεξάρτητες.
2. Η απαριθμήτρια διαδικασία $N(A)$ έχει την κατανομή Poisson, ($P(\mu)$), όπου $\mu = \mu(A)$ και $0 \leq \mu \leq \infty$

Αν το $\mu(A)$ είναι πεπερασμένο, τότε το σύνολο $\Pi \cap A$ είναι με πιθανότητα 1 πεπερασμένο και κενό εάν $\mu(A) = 0$. Εάν $\mu(A) = \infty$ τότε το $\Pi \cap A$ είναι άπειρο και αριθμήσιμο με πιθανότητα 1.

Εξ' ορισμού η μέση τιμή της απαριθμήτριας διαδικασίας είναι

$$E\{N(A)\} = \mu(A)$$

Στην περίπτωση που έχουμε σύνολα $A_1, A_2, \dots$ τα οποία είναι ξένα μεταξύ τους και ισχύει ότι $A = \bigcup_{n=1}^{\infty} A_n$ τότε αφού $N(A) = \sum_{n=1}^{\infty} N(A_n)$ είναι προφανές ότι η μέση τιμή της απαριθμήτριας διαδικασίας θα είναι το άθροισμα:

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n)$$

Εάν γνωρίζουμε το μ μπορούμε να βρούμε την από κοινού κατανομή των $N(A)$ για κάθε σύνολο A .

Για να το δούμε αυτό υποθέτουμε ότι έχουμε τα μετρήσιμα σύνολα $A_1, A_2, \dots, A_n$ όχι κατ' ανάγκη είναι ξένα μεταξύ τους τότε παίρνουμε:

$$B_i = A_1^* \cap A_2^* \cap \dots \cap A_n^*$$

όπου $A_j^*, j = 1, 2, \dots, n$ είναι τα είτε τα A_j ή το συμπλήρωμα τους.

Συνεπώς τα B_i είναι 2^n το πλήθος.

Έχουμε λοιπόν ότι:

$$A_j = \bigcup_{i \in \gamma_j} B_i \text{ με } \gamma_j \subseteq \{1, 2, \dots, 2^n\} \Rightarrow N(A) = \sum_{i \in \gamma_j} N(B_i)$$

Από την κατασκευή τους οι τυχαίες μεταβλητές $N(B_i)$ είναι ανεξάρτητες και ακολουθούν την κατανομή $\mathcal{P}(\mu(B_i))$. Η μέση τιμή θα είναι

$$\mu(A) = \sum_{n \in \gamma_j} \mu(B_n)$$

Έτσι η μπορούμε να ορίσουμε την πιθανότητα κάθε γεγονότος χρησιμοποιώντας τις τυχαίες μεταβλητές $N(A_1), N(A_2), \dots, N(A_n)$.

Ονομάζουμε το μ , μέσο μέτρο για τη διαδικασία Poisson.

Είναι σημαντικό να παρατηρήσουμε ότι δεν μπορεί κάθε μέτρο πιθανότητας να είναι μέσο μέτρο. Για παράδειγμα, έστω ότι το μέτρο μ στον S έχει ένα άτομο στο x τέτοιο ώστε για το μονοσύνολο $\{x\}$ θα είναι $m = \mu(\{x\}) > 0$. Από το (2) του ορισμού 2.2.1. για τη διαδικασία Poisson αν πάρουμε $A = \{x\}$ τότε η διαδικασία Poisson θα έχει μέσο μέτρο μ και θα είναι τέτοιο ώστε

$$\mathbb{P}\{N(\{x\}) \geq 2\} = 1 - e^{-m} - me^{-m} > 0$$

Καθώς όμως το A είναι μονοσύνολο και $N(A) = \#\{\Pi \cap A\} \leq 1$ από τον τρόπο που ορίσαμε την τυχαία μεταβλητή $N(A)$ έπεται ότι το μ δεν μπορεί να είναι μέσο μέτρο.

Για να μπορεί το μέτρο να είναι μέσο μέτρο πρέπει να είναι μη ατομικό υπό την έννοια ότι $\mu(\{x\}) = 0$ για κάθε $x \in S$

Στην περίπτωση που ο χώρος καταστάσεων είναι ο d -διάστατος Ευκλείδειος χώρος τότε το μέσο μέτρο ονομάζεται ρυθμός (rate) ή ένταση (intensity).

Αν το λ είναι μια θετικά ορισμένη μετρήσιμη συνάρτηση του S , τότε:

$$\mu(A) = \int_A \lambda(x) dx \text{ (με } dx \text{ συμβολίζουμε το } dx_1 dx_2 \dots dx_n).$$

Αν η συνάρτηση λ είναι συνεχής τότε σε μικρή περιοχή του A στο x τότε ισχύει ότι

$$\mu(A) \sim \lambda(x)|A|$$

όπου με $|A|$ συμβολίζουμε το μέτρο (αν $d=1$ είναι μήκος, αν $d=2$ εμβαδόν, αν $d=3$ όγκος) .

Το $\lambda(x)|A|$ είναι προσεγγιστικά η πιθανότητα ένα σημείο της Π να πέσει στο σύνολο A . Με αυτή την έννοια η πιθανότητα αυτή είναι μεγαλύτερη όταν το λ είναι μεγάλο.

Αν το λ είναι σταθερά τότε $\mu(A) = \lambda|A|$ και η διαδικασία αυτή ονομάζεται ομογενής ή ομοιόμορφη διαδικασία Poisson.

Παράδειγμα

Έστω ότι το $S = \mathbb{R}$ και υποθέτουμε ότι το μέσο μέτρο μ είναι πεπερασμένο στο φραγμένο σύνολο A . Το μ τότε καθορίζεται μονοσήμαντα από τις τιμές στο διάστημα $(a, b]$

Γράφουμε

$$M(t) = \begin{cases} \mu(0, t] = \mathbb{E}\{N(0, t]\} & \text{αν } t \geq 0 \\ -\mu(0, t] = -\mathbb{E}\{N(0, t]\} & \text{αν } t < 0 \end{cases}$$

όπου M αύξουσα συνάρτηση.

Το μέσο μέτρο μ ορίζεται ως

$$\mu(a, b] = M(b) - M(a), \quad (a < b)$$

και ονομάζεται μέτρο Stieltjes

Όπως αναφέραμε πιο πάνω απαιτούμε το μέτρο να είναι μη ατομικό. Στο παράδειγμα αυτό το μέτρο μ είναι μη ατομικό αν και μόνο αν η συνάρτηση M είναι συνεχής. Πολλές φορές η συνάρτηση M μπορεί να εκφραστεί ως:

$$M(t) = \int_0^t \lambda(x) dx$$

Όταν η M εκφράζεται σε αυτή τη μορφή τότε $\mu = M$. Σε ομογενή διαδικασία Poisson η συνάρτηση λ είναι, όπως αναφέραμε και πιο πάνω, σταθερά. Άρα το μέσο μέτρο της διαδικασίας θα είναι:

$$\mu(t) = M(t) = \lambda t$$

□

Η στοχαστική διαδικασία Poisson έχει το πλεονέκτημα ότι έχει ενδιαφέρουσες ιδιότητες οι οποίες επιτρέπουν τη χρήση της σε πολλές εφαρμογές ενώ ταυτόχρονα οι υπολογισμοί των πιθανοτήτων και άλλων στοιχείων που μας ενδιαφέρουν γίνονται πολύ απλά. Στις επόμενες σελίδες θα αναφέρουμε θεωρήματα και επιπλέον ιδιότητες της διαδικασίας Poisson.

2.2.2 Το θεώρημα της Υπέρθεσης

Λήμμα 2.2.1. Αμοιβαία ξένες διαδικασίες

The disjointness lemma

Έστω Π_1, Π_2 ανεξάρτητες Poisson διαδικασίες σε χώρο καταστάσεων S και A ένα μετρήσιμο σύνολο δοκιμών. Έστω επίσης $\mu_1(A), \mu_2(A)$ οι παράμετροι των διαδικασιών. Θεωρούμε ότι οι $\mu_1(A), \mu_2(A)$ είναι πεπερασμένες. Τότε οι Π_1, Π_2 ξένες διαδικασίες με πιθανότητα 1 στον A δηλαδή

$$P\{\Pi_1 \cap \Pi_2 \cap A = \emptyset\} = 1$$

Απόδειξη

Έστω A^f το σύνολο όλων των πεπερασμένων υποσυνόλων Λ του A . Το A^f το κατασκευάζουμε έτσι ώστε να περιέχει την μικρότερη οικογένεια των μετρήσιμων συνόλων που κάνουν τη συνάρτηση $\Lambda \mapsto \#(\Lambda \cap B)$ μετρήσιμη για όλα τα σύνολα $B \subseteq A$.

Τότε το $\Pi_1 \cap A$ είναι τυχαίο στοιχείο του A^f με κατανομή P_1 στον A^f . Όμοια το $\Pi_2 \cap A$ έχει κατανομή P_2 .

Τα Π_1, Π_2 είναι ανεξάρτητα και η από κοινού κατανομή των $\Pi_1 \cap A, \Pi_2 \cap A$ είναι η $P_1 \times P_2$ στο χώρο γινόμενο $A^f \times A^f$

Ορίζουμε την απεικόνιση

$$\eta : A^f \times A^f \rightarrow (A \times A)^f \text{ με } \eta(\Lambda_1, \Lambda_2) = \Lambda_1 \times \Lambda_2$$

Θέλουμε να δείξουμε ότι η η είναι μετρήσιμη. Αρκεί να δείξουμε ότι η $\eta^{-1} \{ \Lambda; \#(\Lambda \cup C) = n \}$ είναι μετρήσιμη στο χώρο $A^f \times A^f$ για κάθε μετρήσιμο σύνολο $C \subseteq A \times A$.

Για $C = B_1 \times B_2$ το ζητούμενο ισχύει για κάποια n_1, n_2 αφού αν $n_1 n_2 = n$ και

$$\#(\Lambda_1 \cup B_1) = n_1, \quad \#(\Lambda_2 \cup B_2) = n_2$$

τότε το πλήθος $\# \{ \eta(\Lambda_1, \Lambda_2) \cup (B_1 \times B_2) \} = n$

Επίσης η τάξη των συνόλων C είναι κλειστή και περιέχει όλα τα μετρήσιμα C στον χώρο γινόμενο $A \times A$.

Ειδικότερα, έστω διαγώνιο σύνολο D είναι μετρήσιμο στο χώρο $S \times S$. Ο περιορισμός $D_A = D \cap (A \times A)$ είναι μετρήσιμο στο χώρο $A \times A$ έτσι ώστε το $J = \eta^{-1} \{ \Lambda; \#(\Lambda \cup D_A) = 0 \}$ να είναι μετρήσιμο στο χώρο $A^f \times A^f$.

Από τον τρόπο κατασκευής των P_1, P_2, J η σχέση που θέλουμε να αποδείξουμε (δηλαδή η $P \{ \Pi_1 \cap \Pi_2 \cap A = \emptyset \} = 1$) είναι ισοδύναμη με τη σχέση $(P_1 \times P_2)(J) = 1$

Για όλα τα Λ_1 που βρίσκονται εκτός του P_1 θα ισχύει

$$P_2 \{ \Lambda_2; (\Lambda_1, \Lambda_2) \in J \} = 1 \Rightarrow P_2 \{ N_2(\Lambda_1) = 0 \} = 1$$

Η τελευταία σχέση ισχύει καθώς Λ_1 είναι πεπερασμένο και επομένως $\mu(\Lambda_1) = 0$ για όλα τα Λ_1 .

Η απόδειξη έχει ολοκληρωθεί. $\square$

Θεώρημα 2.2.1. Το θεώρημα της Υπέρθεσης

The superposition Theorem

Το θεώρημα της υπέρθεσης (ή αλλιώς το θεώρημα της Επαλληλίας) βρίσκεται σε αντιστοιχία με το θεώρημα της αριθμήςιμης προσθετικότητας.

Έστω $\Pi_1, \Pi_2, \dots$ μία αριθμήςιμη συλλογή από ανεξάρτητες Poisson διαδικασίες στο χώρο καταστάσεων S και έστω ότι οι Π_n έχουν μέσο μέτρο μ_n για κάθε n . Τότε η ένωση των διαδικασιών $\Pi = \bigcup_{n=1}^{\infty} \Pi_n$ είναι επίσης διαδικασία Poisson με μέσο μέτρο $\mu = \sum_{n=1}^{\infty} \mu_n$.

Απόδειξη

Έστω $N_n(A)$ ο αριθμός σημείων της διαδικασίας Π_n σε ένα μετρήσιμο σύνολο A .

Αν ισχύει ότι $\mu(A) < \infty$ για όλα τα n από το λήμμα των αμοιβαία ξένων διαδικασιών τα τυχαία σύνολα Π_n είναι ξένα μεταξύ τους στο A και τότε

$$N(A) = \sum_{n=1}^{\infty} N_n(A)$$

Από το θεώρημα αριθμήςιμης προσθετικότητας η $N(A)$ έχει την κατανομή $\mathcal{P}(\mu)$ με $\mu = \mu(A) = \sum_{n=1}^{\infty} \mu_n$.

Αν $\mu_n(A) = \infty$ για κάποια n τότε

$$N_n(A) = N(A) = \infty \text{ και η } N(A) = \sum_{n=1}^{\infty} N_n(A)$$

ισχύει τετριμμένα.

Επομένως για να δείξουμε το θεώρημα αρκεί να δείξουμε ότι οι τυχαίες μεταβλητές $N(A_1), N(A_2), \dots, N(A_k)$ είναι ανεξάρτητες αν τα σύνολα A_j είναι ξένα μεταξύ τους. Αυτό όμως είναι προφανές καθώς τα $N_n(A_j)$, $n =$

$1, 2, \dots, j = 1, 2, \dots, k$ είναι ανεξάρτητες και από τον ορισμό των $N(A_j)$ είναι υποσύνολα των ξένων μεταξύ τους μεταβλητών για διαφορετικά j . $\square$

Θεώρημα 2.2.2. Το θεώρημα του Περιορισμού

The restriction theorem

Έστω Π μία διαδικασία Poisson με μέσο μέτρο μ σε χώρο καταστάσεων S και έστω S_1 ένα μετρήσιμο υποσύνολο του S . Τότε το τυχαίο αριθμήσιμο σύνολο $\Pi_1 = \Pi \cap S_1$ είναι επίσης διαδικασία Poisson με μέσο μέτρο $\mu_1(A) = \mu(A \cap S_1)$ ή διαδικασία Poisson στο χώρο καταστάσεων S_1 που το μέσο μέτρο του είναι ο περιορισμός του μ στο S_1 .

Απόδειξη

Η απόδειξη είναι άμεση καθώς από τον ορισμό (2.2.1) για τη διαδικασία Poisson βλέπουμε ότι αν τα 1, 2 ισχύουν για την Π τότε θα ισχύουν και για τον περιορισμό Π_1 . $\square$

2.2.3 Το θεώρημα της Απεικονίσεως

Μία ιδιότητα των διαδικασιών Poisson είναι ότι αν ο χώρος καταστάσεων αλλάξει και μεταφερθούμε σε άλλο χώρο καταστάσεων, τότε τα σημεία που αντιστοιχούν στον καινούργιο χώρο καταστάσεων σχηματίζουν επίσης μία διαδικασία Poisson.

Πριν προχωρήσουμε στο θεώρημα για την μεταφορά της διαδικασίας Poisson από ένα χώρο καταστάσεων σε κάποιο άλλο χώρο, κρίνουμε αναγκαίο να αναφερθούμε σε κάποια στοιχεία για την καλύτερη κατανόηση του θεωρήματος.

Έστω διαδικασία Poisson Π με χώρο καταστάσεων S , μέσο μέτρο μ και έστω

f μία συνάρτηση από το χώρο καταστάσεων S στο χώρο T , $f : S \rightarrow T$. Υποθέτουμε ότι οι πιο πάνω χώροι ικανοποιούν τις συνθήκες που αναφέρθηκαν πριν για να είναι η διαδικασία Π , Poisson. Απαιτούμε η συνάρτηση f να είναι μετρήσιμη υπό την έννοια ότι η αντίστροφη εικόνα

$$f^{-1}(B) = \{x \in S; f(x) \in B\}$$

είναι μετρήσιμο υποσύνολο του S για κάθε μετρήσιμο υποσύνολο $B \subseteq T$.

Τα σημεία $f(x)$, $x \in \Pi$ συγκροτούν ένα τυχαίο αριθμήσιμο σύνολο $f(\Pi) \subseteq T$. Το ερώτημα είναι αν στο σύνολο T θα έχουμε και πάλι διαδικασία Poisson

Θεωρούμε το $N^*(B) = \#\{f(\Pi \cap B)\}$ το πλήθος των σημείων του $f(\Pi)$ στο σύνολο B .

Αν τα σημεία είναι διακριτά τότε μπορούμε να γράψουμε τον πιο πάνω αριθμό ως: $N^*(B) = \#\{x \in \Pi; f(x) \in B\} = N(f^{-1}(B))$

Από τον ορισμό τότε, το $N^*(B)$ είναι διαδικασία Poisson, $\mathcal{P}(\mu^*)$, (όπου $\mu^* = \mu^*(B) = \mu(f^{-1}(B))$) καθώς τα $B_1, B_2, \dots, B_k$ είναι ξένα, και άρα και οι αντίστροφες εικόνες τους $N^*(B_j)$ είναι ανεξάρτητες.

Έτσι για να είναι η $f(\Pi)$ είναι διαδικασία Poisson στο T , αρκεί τα σημεία $f(x)$ να είναι διακριτά.

Εντούτοις η υπόθεση ότι τα σημεία είναι διακριτά, δεν είναι τετριμμένη καθώς μπορούμε να βρούμε συνάρτηση (έστω μια σταθερή συνάρτηση) που αντιστοιχεί τα σημεία του S μέσω της Π σε ένα σταθερό σημείο του T .

Γενικότερα αν τα μ, f είναι τέτοια ώστε το επαγόμενο μέτρο μ^* να έχει ένα άτομο στο $t \in T$ τότε το $A = f^{-1}(t)$ έχει μη-μηδενικό μέτρο $m = \mu(A) = \mu^*({t})$ και με μη μηδενική πιθανότητα $1 - e^{-m} - me^{-m}$ τουλάχιστον δύο σημεία της Π που είναι στο A αντιστοιχούν στο ίδιο t .

Άρα είναι απαραίτητο τουλάχιστον να υποθέσουμε ότι το μ^* είναι μη-ατομικό. Χρειαζόμαστε και μία συνθήκη ώστε το μέσο μέτρο να μην είναι άπειρο και είναι αρκετό να υποθέσουμε ότι το μ είναι σ-πεπερασμένο υπό την έννοια ότι το S μπορεί να γραφεί ως αριθμήσιμη ένωση των μετρήσιμων συνόλων S_n , με $\mu(S_n) < \infty$.

Τελικά έχουμε το επόμενο θεώρημα:

Θεώρημα 2.2.3. Το θεώρημα της απεικονίσεως

The mapping theorem

Έστω Π μια διαδικασία Poisson με σ-πεπερασμένο μέσο μέτρο μ στο χώρο καταστάσεων S και έστω $f : S \rightarrow T$ μία μετρήσιμη συνάρτηση τέτοια ώστε το επαγόμενο μέτρο $\mu(A) = E[N(A)]$ να είναι μη ατομικό. Τότε η $f(\Pi)$ είναι διαδικασία Poisson στο T με μέσο μέτρο $\mu^* = \mu^*(B) = \mu(f^{-1}(B))$, $B \subseteq T$.

Απόδειξη

Έστω ότι $\mu(S) < \infty$ και έστω A κάποιο μετρήσιμο σύνολο. Συμβολίζουμε με A^c το συμπλήρωμα του συνόλου A . Ορίζουμε με Π_1, Π_2 τους περιορισμούς της Π στα σύνολα A, A^c αντίστοιχα. Οι διαδικασίες Π_1, Π_2 είναι ανεξάρτητες Poisson διαδικασίες.

Αν εφαρμόσουμε τον ίδιο συλλογισμό όπως και στην απόδειξη του λήμματος των αμοιβαία ξένων διαδικασιών στο τυχαίο σύνολο

$$f(\Pi_1) \times f(\Pi_2) \subseteq T \times T$$

τότε μπορούμε να δούμε ότι οι διαδικασίες $f(\Pi_1), f(\Pi_2)$ είναι με πιθανότητα 1, ξένες.

Ορίζουμε με M το μέτρο στο χώρο $S \times S$ με την έννοια ότι το $M(C)$ είναι ο αναμενόμενος αριθμός των ζευγαριών $(x, \psi) \in C$ για $C \subseteq S \times S$ και

$$x \in \Pi, \psi \in \Pi, f(x) = f(\psi).$$

Τότε αφού οι $f(\Pi_1), f(\Pi_2)$ είναι ξένες θα έχουμε $M(A \times A^c) = 0$ για όλα τα A .

Για κάθε $A, B \subseteq S$, το σύνολο $A \times B$ μπορεί να γραφεί ως $(A \cap B) \times (A \cap B)$ και το μέτρο του $A \times B$ ως $M(A \times B) = m(A \cap B)$, όπου $m(A) = M(A \times A)$. Αν $B = S$ τότε ισχύει ότι $m(A) = m(A \times S)$ και επομένως το m είναι μέτρο. Έστω M_1 το επαγόμενο μέτρο από το m στο διαγώνιο $D \subseteq S \times S$ μέσω της συνάρτησης $x \mapsto (x, x)$. Τότε:

$$M_1(A \times B) = m(A \cap B) \Rightarrow M(A \times B) \text{ για όλα τα } A, B.$$

Από τη μοναδικότητα των μέτρων επεται ότι $M_1 = M$. Επομένως με πιθανότητα 1 δεν υπάρχουν διαφορετικά $x \in \Pi, \psi \in \Pi$ με $f(x) = f(\psi)$.

Αν καταργήσουμε την υπόθεση ότι $\mu(S) < \infty$ και αντί αυτού υποθέσουμε ότι τα S_n είναι ξένα και $S = \bigcup_{n=1}^{\infty} S_n$, $\mu(S_n) < \infty$.

Έστω επίσης ο περιορισμός Π_n του Π έτσι ώστε οι Π_n είναι ανεξάρτητες διαδικασίες Poisson με πεπερασμένα μέτρα μ_n . Τότε τα $f(\Pi_n)$ είναι ανεξάρτητες διαδικασίες Poisson με μέτρα $\mu_n^*(B) = \mu(f^{-1}(B))$, και η ένωση τους

$$f(\Pi) = f\left(\bigcup_{n=1}^{\infty} \Pi_n\right) = \bigcup_{n=1}^{\infty} f(\Pi_n)$$

είναι διαδικασία Poisson με μέσο μέτρο

$$\mu^* = \mu^*(B) = \mu(f^{-1}(B))$$

⊠

Παράδειγμα

Έστω Π μία ομοιόμορφη διαδικασία με ρυθμό λ . Ο μετασχηματισμός των συντεταγμένων (x, y) στις πολικές συντεταγμένες του (x, y) θα δίνεται από τη συνάρτηση $f(x, y) = ((x^2 + y^2)^{1/2}, \tan^{-1}(x, y))$.

Το μέσο μέτρο θα είναι $\mu^*(B) = \int \int_{f^{-1}(B)} \lambda dx dy = \int \int_B \lambda r dr d\theta$.

Τα σημεία (r, θ) σχηματίζουν μία διαδικασία Poisson στο χώρο

$$\{(r, \theta); r > 0, 0 \leq \theta \leq 2\pi\}$$

με παράμετρο $\lambda^*(r, \theta) = \lambda r$

Αν αγνοήσουμε τις γωνίες θ το r ακολουθεί διαδικασία Poisson στους θετικούς πραγματικούς αριθμούς με συνάρτηση έντασης

$$\lambda^{**}(r) = \int_0^{2\pi} \lambda^*(r, \theta) d\theta = 2\pi \lambda r$$

⊠

Παράδειγμα

Έστω Π μία διαδικασία Poisson στον $\mathbb{R}^D$ με ρυθμό $\lambda(x_1, x_2, \dots, x_D)$

Έστω επίσης η συνάρτηση προβολής $f: \mathbb{R}^D \rightarrow \mathbb{R}^d$, $d \leq D$ η οποία ορίζεται ως

$$f(x_1, x_2, \dots, x_D) = (x_1, x_2, \dots, x_d)$$

Για $B \subseteq \mathbb{R}^d$, το μέσο μέτρο θα είναι

$$\begin{aligned} \mu^*(B) &= \int \cdots \int_{B \times \mathbb{R}^{D-d}} \lambda(x_1, x_2, \dots, x_D) dx_1 dx_2 \dots dx_D = \\ &= \int \cdots \int_B \lambda^*(x_1, x_2, \dots, x_d) dx_1 dx_2 \dots dx_d \\ \text{όπου } \lambda^*(x_1, x_2, \dots, x_d) &= \int \cdots \int_{\mathbb{R}^{D-d}} \lambda(x_1, x_2, \dots, x_D) dx_{d+1} dx_{d+2} \dots dx_D \quad (*) \end{aligned}$$

Καθώς το μ είναι σ-πεπεραμένο και το μ^* είναι μη-ατομικό το θεώρημα αντιστοίχισης μας δίνει ότι εάν το ολοκλήρωμα $(*)$ συγκλίνει τότε η διαδικασία $f(\Pi)$ είναι διαδικασία Poisson με ρυθμό λ^* στον $\mathbb{R}^d$.

□

2.2.4 Η Διωνυμική-Πολυνυμική διαδικασία

Έστω Π μία διαδικασία Poisson στο S με μέσο μέτρο μ και επίσης $\mu(S) < \infty$. Τότε το Π είναι σχεδόν βεβαίως ένα πεπερασμένο υποσύνολο του S και ο συνολικός αριθμός $N(S)$ έχει την κατανομή $\mathcal{P}(\mu(S))$

Αν δεσμεύσουμε ως προς την τιμή του $N(S)$ και αν συμβολίσουμε με $P_n\{\cdot\} = P\{\cdot | N(S) = n\}$ θα έχουμε τα εξής:

Έστω $A_1, A_2, \dots, A_k$ ξένα υποσύνολα του S .

$$\begin{aligned} \mathbb{P}_n\{N(A_1) = n_1, N(A_2) = n_2, \dots, N(A_k) = n_k\} &= \\ \mathbb{P}\{N(A_0) = n_0, N(A_1) = n_1, \dots, N(A_k) = n_k\} / \mathbb{P}\{N(S) = n\} &= \\ = \prod_{j=1}^k \frac{e^{-\mu(A_j)} \mu(A_j)^{n_j}}{n_j!} / \frac{e^{-\mu(S)} \mu(S)^n}{n!} &= \\ = \frac{n!}{n_0! n_1! \dots n_k!} \left(\frac{\mu(A_0)}{\mu(S)}\right)^{n_0} \left(\frac{\mu(A_1)}{\mu(S)}\right)^{n_1} \dots \left(\frac{\mu(A_k)}{\mu(S)}\right)^{n_k} \end{aligned}$$

όπου $n_0 = n - \sum_{i=1}^k n_i$ και A_0 το συμπλήρωμα του συνόλου $A_1 \cup A_2 \dots \cup A_k$.

Συνεπώς ένα πεπερασμένο σύνολο Π για το οποίο το πλήθος των σημείων του είναι ίσο με n τότε θα είναι:

$$\mathbb{P}\{N(A_1) = n_1, N(A_2) = n_2, \dots, N(A_k) = n_k\} = \frac{n!}{n_0! n_1! \dots n_k!} p(A_0)^{n_0} p(A_1)^{n_1} \dots p(A_k)^{n_k}$$

Η παραπάνω διαδικασία μπορεί να ονομασθεί πολυωνυμική με παραμέτρους n , $p(\cdot) = \frac{\mu(\cdot)}{\mu(S)}$.

Για $k = 2$ προκύπτει η Διωνυμική διαδικασία. Στην βιβλιογραφία η παραπάνω διαδικασία συναντάται και με το όνομα διαδικασία Bernoulli. (Βλέπε J.F.C Kingman (1993)).

Το μέσο μέτρο θα είναι $E\{N(A)\} = np(A)$ έτσι ώστε το p είναι μέτρο πιθανότητας στο S .

Συνεπώς δεσμεύοντας μία διαδικασία Poisson με πεπερασμένο μέσο μέτρο μ , ως προς $N(S) = n$ τότε η διαδικασία γίνεται Πολυωνυμική-Διωνυμική.

Ένας άλλος τρόπος κατασκευής Πολυωνυμική-Διωνυμικής διαδικασίας είναι να θεωρήσουμε n ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, \dots, X_n$ κατανεμημένες στο χώρο S σύμφωνα με το κάποιο μέτρο πιθανότητας p . Αν το p είναι μη ατομικό, οι τυχαίες μεταβλητές είναι διακριτές με πιθανότητα 1 (αποδεικνύεται με το λήμμα των αμοιβαία ξένων διαδικασιών) και έτσι το $\{X_1, X_2, \dots, X_n\}$ είναι ένα τυχαίο σύνολο με n στοιχεία.

Τότε η τυχαία μεταβλητή

$$N(A) = \#\{r : X_r \in A\}$$

ικανοποιεί τη σχέση

$$\mathbb{P}\{N(A_1) = n_1, N(A_2) = n_2, \dots, N(A_k) = n_k\} = \frac{n!}{n_0!n_1!\dots n_k!} p(A_0)^{n_0} p(A_1)^{n_1} \dots p(A_k)^{n_k}$$

Συνεπώς η διαδικασία $N(A)$ είναι Πολυωνυμική διαδικασία.

Γενικά θα έχουμε ότι δεδομένου του αριθμού $N(S)$ των σημείων του χώρου S , τα σημεία μιας πεπερασμένης διαδικασίας Poisson μπορούν να θεωρηθούν ως τυχαίες μεταβλητές, οι οποίες προφανώς θα είναι $N(S)$ το πλήθος, με κοινή κατανομή $p(\cdot) = \frac{\mu(\cdot)}{\mu(S)}$.

2.2.5 Ανεξάρτητες εκλεπτύνσεις της διαδικασίας Poisson

The independent (or Bernoulli) thinning

Η εκλέπτυνση μιας διαδικασίας Poisson είναι πολύ χρήσιμη στις εφαρμογές καθώς μπορούμε να μειώσουμε τον αριθμό των σημείων της διαδικασίας Poisson ώστε να έχουμε πιο βολική και πιο πρακτική συνάρτηση έντασης (λ).

Έστω Π μια διαδικασία Poisson στο χώρο S .

Για κάθε $x \in S$ θέτουμε $1 - \alpha(x)$, $0 \leq \alpha(x) \leq 1$ την πιθανότητα το σημείο x να αφαιρεθεί από το σύνολο των πραγματοποιήσεων της διαδικασίας Poisson. Η καινούργια αυτή διαδικασία είναι μια νέα διαδικασία Poisson με συνάρτηση έντασης $\lambda_\alpha(x) = \alpha(x)\lambda(x)$. Υποθέτουμε ότι ο συνολικός αριθμός n των σημείων της αρχικής διαδικασίας Π είναι άγνωστος. Ονομάζουμε τις εκλέπτυνσεις που προκύπτουν με αυτό τον τρόπο Bernoulli ή ανεξάρτητες εκλεπτύνσεις γιατί η πιθανότητα $\alpha(x)$ εξαρτάται μόνο από το εκάστοτε x .

Στην απλή περίπτωση όπου η πιθανότητα $\alpha(x) = c$ είναι δηλαδή σταθερή στο A τότε θα ισχύουν τα εξής:

$$\mu = \int_A \lambda(x)dx, \quad \mu_\alpha = \int_A \lambda_\alpha(x)dx, \quad \beta = \frac{\mu_\alpha}{\mu}$$

Η καινούργια διαδικασία Π' έχει m σημεία δεδομένου ότι η αρχική διαδικασία είχε n σημεία, με πιθανότητα

$$\mathbb{P}(m|n) = \binom{n}{m} \beta^m (1 - \beta)^{n-m}, \quad m \leq n$$

Καθώς ο αριθμός n είναι άγνωστος θα έχουμε ότι η πιθανότητα τα σημεία της καινούργιας διαδικασίας να είναι m :

$$\begin{aligned} \mathbb{P}(m) &= \sum_{n=m}^{\infty} \binom{n}{m} \beta^m (1 - \beta)^{n-m} \mathbb{P}(n) = \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} \beta^m (1 - \beta)^{n-m} \frac{e^{-\mu} \mu^n}{n!} \\ &= \frac{(\beta\mu)^m}{m!} e^{-\mu} \sum_{n=m}^{\infty} \frac{((1 - \beta)\mu)^{n-m}}{(n-m)!} = \frac{(\beta\mu)^m}{m!} e^{-\beta\mu} \equiv \frac{e^{-\mu_\alpha} \mu_\alpha^m}{m!} \end{aligned}$$

Έχουμε δηλαδή δείξει ότι ο συνολικός αριθμός των σημείων της διαδικασίας $\Pi'(m)$ ακολουθεί την κατανομή Poisson με μέσο μ_α

Αν συμβολίσουμε με $\{x'_1, \dots, x'_m\}$ τα σημεία που κρατήσαμε μετά από τις δοκιμές Bernoulli, τα σημεία αυτά είναι ανεξάρτητα με συνάρτηση πιθανότητας $\lambda_\alpha(x)/\mu_\alpha = \lambda(x)/\mu(x)$

Στην πιο γενική περίπτωση που η πιθανότητα $\alpha(x)$ δεν είναι σταθερή αντιμετωπίζουμε το πρόβλημα με παρόμοιο τρόπο.

Χωρίζουμε το A σε πολλά μικρότερα σύνολα A_t μη επικαλυπτόμενα.

Τότε ορίζουμε:

$$\mu = \int_{A_t} \lambda(x) dx, \quad \mu_\alpha = \int_{A_t} \lambda_\alpha(x) dx, \quad \beta = \frac{\mu_\alpha}{\mu}$$

Η πιθανότητα η διαδικασία Π' να έχει m σημεία μετά από την εκλέπτυνση έχει όπως δείξαμε στην απλή περίπτωση κατανομή $\mathcal{P}(\mu_\alpha)$. Πάλι τα σημεία $\{x'_1, \dots, x'_m\}$ είναι ανεξάρτητα με κατανομή στο R' , $\lambda_\alpha(x)/\mu_\alpha$.

Επεκτείνουμε τη συνάρτηση έντασης λ_α από το R' σε όλο το R θέτοντας την με μηδέν εκτός των A_t . Με αυτό τον τρόπο η συνάρτηση έντασης $\lambda_\alpha(x)$ ορίζεται σε όλο το R .

2.2.6 Το θεώρημα Ύπαρξης

Θεώρημα 2.2.4. Ύπαρξης

The existence theorem

Έστω μ ένα μη-ατομικό μέτρο στο χώρο καταστάσεων S που μπορεί να εκφραστεί στη μορφή

$$\mu = \sum_{n=1}^{\infty} \mu_n, \quad \mu_n(S) < \infty$$

Τότε υπάρχει διαδικασία Poisson στο χώρο καταστάσεων S με μέσο μέτρο μ .

Απόδειξη

Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $\mu_n(S) > 0$ για όλα τα n . Κατασκευάζουμε σε κατάλληλο χώρο πιθανότητας ανεξάρτητες μεταβλητές

$$N_n, X_{nr}, \quad (n, r = 1, 2, 3 \dots)$$

τέτοιες ώστε η κατανομή των μεταβλητών N_n να είναι $\mathcal{P}(\mu_n(S))$ και των μεταβλητών X_{nr} να είναι $p_n(\cdot) = \frac{\mu_n(\cdot)}{\mu_n(S)}$.

Έστω $\Pi_n = \{X_{n_1}, X_{n_2}, \dots, X_{n_{N_n}}\}$ και $\Pi = \bigcup_{n=1}^{\infty} \Pi_n$

Θεωρούμε τις τυχαίες μεταβλητές $N_n(A) = \#\{\Pi_n \cap A\}$ και αν $A_1, A_2, \dots, A_k$ ξένα σύνολα και $A_0 = \left(\sum_{i=1}^k A_i\right)^c$ το συμπλήρωμα της ενωσεως των συνόλων. Έχουμε:

$$\begin{aligned} & \mathbb{P}\{N_n(A_1) = m_1, N_n(A_2) = m_2, \dots, N_n(A_k) = m_k | N_n = m\} \\ &= \frac{m!}{m_0!m_1! \dots m_k!} p_n(A_0)^{m_0} p_n(A_1)^{m_1} \dots p_n(A_k)^{m_k} \end{aligned}$$

όπου $m_0 = m - m_1 - \dots - m_k$.

Άρα παίρνουμε ότι

$$\begin{aligned} & \mathbb{P}\{N_n(A_1) = m_1, N_n(A_2) = m_2, \dots, N_n(A_k) = m_k\} \\ &= \sum_{m=\sum m_j}^{\infty} \frac{e^{-\mu_n(S)} \mu_n(A)^m}{m!} \frac{m!}{m_0!m_1! \dots m_k!} \prod_{j=0}^k p_n(A_j)^{m_j} \\ &= \sum_{m=\sum m_j}^{\infty} \pi_{m-\sum m_j}(\mu_n(A_0)) \prod_{j=0}^k \pi_{m_j}(\mu_n(A_j)) \\ &= \prod_{j=0}^k \pi_{m_j}(\mu_n(A_j)) \end{aligned}$$

Επομένως τα $N_n(A_j)$ είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομές $\mathcal{P}(\mu_n(A_j))$, έτσι ώστε τα Π_n είναι ανεξάρτητες διαδικασίες Poisson με μέσα μέτρα μ_n . Από το θεώρημα της υπέρθεσης η Π ορίζει διαδικασία Poisson με μέσο μέτρο μ .

□

2.3 Αθροίσματα Poisson

Σε αυτή την ενότητα εξετάζουμε τα αθροίσματα της μορφής $\Sigma = \sum_{X \in \Pi} f(X)$ όπου f πραγματική συνάρτηση στο χώρο καταστάσεων της διαδικασίας Poisson, Π .

Έστω για παράδειγμα $A_j = \{x; f(x) = f_j\}$ μετρήσιμο με μέσο $m_j = \mu(A_j) < \infty$, και τα A_j είναι ξένα μεταξύ τους. Τότε τα $N_j = N(A_j)$ είναι ανεξάρτητα με κατανομές $\mathcal{P}(m_j)$ αντίστοιχα. Για το άθροισμα Σ θα ισχύει ότι:

$$\Sigma = \sum_{X \in \Pi} f(X) = \sum_{j=1}^k f_j N_j$$

Η χαρακτηριστική συνάρτηση του Σ θα είναι:

$$\begin{aligned} \mathbb{E}(e^{\theta \Sigma}) &= \prod_{j=1}^k \mathbb{E}(e^{\theta f_j N_j}) \\ &= \prod_{j=1}^k \exp \{ (e^{\theta f_j} - 1) m_j \} \\ &= \exp \left\{ \sum_{j=1}^k \int_{A_j} (e^{\theta f(x)} - 1) \mu(dx) \right\} \\ &= \exp \left\{ \int_S (e^{\theta f(x)} - 1) \mu(dx) \right\} \end{aligned}$$

Αν η συνάρτηση f μπορεί να πάρει και αρνητικές και θετικές τιμές, είναι καλύτερα να παίρνουμε τον θ ως καθαρά μιγαδικό αριθμό και τότε η χαρακτηριστική συνάρτηση γίνεται:

$$\mathbb{E}(e^{it\Sigma}) = \exp \left\{ \int_S (e^{itf(x)} - 1) \mu(dx) \right\}$$

Αν η συνάρτηση f παίρνει μόνο θετικές τιμές τότε συνήθως παίρνουμε το θ πραγματικό και αρνητικό και τότε:

$$\mathbb{E}(e^{-u\Sigma}) = \exp \left\{ - \int_S (1 - e^{-uf(x)}) \mu(dx) \right\} \text{ για όλα τα } u \geq 0$$

Η χαρακτηριστική συνάρτηση είναι πολύ χρήσιμη καθώς μέσω αυτής της σχέσης μπορούμε να υπολογίσουμε τη μέση τιμή και διασπορά

$$\begin{aligned} \mathbb{E}(\Sigma) &= \int_S f(x) \mu(dx) \\ Var(\Sigma) &= \int_S (f(x))^2 \mu(dx) \end{aligned}$$

της τυχαίας μεταβλητής.

Έχουμε ότι αν οι f_1, f_2 είναι πραγματικές συναρτήσεις ισχύει ότι

$$\mathbb{E}(e^{\theta_1 \Sigma_1 + \theta_2 \Sigma_2}) = \exp \left\{ \int_S (e^{\theta_1 f_1(x) + \theta_2 f_2(x)} - 1) \mu(dx) \right\}$$

Η πιο πάνω εξίσωση καθορίζει την από κοινού κατανομή των

$$\Sigma_1 = \sum_{X \in \Pi} f_1(X), \Sigma_2 = \sum_{X \in \Pi} f_2(X)$$

Επίσης,

$$\mathbb{E}(\Sigma_1 \Sigma_2) = \int f_1 d\mu \int f_2 d\mu + \int f_1 f_2 d\mu$$

ο δεύτερος όρος της ισότητας δίνει τη συνδιακύμανση του Σ_1, Σ_2

$$\begin{aligned} \mathbb{E} \left(\sum_{X_1, X_2 \in \Pi} f_1(X_1) f_2(X_2) \right) &= \mathbb{E} \left(\sum_{X_1 \in \Pi} f_1(X_1) f_2(X_2) \right) + \mathbb{E} \left(\sum_{X_1 \neq X_2} f_1(X_1) f_2(X_2) \right) \\ \mathbb{E} \left(\sum_{\substack{X_1, X_2 \in \Pi \\ X_1 \neq X_2}} f_1(X_1) f_2(X_2) \right) &= \mathbb{E} \left(\sum_{X_1 \in \Pi} f_1(X_1) \right) \mathbb{E} \left(\sum_{X_2 \in \Pi} f_2(X_2) \right) \end{aligned}$$

Γενικότερα:

$$\mathbb{E} \left(\sum_{\substack{X_1, X_2, \dots, X_n \in \Pi \\ X_i \text{ διακριτά}}} f_1(X_1) f_2(X_2) \dots f_n(X_n) \right) = \prod_{j=1}^n \mathbb{E} \left(\sum_{X_j \in \Pi} f_j(X_j) \right)$$

Θεώρημα 2.3.1. Το θεώρημα του Campbell

Cambell's theorem

Έστω Π μία διαδικασία Poisson στο χώρο καταστάσεων S με μέσο μέτρο μ και έστω $f : S \rightarrow \mathbb{R}$ μετρήσιμη συνάρτηση. Τότε το άθροισμα

$$\Sigma = \sum_{X \in \Pi} f(X)$$

συγκλίνει απολύτως κατά πιθανότητα αν και μόνον αν

$$\int_S \min(|f(x)|, 1) \mu(dx) < \infty$$

Αν πιο πάνω συνθήκη ισχύει τότε θα ισχύει ότι:

$$\mathbb{E}(e^{\theta \Sigma}) = \exp \left\{ \int_S (e^{\theta f(x)} - 1) \mu(dx) \right\}$$

για κάθε θ για το οποίο το ολοκλήρωμα στο δεξί μέρος της ισότητας συγκλίνει. Στην ειδική περίπτωση που το θ είναι καθαρά μιγαδικός αριθμός η ισότητα ισχύει πάντα.

Επίσης αν το ολοκλήρωμα $\int_S f(x)\mu(dx)$ συγκλίνει τότε θα ισχύει και για τη μέση τιμή ότι

$$\mathbb{E}(\Sigma) = \int_S f(x)\mu(dx)$$

και για τη διασπορά

$$Var(\Sigma) = \int_S f(x)^2\mu(dx)$$

και είναι άπειρη ή πεπερασμένη.

Η απόδειξη παραλείπεται καθώς ξεφεύγει από τους σκοπούς της μελέτης μας.³

Πόρισμα 2.3.1. Πόρισμα

Αν οι συναρτήσεις $f_1, f_2, \dots, f_n$ ικανοποιούν τη συνθήκη

$$\int_S \min(|f(x)|, 1)\mu(dx) < \infty$$

έτσι ώστε αν τα $\Sigma_j = \sum_{X \in \Pi} f_j(X)$ συγκλίνουν με πιθανότητα 1 τότε

$$E(e^{it_1 \Sigma_1 + it_2 \Sigma_2 + \dots + it_n \Sigma_n}) = \exp \left\{ \int_S (e^{it_1 f_1(x) + \dots + it_n f_n(x)} - 1) \mu(dx) \right\}$$

Αν η f_j ικανοποιεί την ανισότητα:

$$\int_S f_j(x)^2 \mu(dx) < \infty$$

τότε

$$cov(\Sigma_j, \Sigma_k) = \int_S f_j(x)f_k(x)\mu(dx)$$

³Ο αναγνώστης μπορεί να βρει την απόδειξη στο βιβλίο J.F.C Kingman (1993)

Στην ειδική περίπτωση όπου $\theta = -1$ η εξίσωση

$$\mathbb{E}(e^{\theta \Sigma}) = \exp \left\{ \int_S (e^{\theta f(x)} - 1) \right\}$$

γίνεται

$$\mathbb{E}(e^{-\Sigma_f}) = \exp \left\{ - \int_S (1 - e^{-f(x)}) \mu(dx) \right\}$$

$$\text{όπου } \Sigma_f = \sum_{X \in \Pi} f(X)$$

Ορίζουμε τη μέση τιμή $\mathbb{E}(e^{-\Sigma_f})$ ως τη **χαρακτηριστική συναρτησοειδή**.

Θεώρημα 2.3.2. Θεώρημα: Το θεώρημα του Renyi

Renyi's theorem

Έστω μ μη-ατομικό μέτρο στον ευκλείδειο χώρο $\mathbb{R}^d$ πεπερασμένο σε φραγμένα σύνολα. Έστω επίσης Π ένα τυχαίο αριθμήσιμο υποσύνολο του $\mathbb{R}^d$ τέτοιο ώστε, αν το A είναι άπειρη ένωση παραλληλογράμμων θα ισχύει ότι

$$\mathbb{P}\{\Pi \cup A = \emptyset\} = e^{-\mu(A)}$$

Τότε η Π είναι διαδικασία Poisson με μέσο μέτρο μ .

Η απόδειξη παραλείπεται.⁴

2.4 Η διαδικασία Poisson στην ευθεία των πραγματικών αριθμών

Έστω ότι Π είναι διαδικασία Poisson με χώρο καταστάσεων τους πραγματικούς αριθμούς και μέσο μέτρο $\lambda \times \text{Lebesgue}$ μέτρο. Η απαριθμήτρια διαδικασία $N(A)$, των σημείων του Π σε ένα Borel σύνολο A έχει την κατανομή

⁴Ο αναγνώστης παραπέμπεται στο βιβλίο J.F.C Kingman (1993)

$\mathcal{P}(\lambda(|A|))$. Συγκεκριμένα σε ένα σύνολο (a, b) η τυχαία μεταβλητή $N(a, b)$ έχει μέση τιμή $\lambda(b - a)$. Όπως έχουμε αναφέρει και προηγουμένως η Π είναι διαδικασία Poisson με παράμετρο λ .

Σε φραγμένο σύνολο, το σύνολο των σημείων του Π είναι πεπερασμένο με πιθανότητα 1 και δεν έχει πεπερασμένα σημεία συσσώρευσης.

Στο σύνολο $(0, \infty)$ υπάρχουν άπειρα σημεία τα οποία μπορούν να διαταχθούν και να γραφτούν ως:

$$0 < X_1 < X_2 < X_3 < \dots$$

Όμοια στο σύνολο $(-\infty, 0)$ υπάρχουν άπειρα σημεία τα οποία μπορούν να διαταχθούν και να γραφτούν ως:

$$\dots < X_{-3} < X_{-2} < X_{-1} < 0$$

Τα σημεία αυτά καλύπτουν όλο τα σημεία της διαδικασίας Π καθώς η πιθανότητα $\mathbb{P}\{0 \in \Pi\} = 0$

Ορίζουμε λοιπόν την ακόλουθη διαδικασία:

Έστω $\{X_n, n \geq 1\}$ ακολουθία τυχαίων μεταβλητών. Οι τυχαίες μεταβλητές αυτές εκφράζουν τον χρόνο πραγματοποίησης κάποιου γεγονότος (για παράδειγμα σε μία ουρά αναμονής μπορεί να εκφράζουν την άφιξη κάποιου πελάτη). Η αпарιθμήτρια διαδικασία N για την οποία ισχύει ότι $N(0, x] \geq n$ είναι ισοδύναμη με το $X_n \leq x$ για κάθε $x > 0$.

Συνεπώς η X_n είναι τυχαία μεταβλητή αφού θα ισχύει ότι

$$\{\omega; X_n(\omega) \leq x\} = \{\omega; N(0, x] \neq n\} \in \mathcal{F}$$

Όμοια και η X_{-n} είναι τυχαία μεταβλητή και επομένως θα έχουμε ότι το σύνολο

$$\Pi = \{\dots, X_{-2}, X_{-1}, X_1, X_2, \dots\}$$

είναι τυχαίο σύνολο.

Τα σημεία των $\Pi_n(-\infty, 0)$ και $\Pi \cap (0, \infty)$ ορίζουν ανεξάρτητες διαδικασίες Poisson στους χώρους $(-\infty, 0)$, $(0, \infty)$ αντίστοιχα και οι υπακολουθίες $(X_n; n \leq -1)$, $(X_n; n \geq 1)$ είναι ανεξάρτητες.

Εφαρμόζοντας το θεώρημα αντιστοίχισης με $f(x) = -x$ θα έχουμε ότι η από κοινού κατανομή των ακολουθιών

$$(-X_{-n}; n = 0, 1, 2, \dots)$$

είναι όμοια με την κατανομή των

$$(X_n; n = 0, 1, 2, 3, \dots)$$

Έτσι η $\Pi = \{\dots, X_{-2}, X_{-1}, X_1, X_2, \dots\}$ μπορεί να κατασκευαστεί από δύο σύνολα διατεταγμένων τυχαίων μεταβλητών $0 < X_1 < X_2 < X_3 < \dots$ εκ των οποίων το ένα είναι αντεστραμμένο και παίρνει τιμές στο σύνολο $(-\infty, 0)$.

Στις εφαρμογές πολύ συχνά χρησιμοποιούνται οι παρακάτω τυχαίες μεταβλητές, γύρω από τις οποίες έχει αναπτυχθεί η θεωρία των διαδικασιών Poisson .

Ορίζουμε:

- Τις τυχαίες μεταβλητές που εκφράζουν τους χρόνους πραγματοποίησης γεγονότων $\{X_n, n \geq 1\}$.
- Τις τυχαίες μεταβλητές $Y_1 = X_1, Y_n = X_n - X_{n-1}, n \geq 2$ που εκφράζουν τους ενδιάμεσους χρόνους πραγματοποίησεων των γεγονότων.
- Την απεριθμήτρια στοχαστική διαδικασία των X_n ως $N(x) = \sup\{n \geq 0; X_n \leq x\}, x \geq 0$. Δηλαδή η $N(x)$ είναι ο αριθμός των γεγονότων που συνέβησαν στο διάστημα $(0, x]$.

Θεώρημα 2.4.1. Το θεώρημα των διαστημάτων

Interval Theorem

Έστω Π μια διαδικασία Poisson με παράμετρο $\lambda \in (0, \infty)$ και έστω ότι στα σημεία της Π μπορούν να διαταχθούν ως $0 < X_1 < X_2 < X_3 < \dots$. Τότε οι τυχαίες μεταβλητές $Y_1 = X_1, Y_n = X_n - X_{n-1}, n \geq 2$ είναι ανεξάρτητες και ισόνομες με συνάρτηση πυκνότητας πιθανότητας

$$g(\psi) = \lambda e^{-\lambda\psi}, \psi > 0$$

έχουν δηλαδή την εκθετική κατανομή.

Απόδειξη

Θεωρούμε το υποσύνολο των θετικών πραγματικών αριθμών

$$\Pi' = \{X_2 - X_1, X_3 - X_1, X_4 - X_1, \dots\}$$

Θέλουμε να δείξουμε ότι τα Π' , X_1 είναι ανεξάρτητα και ότι η Π' είναι επίσης διαδικασία Poisson με παράμετρο λ .

Έστω τώρα f μια συνεχής συνάρτηση στο $(0, \infty)$ για την οποία θα ισχύει ότι $f(0) \equiv 0$. Θεωρούμε το άθροισμα

$$\Sigma' = \sum_{n=2}^{\infty} f(X_n - X_1)$$

Αν αποδείξουμε ότι για κάθε τυχαία επιλογή της f οι X_1 , Σ' είναι ανεξάρτητες ότι η Σ' έχει την ίδια κατανομή με τη $\Sigma = \sum_{n=1}^{\infty} f(X_n)$, τότε έπεται το ζητούμενο.

Ορίζουμε με ξ_k τον ελάχιστο ακέραιο αριθμό πολλαπλάσιο του 2^{-k} που είναι μεγαλύτερος από τον X_1 . Το ξ_k είναι τυχαία μεταβλητή που είναι φθίνει καθώς το $k \rightarrow \infty$ έτσι ώστε:

$$\Sigma' = \lim_{k \rightarrow \infty} \Sigma^k, \Sigma^k = \sum_{n=2}^{\infty} f(X_n - \xi_k)$$

με την προϋπόθεση ότι $f(x) = 0$ αν $x < 0$.

Για τυχαία z, x θα έχουμε ότι:

$$P\{\Sigma^k \leq z, X_1 \leq x\} = \sum_{l=1}^{\infty} P\{\Sigma^k \leq z, X_1 \leq x, \xi_k = l2^{-k}\} \quad (*)$$

Όταν $\xi_k = l2^{-k}$ τα σημεία $X_n \in (l2^{-k}, \infty)$ σχηματίζουν μία διαδικασία Poisson σε αυτό το διάστημα και είναι ανεξάρτητα από την $X_1 < l2^{-k}$.

Επομένως τα

$$\Sigma^k = \sum_{n=2}^{\infty} f(X_n - l2^{-k})$$

έχουν την ίδια κατανομή με το Σ και θα ισχύει ότι

$$\mathbb{P}\{\Sigma^k \leq z, X_1 \leq x, \xi_k = l2^{-k}\} = \mathbb{P}\{\Sigma \leq z\} \mathbb{P}\{X_1 \leq x, \xi_k = l2^{-k}\}$$

αντικαθιστούμε στην (*) και παίρνουμε τα εξής:

$$\begin{aligned}\mathbb{P}\{\Sigma^k \leq z, X_1 \leq x\} &= \mathbb{P}\{\Sigma \leq z\} \mathbb{P}\{X_1 \leq x\} \\ &= \mathbb{P}\{\Sigma \leq z\} \mathbb{P}\{N(0, x] \neq 1\} \\ &= \mathbb{P}\{\Sigma \leq z\} (1 - e^{-\lambda x})\end{aligned}$$

Καθώς $k \rightarrow \infty$ παίρνουμε ότι

$$P\{\Sigma' \leq z, X_1 \leq x\} = P\{\Sigma \leq z\} \int_0^x \lambda e^{-\lambda \psi}$$

Έτσι δείξαμε ότι το Σ' είναι ανεξάρτητο του X_1 έχει την ίδια κατανομή με το Σ και το $X_1 = Y_1$ έχει την εκθετική κατανομή.

Επαναλαμβάνοντας την ίδια διαδικασία μπορούμε να δείξουμε με επαγωγή ότι για κάθε m θα ισχύει ότι τα $Y_1, Y_2, \dots, Y_m, \Pi^m$ είναι ανεξάρτητα και ότι οι τυχαίες μεταβλητές Y_j ακολουθούν την εκθετική κατανομή.

Σημειώνεται ότι με Π^m συμβολίζουμε το σύνολο $\Pi^m = \{X_{m+1} - X_m, X_{m+2} - X_m, \dots\}$ $\square$

Στην απόδειξη είναι απαραίτητο να υποθέσουμε διάταξη για τις τυχαίες μεταβλητές X_i στο σύνολο των πραγματικών αριθμών.

Για παράδειγμα δεν ισχύει ότι τα σημεία της διαδικασίας $X_n - X_1, n \neq 1$ σχηματίζουν Poisson διαδικασία.

Αν ισχύει το πιο πάνω τότε η $L = X_1 - X_{-1}$ θα είχε πυκνότητα πιθανότητας την $g(\psi) = \lambda e^{-\lambda \psi}, \psi > 0$ ενώ στην πραγματικότητα η L έχει την ίδια κατανομή με την X_2 και θα ισχύει ότι $g_2(\psi) = \int_0^\psi g(u)g(\psi - u)du = \lambda^2 \psi e^{-\lambda \psi}$

Το πιο πάνω αντιπαράδειγμα αναφέρεται συχνά στη βιβλιογραφία ως το ανανεωτικό παράδοξο. Για οποιοδήποτε x η απόσταση από το x στο επόμενο σημείο του Π (στα δεξιά) και η απόσταση από το x στο προηγούμενο σημείο του Π (στα αριστερά) είναι τυχαίες μεταβλητές ανεξάρτητες και ισόνομες με την εκθετική κατανομή. Το άθροισμα τους όμως έχει την κατανομή $g_2(\psi) = \int_0^\psi g(u)g(\psi - u)du = \lambda^2 \psi e^{-\lambda \psi}$. Αυτό συμβαίνει γιατί τα μεγαλύτερα διαστήματα επιλέγονται με μεγαλύτερη πιθανότητα από τα πιο μικρά διαστήματα και αυτό το δειγματοληπτικό σφάλμα εκφράζεται μέσω

της $g_2(\psi) = \lambda \psi g(\psi)^5$.

2.4.1 Ο νόμος των μεγάλων αριθμών

Από το προηγούμενο θεώρημα είναι φανερό ότι το n -οστό σημείο X_n της ομογενούς διαδικασίας Poisson στο $(0, \infty)$ μπορεί να γραφτεί ως άθροισμα $X_n = Y_1 + Y_2 + \dots + Y_n$ ανεξάρτητων τυχαίων εκθετικών κατανομών.

Με αυτό τον τρόπο μπορούμε να βρούμε εύκολα την κατανομή της X_n καθώς η θεωρία για το άθροισμα ανεξάρτητων τυχαίων μεταβλητών μπορεί να εφαρμοστεί και σε αυτή την περίπτωση.

Έχουμε λοιπόν

$$P\{X_n \leq x\} = P\{N(0, x] \geq n\} = \sum_{r=n}^{\infty} \pi_r(\lambda x) = \int_0^x \lambda \pi_{n-1}(\lambda u) du$$

Μπορεί να αποδειχθεί με επαγωγή στο n ότι η X_n έχει συνάρτηση πυκνότητας πιθανότητας

$$g_n(x) = \lambda \pi_{n-1}(\lambda x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}$$

Επίσης είναι εύκολο πλέον να δείχθεί ότι $E(X_n) = n\lambda^{-1}$, $Var(X_n) = n\lambda^{-2}$. Από το νόμο των μεγάλων αριθμών, με πιθανότητα 1 θα ισχύει ότι

$$\lim_{n \rightarrow \infty} X_n/n = \lambda^{-1}$$

Αφού $N(0, t] \rightarrow \infty$ αν $t \rightarrow \infty \Rightarrow \lim_{t \rightarrow \infty} X_{N(0, t]}/n(0, t] = \lambda^{-1}$

Καθώς ισχύει ότι $X_{N(0, t]} \leq t \leq X_{N(0, t]+1} \Rightarrow \lim_{t \rightarrow \infty} t/N(0, t] = \lambda^{-1}$

Καταλήγουμε λοιπόν στο ακόλουθο θεώρημα:

⁵Για περισσότερες πληροφορίες, παραδείγματα και εφαρμογές της διαδικασίας Poisson στην ευθεία των πραγματικών αριθμών προτείνουμε στον αναγνώστη το βιβλίο Vidyadhar G. Kulkarni (1995)

Θεώρημα 2.4.2. Ο νόμος των μεγάλων αριθμών

The law of large numbers

Έστω Π διαδικασία Poisson με σταθερή παράμετρο λ στο $(0, \infty)$. Ο αριθμός $N(0, t]$ των σημείων της Π στο $(0, t]$ ικανοποιεί τη σχέση

$$\lim_{t \rightarrow \infty} N(0, t]/t = \lambda$$

με πιθανότητα 1.

Απόδειξη

Όπως έχουμε δει η τυχαία μεταβλητή $N(0, t]$ έχει κατανομή $\mathcal{P}(\lambda t)$ και με

$$E\{N(0, t]\} = \lambda t, \quad \text{Var}\{N(0, t]\} = \lambda t$$

Από την ανισότητα του Chebychev έχουμε για κάθε ϵ θετικό

$$P\left\{\left|\frac{N(0, t]}{t} - \lambda\right| \geq \epsilon\right\} \leq \frac{\lambda}{\epsilon^2 t} \quad (*)$$

Θέτουμε $t_k = k^2$ και τότε η $(*)$ γίνεται

$$\sum_{k=1}^{\infty} P\left\{\left|\frac{N(0, k^2]}{k^2} - \lambda\right| \geq \epsilon\right\} < \infty$$

Από το Borell-Cantelli λήμμα με πιθανότητα 1 θα ισχύει ότι

$$\left|\frac{N(0, k^2]}{k^2} - \lambda\right| \geq \epsilon$$

Για τουλάχιστον ένα πεπερασμένο αριθμό ακέραιων αριθμών k . Αφού το ϵ ήταν αυθαίρετο τότε θα ισχύει ότι:

$$\lim_{k \rightarrow \infty} \frac{N(0, k^2]}{k^2} = \lambda$$

με πιθανότητα 1.

Για να ολοκληρωθεί η απόδειξη παίρνουμε το k να είναι το ακέραιο μέρος του $t^{1/2}$ έτσι ώστε για $t > 1$

$$N(0, k^2] \leq N(0, t] \leq N(0, (k+1)^2] \text{ και } k^2 \leq t \leq (k+1)^2$$

Η απόδειξη είναι πλήρης αφού αρκεί να παρατηρήσουμε ότι $(k+1)^2/k^2 \rightarrow 1$
 $\square$

2.4.2 Οι μη ομογενείς διαδικασίες Poisson στην ευθεία των πραγματικών αριθμών

Στις προηγούμενες παραγράφους του κεφαλαίου αυτού αυτή ασχοληθήκαμε με τις ομογενείς διαδικασίες Poisson. Τι γίνεται όμως όταν έχουμε μη ομογενή διαδικασία στην ευθεία των πραγματικών αριθμών ;

Σε μία διάσταση η μη ομογενής διαδικασία Poisson μπορεί να μετατραπεί σε ομογενή διαδικασία μέσω κάποιου μονότονου μετασχηματισμού.

Συγκεκριμένα θεωρούμε μία διαδικασία Poisson, Π , στο $\mathbb{R}$ με μέσο μέτρο μ σε το οποίο είναι πεπερασμένο φραγμένα διαστήματα. Τότε η συνάρτηση $M(t)$ όπου η $M(t)$ είναι συνεχής αύξουσα συνάρτηση και δίνεται από

$$\begin{aligned} M(t) &= \mu(0, t] = E\{N(0, t]\} , \quad (t \geq 0) \\ M(t) &= -\mu(t, 0] = -E\{N(t, 0]\} , \quad (t < 0) \\ \mu(a, b] &= M(b) - M(a) , \quad a < b \end{aligned}$$

Η συνάρτηση M ορίζει μονοσήμαντα το μ .
Το τυχαίο σύνολο

$$\Pi_1 = M(\Pi) = \{M(X); X \in \Pi\}$$

είναι Poisson διαδικασία με παράμετρο 1 στο πεπερασμένο ή άπειρο σύνολο $(M(-\infty), M(\infty))$ όπου με $M(\pm\infty) = \lim_{t \rightarrow \pm\infty} M(t)$

Όλα τα θεωρήματα της παραγράφου αυτής ισχύουν για τη διαδικασία Π_1 αφού από την Π_1 μπορούμε να καταλήξουμε στην Π αντιστρέφοντας την M . Συνεπώς οι ιδιότητες της Π μπορούν σε πολλές περιπτώσεις να ισχύουν και για την Π_1 αρκεί ο μετασχηματισμός M^{-1} να είναι γραμμικός. Μη γραμμικοί μετασχηματισμοί καταστρέφουν πολλές από τις ιδιότητες των διαδικασιών Poisson.

2.5 Μαρκαρισμένες διαδικασίες Poisson

2.5.1 Το θεώρημα Χρωματισμού

Έστω Π μία διαδικασία Poisson με χώρο καταστάσεων S και μέσο μέτρο μ . Ας υποθέσουμε ότι χρωματίζουμε τυχαία τα σημεία της Π με κόκκινο ή πράσινο χρώμα με πιθανότητες p , $q = 1 - p$ αντίστοιχα. Υποθέτουμε επίσης ότι τα χρώματα των σημείων είναι ανεξάρτητα μεταξύ τους.

Για κάθε $A \subseteq S$ συμβολίζουμε τα σημεία της Π στο A με $N(A)$ με $N_r(A)$ τα κόκκινα σημεία του Π και με $N_g(A)$ τα πράσινα σημεία. Τότε η τυχαία μεταβλητή $N(A)$ ακολουθεί την κατανομή $\mathcal{P}(\mu(A))$ και δεδομένου του $N(A)$ η δεσμευμένη κατανομή του $N_r(A)$ είναι διωνυμική με παραμέτρους $N(A)$, p . Συνεπώς είναι εύκολο να δούμε ότι οι τυχαίες μεταβλητές $N_r(A)$, $N(A) - N_r(A) = N_g(A)$ είναι ανεξάρτητες με κατανομές $\mathcal{P}(p\mu(A))$, $\mathcal{P}(q\mu(A))$ αντίστοιχα.

Επιπλέον αν έχουμε $A_1, \dots, A_n$ ξένα μεταξύ τους υποσύνολα του S , τότε οι τριάδες $(N(A_j), N_r(A_j), N_g(A_j))$, $j = 1, 2, \dots, n$ είναι ανεξάρτητες. Έτσι καταλήγουμε ότι οι $2n$ τυχαίες μεταβλητές $N_r(A_j)$, $N_g(A_j)$, $j = 1, 2, \dots, n$ είναι ανεξάρτητες.

Έπεται λοιπόν ότι το σύνολο των κόκκινων σημείων και το σύνολο των πράσινων σημείων είναι ανεξάρτητες διαδικασίες Poisson με μέσους $p\mu$, $q\mu$.

Θεώρημα 2.5.1. Το θεώρημα του Χρωματισμού

The colouring theorem

Έστω Π μία διαδικασία Poisson στο χώρο S με μέσο μέτρο μ . Έστω ότι τα σημεία της Π χρωματίζονται τυχαία με k χρώματα. Η πιθανότητα ότι ένα σημείο χρωματίζεται με το i -οστό χρώμα p_i και τα χρώματα για διαφορετικά σημεία είναι ανεξάρτητα. Έστω Π_i το σύνολο των σημείων του i -οστού χρώματος. Τότε οι Π_i είναι ανεξάρτητες διαδικασίες Poisson με μέσο μέτρο $\mu_i = p_i\mu$.

2.5.2 Η διαδικασία Poisson στο χώρο γινόμενο

Έστω Π μία διαδικασία Poisson στο χώρο S με μέσο μέτρο μ . Υποθέτουμε ότι κάθε σημείο X του τυχαίου συνόλου Π συνδέεται με μια τυχαία μεταβλητή (σημαδι -mark) m_X παίρνοντας τιμές σε κάποιο χώρο M . Οι κατανομές του m_X για διαφορετικά σημεία X είναι ανεξάρτητες.

Τότε μπορούμε να θεωρήσουμε το ζεύγος (X, m_X) ως ένα σημείο X^* του χώρου γινόμενου $S \times M$. Το σύνολο των σημείων X^* συγκροτούν ένα τυχαίο αριθμήσιμο υποσύνολο $\Pi^* = \{(X, m_X); X \in \Pi\}$ του χώρου (X, m_X) .

Τότε η Π^* είναι διαδικασία Poisson στο χώρο γινόμενο $S \times M$.

Ακολουθεί αυστηρή διατύπωση του marking theorem.

Έστω μία διαδικασία Poisson, Π στο χώρο S με μέσο μέτρο μ και κατανομή πιθανότητας $p(x, \cdot)$ στον M . Η κατανομή πιθανότητας $p(x, \cdot)$ εξαρτάται από το $x \in S$ με τρόπο τέτοιο ώστε για $B \subseteq M$ η $p(\cdot, B)$ να είναι μετρήσιμη συνάρτηση στο χώρο S . Ένα 'σημαδιών' της Π είναι το τυχαίο σύνολο $\Pi^* = \{(X, m_X); X \in \Pi\}$ του χώρου (X, m_X) και του οποίου η προβολή στον S είναι η Π και η δεσμευμένη κατανομή του Π^* δεδομένου της Π κάνει τις m_X ανεξάρτητες με κατανομές $p(X, \cdot)$.

Θεώρημα 2.5.2. Μαρκαρισμένες διαδικασίες Poisson

The marking theorem

Το τυχαίο υποσύνολο Π^* είναι διαδικασία Poisson στον $S \times M$ με μέσο μέτρο μ^* δίνεται από :

$$\mu^*(C) = \int \int_{(x,m) \in C} \mu(dx) p(x, dm)$$

Απόδειξη

Για κάθε μετρήσιμη συνάρτηση f στο $S \times M$ θεωρούμε τη διαδικασία

$$\Sigma^* = \sum_{X \in \Pi} f(X, m_X)$$

Δεδομένου ότι Π , το Σ^* είναι άθροισμα των ανεξάρτητων τυχαίων μεταβλητών $f(X, m_X)$. Άρα θα έχουμε ότι:

$$E\{e^{-\Sigma^*}|\Pi\} = \prod_{X \in \Pi} E\{e^{-f(X, m_x)}|\Pi\} = \prod_{X \in \Pi} \int_M e^{-f(X, m)p(X, dm)}.$$

Συμβολίζουμε με $f_*(x) = -\log \int_M e^{-f(X, m)}p(x, dm)$.

Αν στη χαρακτηριστική συναρτησοειδή για τη διαδικασία Π ,

$$\mathbb{E}(e^{-\Sigma_f}) = \exp\left\{-\int_S (1 - e^{-f(x)})\mu(dx)\right\}$$

αντικαταστήσουμε την f με f_* θα πάρουμε:

$$\begin{aligned} E\{e^{-\Sigma^*}\} &= \exp\left\{-\int_S (1 - e^{-f_*(x)})\mu(dx)\right\} \\ &= \exp\left\{-\int_S \int_M (1 - e^{-f(x, m)})\mu(dx)p(x, dm)\right\} \\ &= \exp\left\{-\int_{S \times M} (1 - e^{-f})d\mu^*\right\} \end{aligned}$$

Δηλαδή δείξαμε ότι η διαδικασία Π^* είναι επίσης διαδικασία Poisson με μέσο μέτρο μ^* . $\square$

Κεφάλαιο 3

Μπεϋζιανή Συμπερασματολογία και Μαρκοβιανές αλυσίδες

Στο κεφάλαιο αυτό θα γίνει μια σύντομη αναφορά στις βασικές αρχές της μπεϋζιανής ανάλυσης καθώς επίσης και σε βασικές έννοιες των μαρκοβιανών αλυσίδων. Η βασική βιβλιογραφία που χρησιμοποιήθηκε σε αυτό το κεφάλαιο είναι: “ C.P.Robert (2007)” , “ L.Meligkotsidou (2008)” για την μπεϋζιανή ανάλυση και ‘ Ο. Χρυσαφίνου (2004)’ , “ S.M. Ross (2006)” , “D.A. Levin, Y. Peres, E.L. Wilmer, (2002)” , “ O.Häggström” “J.Moller (2003)” ,για τις μαρκοβιανές αλυσίδες.

3.1 Μπεϋζιανή στατιστική

Στη στατιστική ο κύριος σκοπός μας είναι βάσει κάποιου δείγματος, ή κάποιων μετρήσεων και γενικότερα βάσει δεδομένων που συλλέξαμε, να εκτιμήσουμε τις παραμέτρους του αντίστοιχου προβλήματος που μας απασχολεί. Στην κλασσική στατιστική χρησιμοποιούμε την πληροφορία του δείγματος, εκτιμούμε την παράμετρο ή τις παραμέτρους που μας ενδιαφέρουν και εξαγάγουμε διάφορα συμπεράσματα. Η μπεϋζιανή στατιστική όμως, έχοντας ως βάση για τη θεμελίωση της το θεώρημα του Bayes, ενώ χρησιμοποιεί τα δεδομένα που έχουμε στη διάθεση μας, λαμβάνει υπόψη και άλλες πληροφορίες που μπορεί να έχουμε για το συγκεκριμένο πρόβλημα. Δηλαδή αν υπάρχουν παλαιότερες μελέτες οι οποίες μπορεί να επηρεάζουν ή να αφορούν το ίδιο

θέμα, λαμβάνονται υπόψη στην παρούσα εργασία με σκοπό τα πιο ασφαλή και με μικρότερα σφάλματα αποτελέσματα.

Θεώρημα 3.1.1. Το θεώρημα του Bayes

Bayes' theorem

Έστω A, B δύο γεγονότα με $P(A) > 0$. Τότε η δεσμευμένη πιθανότητα να συμβεί το γεγονός B δεδομένου του A δίνεται από :

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

Το θεώρημα μπορεί να επεκταθεί και για k ξένα γεγονότα του ίδιου χώρου πιθανότητας.

Συγκεκριμένα έστω $C_1, \dots, C_k$ όπου $C_j \in \Omega$ και $C_i \cap C_j = \emptyset$, $i \neq j$ και $C_1 \cup \dots \cup C_k = \Omega$. Τότε:

$$P(C_i|A) = \frac{P(A|C_i)P(C_i)}{\sum_{j=1}^k P(A|C_j)P(C_j)}, \quad i = 1, \dots, k$$

Γενικότερα για μία τυχαία μεταβλητή θ θα ισχύει ότι:

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)} = \frac{f(x|\theta)f(\theta)}{\int f(\theta)f(x|\theta)d(\theta)}$$

Εάν έχουμε να αντιμετωπίσουμε πολυπαραμετρικά προβλήματα, δηλαδή αν η παράμετρος θ είναι k -διάστατη, $\theta = (\theta_1, \dots, \theta_k)$ τότε η σταθερά κανονικοποίησης του θεωρήματος Bayes είναι ένα k -διάστατο ολοκλήρωμα. Στην περίπτωση αυτή υπολογίζουμε για κάθε διάσταση την περιθώρια εκ των υστέρων κατανομή.

$$f(\theta_i|x) = \int_{\theta_1} \dots \int_{\theta_{i-1}} \int_{\theta_{i+1}} \dots \int_{\theta_k} f(\theta|x) d\theta_1 \dots d\theta_{i-1} d\theta_{i+1} \dots d\theta_k$$

Καθώς είναι δύσκολο πολλές φορές να υπολογιστεί ένα ολοκλήρωμα διάστασης $k - 1$ χρήσιμη είναι η δεσμευμένη εκ των υστέρων κατανομή του θ_i .

Θεωρούμε δηλαδή ως σταθερές όλες τις παραμέτρους εκτός του θ_i (και τις συμβολίζουμε με θ_{-i}) και παίρνουμε την δεσμευμένη εκ των υστέρων κατανομή ως συνάρτηση του θ_i , $f(\theta_i|x, \theta_{-i}) \propto f(\theta|x)$

Στην Μπεϋζιανή στατιστική αντιμετωπίζουμε δηλαδή την άγνωστη παράμετρο ως τυχαία μεταβλητή και για να μπορέσουμε να κάνουμε τη συμπερασματολογία ακολουθούμε τα ακόλουθα βήματα:

- Κατασκευή του μοντέλου πιθανοφάνειας του μοντέλου $f(x|\theta)$.
- Επιλογή κατάλληλης εκ των προτέρων κατανομής (prior distribution) $f(\theta)$.
- Υπολογισμός της εκ των υστέρων κατανομής (posterior distribution) $f(\theta|x)$ χρησιμοποιώντας το θεώρημα του Bayes ότι δηλαδή $f(\theta|x) \propto f(\theta)f(x|\theta)$.
- Εξάγουμε τα συμπεράσματά μας από την εκ των υστέρων κατανομή.

Παράδειγμα

Έστω το μοντέλο πιθανοφάνειας $X \sim \text{Binomial}(n, \theta)$. Θέλουμε να εκτιμήσουμε την εκ των υστέρων κατανομή της παραμέτρου θ .

Έχουμε:

$$f(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \quad x = 0, 1, 2, \dots, n \quad 0 \leq \theta \leq 1$$

Έστω ότι οι εκ των προτέρων πεποιθήσεις μας για το θ ακολουθούν κατανομή $\text{Beta}(p, q)$.

$$\begin{aligned} (\theta) &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \theta^{p-1} (1 - \theta)^{q-1} \quad (0 < \theta < 1) \\ &\propto \theta^{p-1} (1 - \theta)^{q-1} \end{aligned}$$

Συνεπώς από το θεώρημα του Bayes η εκ των υστέρων κατανομή του θ είναι:

$$\begin{aligned} f(\theta|x) &\propto f(\theta)f(x|\theta) \\ &\propto \theta^{p-1} (1 - \theta)^{q-1} \times \theta^x (1 - \theta)^{n-x} \\ &= \theta^{p+x-1} (1 - \theta)^{q+n-x-1} \end{aligned}$$

Η παράμετρος θ δηλαδή έχει εκ των υστέρων κατανομή $Beta(p+x, q+n-x)$.

□

Παράδειγμα

Έστω τυχαίο δείγμα μεγέθους n από Poisson, $\mathcal{P}(\theta)$.

Η κατανομή του X (για μία παρατήρηση) θα είναι:

$$f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad \theta \geq 0$$

Η πιθανοφάνεια θα είναι:

$$f(x|\theta) = \prod_{i=1}^n \frac{\theta^{x_i} e^{-\theta}}{x_i!} \propto \theta^{\sum_{i=1}^n x_i} e^{-n\theta}$$

Έστω ότι οι εκ των προτέρων πεποιθήσεις μας χαρακτηρίζονται από τη Γαμμα κατανομή με παραμέτρους p, q .

$$\begin{aligned} f(\theta) &= \frac{p^q}{\Gamma(p)} \theta^{p-1} \exp\{-q\theta\}, \quad \theta > 0 \\ f(\theta|x) &\propto \theta^{p-1} \exp\{-q\theta\} \theta^{\sum_{i=1}^n x_i} \exp\{-n\theta\} \\ &= \theta^{\sum_{i=1}^n x_i + p - 1} \exp\{-(q+n)\theta\} \end{aligned}$$

Η εκ των υστέρων κατανομή δηλαδή του θ θα είναι $Gamma(p + \sum x_i, q+n)$.

□

3.1.1 Καθορισμός των εκ των προτέρων κατανομών

Στην Μπεϋζιανή στατιστική ένα από τα πιο κρίσιμα σημεία για την ανάλυση των δεδομένων και ίσως το πιο αμφιλεγόμενο σημείο για τους επιστήμονες είναι η επιλογή της εκ των προτέρων κατανομής. Η εκ των προτέρων κατανομή παίζει σημαντικό ρόλο στον καθορισμό της εκ των υστέρων κατανομής και επομένως έχει επίδραση στη συμπερασματολογία. Η επιλογή της κατανομής αυτή δεν είναι μοναδική και μπορεί να έχει μικρή, μέτρια ή τεράστια

επίδραση στην εξαγωγή των συμπερασμάτων. Αυτό το επιχείρημα χρησιμοποιούν οι πιο δύσπιστοι επιστήμονες όσον αφορά την μπεϋζιανή στατιστική για να την κατακρίνουν. Δηλαδή το ότι κάποιος επιστήμονας μπορεί να πάρει στην ουσία όποιο αποτέλεσμα θέλει, ανάλογα με την επιλογή της εκ των προτέρων κατανομής, είναι το σημείο στο οποίο οι “ κλασσικοί ” στατιστικοί απορρίπτουν την μπεϋζιανή στατιστική. Εντούτοις οι μπεϋζιανοί επιστήμονες αντιμετωπίζουν αυτό το πρόβλημα με διάφορους τρόπους. Εάν η πληροφορία που έχουμε δεν είναι αρκετή ή δεν έχουμε καθόλου εκ των προτέρων πληροφορία για την εκτίμηση της παραμέτρου, μπορούμε να χρησιμοποιήσουμε για παράδειγμα τη μη πληροφοριακή εκ των προτέρων κατανομή του Jeffrey και τότε η συμπερασματολογία δεν εξαρτάται τόσο από την εκ των προτέρων κατανομή. Επίσης υπάρχουν μέθοδοι όπως η ανάλυση ευαισθησίας με την οποία μπορεί να ελεγχθεί η επίδραση της εκ των προτέρων κατανομής στη συμπερασματολογία.

Οι συζυγείς εκ των προτέρων κατανομές

Conjugate priors

Ορισμός 3.1.1. Συζυγής συνάρτηση πιθανοφάνειας

Μία οικογένεια κατανομών πιθανότητας F στον Ω ονομάζεται συζυγής της συνάρτησης πιθανοφάνειας $f(x|\theta)$ αν για κάθε $\pi \in F$ η εκ των υστέρων κατανομή $\pi(\theta|x)$ ανήκει επίσης στην οικογένεια F .

Οι συζυγείς εκ των προτέρων κατανομές συνδέονται άμεσα με την εκθετική οικογένεια κατανομών, καθώς εάν μια κατανομή ανήκει στην εκθετική οικογένεια τότε η συζυγής εκ των προτέρων κατανομή της είναι χαρακτηριστική και μπορεί να υπολογιστεί άμεσα.

Ορισμός 3.1.2. Εκθετική οικογένεια κατανομών

Έστω μ ένα σ -πεπερασμένο μέτρο στον δειγματικό χώρο Ω και έστω Θ ο παραμετρικός χώρος. Έστω h, g συναρτήσεις από τον Ω και Θ στον $\mathbb{R}_+$ αντίστοιχα και έστω t, c συναρτήσεις από το Θ και Ω στον $\mathbb{R}^k$. Η οικογένεια κατανομών για τις οποίες ισχύει ότι:

$$f(x|\theta) = h(x)g(\theta) \exp\{t(x)c(\theta)\}$$

τέτοιες ώστε $\int f(x|\theta) = 1$, καλείται *εκθετική οικογένεια κατανομών διάστασης k* .

Βάσει του ορισμού έστω ότι έχουμε ένα τυχαίο δείγμα από την $(x_1, x_2, \dots, x_n)$ με κατανομή που ανήκει στην εκθετική οικογένεια κατανομών. Σύμφωνα με τον πιο πάνω ορισμό η πιθανοφάνεια του μοντέλου είναι :

$$\begin{aligned} f(x|\theta) &= \left[\prod_{i=1}^n h(x_i) \right] [g(\theta)]^n \exp\left\{ \sum_{i=1}^n t(x_i)c(\theta) \right\} \\ &\propto [g(\theta)]^n \exp\left\{ \sum_{i=1}^n t(x_i)c(\theta) \right\} \end{aligned}$$

Η εκ των προτερων κατανομή θα είναι της μορφής:

$$f(\theta) \propto [g(\theta)]^d \exp\{bc(\theta)\}$$

Η εκ των υστέρων κατανομή θα είναι

$$\begin{aligned} f(\theta|x) &\propto f(\theta)f(x|\theta) \\ &\propto [g(\theta)]^{n+d} \exp\{b + \sum_{i=1}^n t(x_i)c(\theta)\} \end{aligned}$$

Εάν έχουμε συνεπώς τυχαίο δείγμα από κατανομή που ανήκει στην εκθετική οικογένεια και θέλουμε να υπολογίσουμε τη συζυγή εκ των προτέρων κατανομή της τότε:

- Τη φέρνουμε στη μορφή του που αναφέρθηκε στον ορισμό και βρίσκουμε τις συναρτήσεις h, g, t, c .
- Παίρνουμε ως εκ των προτέρων κατανομή τη κατανομή με συνάρτηση $f(\theta) \propto [g(\theta)]^d \exp\{bc(\theta)\}$.

Παράδειγμα

Έστω τυχαίο δείγμα από κατανομή Bernoulli. Η πιθανοφάνεια του μοντέλου θα είναι:

$$\begin{aligned} f(x|\theta) &= \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &= \binom{n}{x} (1-\theta)^n \exp \left\{ x \log \left(\frac{\theta}{1-\theta} \right) \right\} \end{aligned}$$

Οι συναρτήσεις h, g, t, c θα είναι:

$$\begin{aligned} h(x) &= \binom{n}{x} \\ g(\theta) &= (1-\theta)^n \\ t(x) &= x \\ c(\theta) &= \log \left(\frac{\theta}{1-\theta} \right) \end{aligned}$$

Η συζυγής εκ των προτέρων κατανομή θα είναι:

$$f(\theta) \propto [(1-\theta)^n]^d \exp \left\{ b \log \left(\frac{\theta}{1-\theta} \right) \right\} = (1-\theta)^{nd} \left(\frac{\theta}{1-\theta} \right)^b$$

Η εκ των υστέρων κατανομή θα είναι $Beta(\beta = nd + 1, \alpha = b + 1)$

⊠

Οι συζυγείς εκ των προτέρων κατανομές συχνά χρησιμοποιούνται στη με-
ϋζιανή ανάλυση καθώς έχουν είναι εύκολες στον υπολογισμό τους και έχουν
και χρήσιμες ιδιότητες όπως για παράδειγμα αν έχουμε πολλές συζυγείς
εκ των προτέρων κατανομές με διάφορα βάρη πιθανότητας τότε και η μίξη
τους θα είναι συζυγής κατανομή.¹ Εντούτοις οι συζυγείς εκ των προτέρων
κατανομές είναι πιθανόν να μην μπορούν να υπολογιστούν σε διάφορα προ-
βλήματα.

Μη πληροφοριακές εκ των προτέρων κατανομές

Εάν σε ένα πρόβλημα δεν έχουμε κάποια εκ των προτέρων πληροφορία διαθέσιμη
τότε ο υπολογισμός των υπερπαραμέτρων b, d για τις συζυγείς εκ των προτέρων

¹Στη βιβλιογραφία υπάρχουν πίνακες με τις κατάλληλες παραμέτρους των συζυγών
κατανομών. Ένας τέτοιος πίνακας υπάρχει και στο βιβλίο "C.P. Robert, (2007)

κατανομές που αναφέρθηκαν πιο πάνω δεν είναι εφικτές. Έτσι καθώς υπάρχει άγνοια για την παράμετρο που θέλουμε να εκτιμήσουμε χρησιμοποιούμε το δείγμα που έχουμε στη διάθεση μας και κατασκευάζουμε εκ των προτέρων κατανομές τις οποίες ονομάζουμε μη πληροφοριακές.

Η μη πληροφοριακή κατανομή του Jeffrey

Η εκ των προτέρων κατανομή του Jeffrey βασίζεται στην πληροφορία του Fisher, η οποία ορίζεται ως:

$$I(\theta) = -E \left\{ \frac{d^2 l(\theta)}{d\theta^2} \right\} = E \left\{ \left(\frac{dl(\theta)}{d\theta} \right)^2 \right\}$$

όπου $l(\theta) = \log f(x|\theta)$ και Θ ο παραμετρικός χώρος.

Η εκ των προτέρων κατανομή του Jeffrey ορίζεται ως $J_{\Theta}(\theta) \propto |I(\theta)|^{\frac{1}{2}}$

Η κατανομή αυτή είναι αρκετά εύχρηστη καθώς είναι αμετάβλητη ως προς τους 1-1 μετασχηματισμούς, δηλαδή αν έχουμε ένα μετασχηματισμό για την παράμετρο θ , $\phi(\theta)$ τότε ο εκ των προτέρων κατανομή του Jeffrey γίνεται

$$J(\phi) = J(\theta) \left| \frac{d\theta}{d\phi} \right|$$

3.1.2 Η σημειακή εκτίμηση στη μπεϋζιανή στατιστική

Στην κλασσική στατιστική ένας από τους στόχους κάθε μελέτης είναι συνήθως η σημειακή εκτίμηση της παραμέτρου έτσι ώστε να μπορεί να γίνει η κατάλληλη συμπερασματολογία. Στην μπεϋζιανή στατιστική η συμπερασματολογία γίνεται όμως από την εκ των υστέρων κατανομή. Είναι όμως πολλές φορές χρήσιμο να δώσουμε μια σημειακή εκτίμηση για την παράμετρο που μας αφορά. Αν για παράδειγμα έχουμε να πάρουμε μια απόφαση για ένα πρόβλημα θα πρέπει να μπορούμε να δώσουμε μια συγκεκριμένη τιμή στην παράμετρο η οποία θα είναι και η απόφαση που θα πάρουμε. Στην προσπάθεια μας λοιπόν να περιορίσουμε την πληροφορία της εκ των υστέρων κατανομής σε μία μόνο εκτίμηση, χρησιμοποιούμε μια συνάρτηση που ονομάζουμε συνάρτηση απώλειας $L(\theta, \alpha)$ και η οποία μετράει την απώλεια, το

σφάλμα όταν εκτιμούμε την παράμετρο θ με α . Η σημειακή εκτίμηση τότε θα είναι η τιμή η οποία ελαχιστοποιεί τη εκ των υστέρων μέση απώλεια.

$$\hat{\theta} = \operatorname{argmin}_{\alpha} E[L(\theta, \alpha)|x]$$

Αναφέρονται ενδεικτικά οι πιο συνηθισμένες συναρτήσεις απώλειας και η σημειακή εκτίμηση τους:

- Δευτεροβάθμια (ή τετραγωνική) συνάρτηση απώλειας

$$L(\theta, \alpha) = (\theta - \alpha)^2$$

Η εκτιμήτρια είναι η εκ των υστέρων μέση τιμή $\hat{\theta} = E(\theta|\alpha)$

- Συνάρτηση απώλειας απόλυτου σφάλματος

$$L(\theta, \alpha) = |\theta - \alpha|$$

Η εκτιμήτρια σε αυτή την περίπτωση είναι η διάμεσος m η οποία είναι τέτοια ώστε $P(\theta < m|x) = P(\theta > m|x) = \frac{1}{2}$

- 0-1 Συνάρτηση απώλειας

$$L(\theta, \alpha) = \begin{cases} 0, & \text{αν } |\theta - \alpha| \leq \epsilon \\ 1, & \text{αν } |\theta - \alpha| > \epsilon \end{cases}$$

Τότε η εκτιμήτρια είναι η κορυφή της εκ των υστέρων κατανομής.

3.1.3 Διαστήματα και περιοχές αξιοπιστίας

Όπως είναι γνωστό στην κλασσική στατιστική συμπερασματολογία δημιουργούμε τα διαστήματα εμπιστοσύνης. Για να ποσοτικοποιήσουμε την αβεβαιότητα σχετικά με τις εκτιμήσεις στη μπεϋζιανή στατιστική έχουμε τα διαστήματα και τις περιοχές αξιοπιστίας.

Ορισμός 3.1.3. Περιοχή αξιοπιστίας

Ορίζουμε μια περιοχή $C_\alpha(x)$ μια $100(1 - \alpha)\%$ περιοχή αξιοπιστίας για την παράμετρο θ αν ισχύει

$$\int_{C_\alpha(x)} f(\theta|x)d\theta = 1 - \alpha$$

Είναι φανερό ότι η περιοχή αξιοπιστίας δεν ορίζεται μονοσήμαντα. Διαλέγουμε λοιπόν την περιοχή αξιοπιστίας που περιέχει τις πιο πιθανές τιμές του θ κάτω από τον επιπλέον περιορισμό για το μήκος του διαστήματος να είναι το ελάχιστο δυνατό. Έτσι, ζητάμε περιοχή

$$C_\alpha(x) = \{\theta : f(\theta|x) \geq \gamma\} \text{ με } \int_{C_\alpha(x)} f(\theta|x)d\theta = 1 - \alpha$$

Η περιοχή αξιοπιστίας σε αυτή την περίπτωση ονομάζεται **περιοχή αξιοπιστίας υψηλότερης πυκνότητας** (*highest posterior density region*).

Υπάρχει κάποια λεπτή διαφορά στην ερμηνεία των διαστημάτων εμπιστοσύνης στην κλασσική στατιστική με την ερμηνεία των περιοχών αξιοπιστίας στη μπεϋζιανή στατιστική.

Ένα διάστημα εμπιστοσύνης $100(1 - \alpha)\%$ στην κλασσική στατιστική σημαίνει ότι από τα 100 πειράματα που πραγματοποιούμε η πραγματική τιμή της παραμέτρου θα βρίσκεται μέσα στο διάστημα εμπιστοσύνης στις $1 - \alpha$ φορές. Στη μπεϋζιανή στατιστική μία περιοχή αξιοπιστίας $100(1 - \alpha)\%$ είναι η περιοχή εκείνη κατά την οποία με πιθανότητα $1 - \alpha$ (βάσει της εκ των υστέρων κατανομής) η παράμετρος θα βρίσκεται μέσα.

3.1.4 Έλεγχος Υποθέσεων και σύγκριση μοντέλων

Έλεγχος υποθέσεων

Έστω ότι έχουμε τον έλεγχο υποθέσεων $H_1 : \theta \in \Omega_1$ έναντι της $H_0 : \theta \in \Omega_0$.

Βασίζουμε τον έλεγχο υποθέσεων στο λόγο των εκ των υστέρων κατανομών:

$$\lambda_B = \frac{f(\theta \in \Omega_1|x)}{f(\theta \in \Omega_0|x)} = \frac{P(\theta \in \Omega_1)f(x|\theta \in \Omega_1)}{f(\theta \in \Omega_0)f(x|\theta \in \Omega_0)}$$

Μεγάλες τιμές του λ υποδηλώνουν προτίμηση υπέρ της H_1

Σύγκριση Μοντέλων

Ο έλεγχος υποθέσεω ν μπορεί να γενικευθεί και για τη σύγκριση μοντέλων. Εάν για παράδειγμα έχουμε κ διαφορετικά μοντέλα τότε αφού υπολογίσουμε τις εκ των προτέρων πιθανότητες των μοντέλων και τις δεσμευμένες κατανομές $f(\theta|M_i)$ βρίσκουμε τις εκ των υστέρων πιθανότητες $P(M_i|x)$ και τις εκ των υστέρων κατανομές του θ κάτω από τα μοντέλα $f(\theta|x, M_i)$. Σημειώνεται ότι η εκ των υστέρων πιθανότητα του i -οστού μοντέλου υπολογίζεται βάσει του θεωρήματος Bayes ως εξής:

$$P(M_j|x) = \frac{P(M_j)f(x|M_j)}{\sum_{i=1}^k P(M_i)f(x|M_i)} \quad j = 1, \dots, k$$

όπου $f(x|M_j)$ η περιθώρια πιθανοφάνεια του μοντέλου M_j .

Επιλέγουμε συνήθως το πιο πιθανό μοντέλο. Δηλαδή το μοντέλο που έχει τη μεγαλύτερη εκ των υστέρων πιθανότητα.

3.2 Μαρκοβιανές Αλυσίδες - Η μέθοδος MCMC

Στα πολυπαραμετρικά προβλήματα ο υπολογισμός σε κλειστή μορφή της εκ των υστέρων κατανομής είναι πολλές φορές δύσκολο εώς και αδύνατο γίνει. Έτσι εάν πάρουμε με κάποιο τρόπο δείγμα από την κατανομή αυτή, τότε θα μπορεί να γίνει η συμπερασματολογία. Το προαναφερθέν πρόβλημα μπορεί να αντιμετωπιστεί προσομοιώνοντας μία μαρκοβιανή αλυσίδα με τη μέθοδο Marcov Chain Monte Carlo (MCMC), η οποία θα έχει στάσιμη κατανομή, την κατανομή που μας ενδιαφέρει.

3.2.1 Βασικοί ορισμοί μαρκοβιανών αλυσίδων και το Εργοδικό θεώρημα

Έστω τυχαίες μεταβλητές $X_0, X_1, \dots$ όπου με X_n δηλώνεται ο ως η κατάσταση του συστήματος στη χρονική στιγμή n . Υποθέτουμε ότι το σύνολο των δυνατών καταστάσεων του συστήματος είναι το σύνολο S . Αν υπάρχουν αριθμοί p_{ij} , $i, j \in S$ τέτοιοι ώστε όταν η διαδικασία είναι στην κατάσταση i και ανεξάρτητα από τις καταστάσεις που βρισκόταν στο παρελθόν θα βρίσκεται στην κατάσταση j με πιθανότητα μετάβασης p_{ij} τότε το σύνολο των τυχαίων μεταβλητών $\{X_n, n \geq 0\}$ σχηματίζει μία Μαρκοβιανή διαδικασία με πίνακα πιθανοτήτων μετάβασης $P_{ij} = (p_{ij})$, $i, j \in S$ και οι οποίες ικανοποιούν τη σχέση $\sum_{j \in S} p_{ij} = 1$, $i \in S$.

Ορισμός 3.2.1. Αδιαχώριστο η Μαρκοβιανή αλυσίδα

Μια μαρκοβιανή αλυσίδα καλείται αδιαχώριστη (ή ανάγωγη) εάν για κάθε ζεύγος καταστάσεων i, j υπάρχει θετική πιθανότητα αν η αλυσίδα ξεκινήσει από την κατάσταση i να πάει κάποτε στην κατάσταση j .

Ορισμός 3.2.2. Έμμονη (επαναληπτική) και μεταβατική μαρκοβιανή αλυσίδα

Έστω $f_j(n) = P(X_n = j, X_r \neq j, r = 1, \dots, n-1 | X_0 = j)$ η πιθανότητα πρώτης επανόδου στην κατάσταση j και $f_j = \sum_{n=1}^{\infty} f_j(n)$ η πιθανότητα το σύστημα να γυρίσει κάποτε στην κατάσταση j . Η μαρκοβιανή αλυσίδα καλείται έμμονη ή επαναληπτική αν $f_j = 1$ και μεταβατική εάν $f_j < 1$.

Ορισμός 3.2.3. Μέσος χρόνος επανόδου

Ο μέσος χρόνος επανόδου της j ορίζεται ως $\mu_j = \sum_{n=1}^{\infty} n f_j(n)$

Στην περίπτωση της επαναληπτικής κατάστασης αν ο $\mu_j < \infty$ η κατάσταση j καλείται θετική ενώ εάν ο μέσος χρόνος επανόδου είναι άπειρος τότε η κατάσταση θα καλείται μηδενική.

Στην περίπτωση της μεταβατικής κατάστασης ο μέσος χρόνος επανόδου θα είναι πάντα άπειρος.

Ορισμός 3.2.4. Περίοδος μαρκοβιανής αλυσίδας

Περίοδος, έστω $d(j)$ μιας κατάστασης j είναι ο μέγιστος κοινός διαιρέτης των ακεραίων n για τους οποίους θα ισχύει $p_{jj}(n) > 0$. Αν $d(j) = 1$ τότες η j καλείται απεριοδική.

Μια κατάσταση που είναι επαναληπτική, θετική και απεριοδική καλείται εργοδική.

Θεώρημα 3.2.1. Έστω μία ανάγωγη και απεριοδική μαρκοβιανή αλυσίδα με πιθανότητες μεταπήδησης (p_{ij}) και χώρο καταστάσεων S . Η αλυσίδα έχει μοναδική στάσιμη κατανομή αν και μόνο αν είναι εργοδική. Τότε η στάσιμη κατανομή συμπίπτει με την οριακή κατανομή (π_j) όπου $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n)$

Για τη στάσιμη κατανομή π ισχύουν:

$$\pi_j = \sum_{i \in S} \pi_i p_{ij} \text{ (εξισώσεις ισορροπίας)}$$

$$\pi_j \geq 0$$

$$\sum_{j \in S} \pi_j = 1 \text{ (εξίσωση κανονικοποίησης)}$$

Τέλος παραθέτουμε το εργοδικό θεώρημα πάνω στο οποίο βασίζεται η ι-δέα τξ προσομοίωσης μέσω Μαρκοβιανής Αλυσίδας Μόντε Κάρλο (MCMC).

Θεώρημα 3.2.2. Εργοδικό θεώρημα

Ergodic Theorem

Έστω f μία πραγματική τυχαία συνάρτηση που ορίζεται στο χώρο S . Αν $\{X_n, n \geq 0\}$ είναι μία αδιαχώριστη μαρκοβιανή αλυσίδα τότε για οποιαδήποτε αρχική κατανομή μ θα ισχύει:

$$P_\mu \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(X_i) = E_\pi(f) \right\} = 1$$

όπου συμβολίζουμε με $E_\mu(f) = \sum_{x \in S} f(x)\mu(x)$, $E_\pi(f) = \sum_{x \in S} f(x)\pi(x)$ και π η στάσιμη κατανομή.

3.2.2 Ο αλγόριθμος Metropolis Hastings

Έστω λοιπόν ότι θέλουμε να προσομοιώσουμε μία τιμή για την τυχαία μεταβλητή X η οποία έχει συνάρτηση πιθανότητας f . Αν προσομοιώσουμε μια αδιαχώριστη απεριοδική μαρκοβιανή αλυσίδα οριακή συνάρτηση πιθανότητας π τότε θα μπορούμε να πρεσεγγίσουμε αυτή την τυχαία μεταβλητή, τρέχοντας την αλυσίδα για n βήματα και παίρνοντας την X_n όπου n αρκετά μεγάλο. Καθώς όμως μια μαρκοβιανή αλυσίδα εξαρτάται από την αρχική της κατανομή, είναι συνετό να τρέξουμε την αλυσίδα για κάποιο αριθμό φορών χωρίς να κρατήσουμε τις τιμές αυτές.

Έστω Q ο πίνακας μίας αδιαχώριστης μαρκοβιανής αλυσίδας με στοιχεία (q_{ij}) , $i, j \in S$. Ορίζουμε μια μαρκοβιανή αλυσίδα $\{X_n, n \geq 0\}$ ως εξής:

Όταν $X_n = i$ η τυχαία μεταβλητή X τέτοια ώστε $\mathbb{P}\{X = j\} = q(i, j)$ $j \in S$ προσομοιώνεται. Αν $X = j$ τότε $X_{n+1} = j$ με πιθανότητα $a(i, j)$ αλλιώς η τιμή παραμένει η ίδια.

Με αυτό τον τρόπο κατασκευάζουμε μια μαρκοβιανή αλυσίδα με πιθανότητες μετάβασης

$$p_{ij} = q(i, j)a(i, j), \quad j \neq i$$

$$p_{ii} = q(i, j) + \sum_{k \neq i} q(i, k)(1 - a(i, k))$$

Η μαρκοβιανή αυτή αλυσίδα θα είναι αντιστέψιμη με στασιμες πιθανότητες $\pi(j)$ αν $\pi(i)p_{ij} = \pi_j p_{ji}$ $j \neq i$

Ισοδύναμα, $\pi(i)q(i, j)a(i, j) = \pi(j)q(j, i)a(j, i)$ και επομένως αρκεί να πάρουμε τις πιθανότητες $a(i, j) = \min\left(\frac{\pi(j)q(j, i)}{\pi(i)q(i, j)}, 1\right)$

Στην Μπεϋζιανή στατιστική ο αλγόριθμος Metropolis Hastings είναι πολύ χρήσιμος στην αντιμετώπιση πολυπαραμετρικών προβλημάτων και ιδιαίτερα στην εύρεση της εκ των υστέρων κατανομής. Ειδικότερα:

Έστω ότι θέλουμε να ότι έχουμε παρατηρήσεις $y = (y_1, \dots, y_n)$ με πιθανοφάνεια $f(y|\theta)$ όπου $\theta = (\theta_1, \dots, \theta_d)$ και εκ των προτέρων κατανομή $f(\theta)$. Αρχίζουμε από αυθαίρετες αρχικές τιμές. Για την $(j+1)$ - επανάληψη του αλγορίθμου έχουμε τα εξής βήματα

- Προτείνουμε τιμή για το θ_1 , θ_1^{can} . Η τιμή θ_1^{can} προτείνεται από αυθαίρετη κατανομή με πυκνότητα πιθανότητας $q(\theta_1^{can}|\theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_d^{(j)})$

- Αποδεχόμαστε την προτεινόμενη τιμή για το θ_1 ως:

$$\theta_1^{(j+1)} = \begin{cases} \theta_1^{can} & \text{με πιθανότητα } p \\ \theta_1^{(j)} & \text{με πιθανότητα } 1 - p \end{cases}$$

όπου

$$p = \min \left\{ \frac{f(\theta_1^{can}|y, \theta_2^{(j)}, \dots, \theta_d^{(j)}) q(\theta_1^{(j)}|\theta_1^{can}, \theta_2^{(j)}, \dots, \theta_d^{(j)})}{f(\theta_1^{(j)}|y, \theta_2^{(j)}, \dots, \theta_d^{(j)}) q(\theta_1^{can}|\theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_d^{(j)})}, 1 \right\}$$

- Επαναλαμβάνουμε τα ίδια βήματα και για τις υπόλοιπες διαστάσεις και ακολούθως προχωράμε στην επόμενη επανάληψη του αλγορίθμου.

Όπως προαναφέρθηκε τρέχουμε τον αλγόριθμο για κάποιο αριθμό φορών που καθορίζουμε εμείς, πετάμε αυτές τις τιμές και έπειτα κρατάμε τις τιμές για τη συμπερασματολογία έτσι ώστε να επιτευχθεί η σύγκλιση του αλγορίθμου.

- Μια συνήθης επιλογή της κατανομής για τις προτεινόμενες τιμές είναι η κανονική κατανομή με μέσο την τιμή κάθε φορά της παραμέτρου (θ_i) και διασπορά τέτοια ώστε το ποσοστό αποδοχής των προτεινόμενων τιμών να είναι γύρω στο 30%. Στην περίπτωση αυτή ο αλγόριθμος ονομάζεται Αλγόριθμος Metropolis-Hastings τυχαίου περιπάτου. Έύκολα κανείς μπορεί να δει ότι σε αυτή την περίπτωση η πιθανότητα αποδοχής θα είναι

$$p = \min \left\{ \frac{f(\theta_1^{can}|y, \theta_2^{(j)}, \dots, \theta_d^{(j)})}{f(\theta_1^{(j)}|y, \theta_2^{(j)}, \dots, \theta_d^{(j)})}, 1 \right\}$$

3.2.3 Ο αλγόριθμος Gibbs

Έστω παρατηρήσεις $y = (y_1, \dots, y_n)$ με πιθανοφάνεια $f(y|\theta)$ όπου $\theta = (\theta_1, \dots, \theta_d)$ και εκ των προτέρων κατανομή $f(\theta)$. Εάν η εκ των υστέρων κατανομή μπορεί να βρεθεί σε κλειστή μορφή και είναι γνωστή τότε η προσομοίωση μπορεί γίνει ακολουθιακά και επαναληπτικά, προσομοιώνοντας από τις δεσμευμένες εκ των υστέρων κατανομές των επιμέρους παραμέτρων. Τα βήματα του αλγορίθμου είναι:

- Ξεκινάμε με τυχαίες τιμές $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$
- Στην j -οστή επανάληψη του αλγορίθμου προσομοιώνουμε νέα τιμή $\theta_1^{(j)}$ από τη δεσμευμένη εκ των υστέρων κατανομή $f(\theta_1|\theta_2^{(j-1)}, \dots, \theta_d^{(j-1)})$
- Προσομοιώνουμε νέα τιμή για το $\theta_2^{(j)}$ από τη δεσμευμένη εκ των υστέρων κατανομή $f(\theta_2|\theta_1^{(j)}, \theta_2^{(j-1)}, \dots, \theta_d^{(j-1)})$
- $\vdots$
- Προσομοιώνουμε νέα τιμή για το $\theta_d^{(j)}$ από τη δεσμευμένη εκ των υστέρων κατανομή $f(\theta_d|\theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{d-1}^{(j)})$

- Επαναλαμβάνουμε τη διαδικασία μέχρι να επιτευχθεί σύγκλιση.

Όπως και στον αλγόριθμο Metropolis-Hastings τρέχουμε τον αλγόριθμο για κάποιο αριθμό φορών μέχρι να επιτευθεί η σύγκλιση και μετά κρατάμε τις τιμές.

Παρατηρούμε ότι αυτός ο πιο απλός αλγόριθμος είναι ειδική περίπτωση του Metropolis-Hastings με πιθανότητες αποδοχής των προτεινόμενων τιμών ίσες με μονάδα.

Κεφάλαιο 4

Το πρόβλημα της Αστροφυσικής

4.1 Μια εισαγωγή στο πρόβλημα

Η συνάρτηση φωτεινότητας των γαλαξιών παίζει σημαντικό ρόλο στην κατανόηση του σύμπαντος. Μέσω της συνάρτησης φωτεινότητας μπορούμε να πάρουμε πολλά στοιχεία όσον αφορά τον πληθυσμό των γαλαξιών, την εξέλιξη των ουράνιων σωμάτων καθώς και την ιστορία της δημιουργίας και ανάπτυξης των άστρων. Τα παρατηρησιακά δεδομένα που υπάρχουν σήμερα δείχνουν ότι η εξέλιξη των γαλαξιών και η δημιουργία των υπέρμαζων μελανών οπών (μαύρες τρύπες) στο κέντρο τους σχετίζονται άμεσα. Η συνάρτηση φωτεινότητας είναι λοιπόν χρήσιμη και για την κατανόηση της δημιουργίας των μελανών οπών αλλά και για την κατανομή της μάζας τους. Ακόμη, η συνάρτηση φωτεινότητας μας επιτρέπει να μελετήσουμε την εξάρτηση των ιδιοτήτων των γαλαξιών με το περιβάλλον.

Τα παρατηρησιακά δεδομένα που έχουμε στη διάθεση μας περιέχουν σφάλματα και αυτό μπορεί να οδηγήσει σε λανθασμένους υπολογισμούς όσον αφορά τις παραμέτρους της συνάρτησης φωτεινότητας των γαλαξιών ή ακόμα και τη λανθασμένη επιλογή μοντέλου που περιγράφει τη φωτεινότητα των γαλαξιών. Συνεπώς η συνεργασία μεταξύ αστροφυσικών και στατιστικών κρίνεται αναγκαία ώστε να γίνει μία καλή εκτίμηση της συνάρτησης φωτεινότητας των γαλαξιών.

Έχουν προταθεί διάφορες μέθοδοι για την εκτίμηση της εν λόγω συνάρτησης, οι οποίες ασφαλώς παρουσιάζουν πλεονεκτήματα και μειονεκτήματα. Μια μέθοδος είναι η μέθοδος διακριτικοποίησης σε ίσα διαστήματα (binning method). Η συμπερασματολογία που βασίζεται σε αυτή τη μέθοδο είναι περιοριστική καθώς δεν μπορούμε να επεκταθούμε σε δεδομένα που έχουν δια-

φορετικό στήριγμα ενώ ταυτόχρονα δεν μπορούμε να κάνουμε την υπόθεση ότι η φωτεινότητα και η ερυθρομετατόπιση είναι εξαρτημένες. Η μέθοδος αυτή παρουσιάζει και επιπλέον μειονέκτημα καθώς η αυθαίρετη επιλογή του μήκους του διαστήματος για το διαχωρισμό μπορεί να οδηγήσει σε επιπλέον σφάλματα. Άλλη μέθοδος για την εκτίμηση της συνάρτησης φωτεινότητας, η οποία χρησιμοποιείται κατά κόρον στην αστροφυσική κοινότητα, είναι η μέθοδος της μέγιστης πιθανοφάνειας. Η μέθοδος αυτή είναι επίσης περιοριστική καθώς είναι πολλές φορές δύσκολη στους υπολογισμούς και απαιτούνται πολύπλοκοι χειρισμοί για τις κατανομές. Τα σφάλματα σε αυτή την περίπτωση θεωρούμε, τις πιο πολλές φορές, ότι ακολουθούν την κανονική κατανομή, μία υπόθεση που δεν είναι πάντα ρεαλιστική. Έτσι λόγω της πολυπλοκότητας της μεθόδου και της αναγκαιότητας επιπλέον υποθέσεις για τις μεταβλητές, προκύπτουν εκτιμήσεις με μεγάλα διαστήματα εμπιστοσύνης και μικρή ακρίβεια.

Η διάδοση και η αυξανόμενη αποδοχή της μπεϋζιανής στατιστικής από την επιστημονική κοινότητα, άρχισε να δημιουργεί και στους αστροφυσικούς ερωτήματα όσον αφορά την ορθότητα των αποτελεσμάτων τους με τις προαναφερθέντες μεθόδους. Πρόσφατες μελέτες έχουν δείξει ότι η αντιμετώπιση προβλημάτων με μπεϋζιανά εργαλεία δίνουν αποτελέσματα αρκετά διαφορετικά από αυτά που προέκυψαν με προηγούμενες μεθόδους. Στην παρούσα εργασία αναπτύσσουμε μία μπεϋζιανή προσέγγιση για την εκτίμηση των παραμέτρων της συνάρτησης φωτεινότητας των γαλαξιών. Η μέθοδος αυτή προσφέρει αρκετά πλεονεκτήματα καθώς μπορούμε να είμαστε πιο ευέλικτοι στις αρχικές υποθέσεις μας για την παραμετρική μορφή της συνάρτησης, μπορούμε να χειριστούμε πολυπαραμετρικά προβλήματα ευκολότερα με τη βοήθεια μεθόδων προσομοίωσης (MCMC) και οι εκτιμήσεις μας έχουν στενότερα διαστήματα αξιοπιστίας.

Στην εργασία αυτή παρουσιάζουμε τη μπεϋζιανή μεθοδολογία για την εκτίμηση της συνάρτησης φωτεινότητας. Αρχικά κάνουμε την ανάλυση για μία απλοποιημένη συνάρτηση φωτεινότητας και ακολούθως για πιο περίπλοκη συνάρτηση φωτεινότητας. Επίσης συγκρίνουμε τα αποτελέσματα που προκύπτουν μετά από μοντελοποίηση με διαδικασίες Poisson και διωνυμικές.

4.2 Το πρόβλημα της αστροφυσικής

4.2.1 Βασικές έννοιες της αστροφυσικής

Ορισμός 4.2.1. Γαλαξίας

Galaxy

Γαλαξίας είναι ένα Σύνολο αστεριών, σκόνης και αερίου τα οποία αλληλεπιδρούν μεταξύ τους βαρυτικά.

Ορισμός 4.2.2. Ενεργός γαλαξιακός πυρήνας

Active galactic nuclei (AGN)

Οι ενεργοί γαλαξιακοί πυρήνες είναι η κεντρική περιοχή ορισμένων γαλαξιών, όπου λαμβάνουν χώρα εξαιρετικά ενεργειακά και βίαια φαινόμενα. Έχουν πολύ ψηλές λαμπρότητες (ακτινοβολία) σε όλα τα μήκη κύματος, φαρδιές γραμμές εκπομπής στο οπτικό, υπεριώδες και ακτίνες Χ. Σχετίζονται με την πιθανή ύπαρξη μελανών οπών (μαύρες τρύπες) στο κέντρο τους και έχουν μικρό μέγεθος.

Ορισμός 4.2.3. Ερυθρομετατόπιση

Redshift (z)

Η ερυθρομετατόπιση είναι η παρατηρούμενη αλλαγή στη συχνότητα ενός κύματος όταν ο παρατηρητής βρίσκεται σε σχετική κίνηση με την πηγή που παράγει το κύμα. όσο μεγαλύτερη είναι η σχετική ταχύτητα τόσο μεγαλύτερη είναι και η ερυθρομετατόπιση. Στο σύμπαν όλα τα αντικείμενα απομακρύνονται μεταξύ τους (διαστολή σύμπαντος) και μάλιστα όσο μακρύτερα πηγαίνουμε τόσο μεγαλύτερη είναι η ταχύτητα απομάκρυνσης και άρα τόσο μεγαλύτερη είναι η ερυθρομετατόπιση. Επομένως αντικείμενα με μεγάλη ερυθρομετατόπιση βρίσκονται μακρύτερα. Έτσι με αυτό τον τρόπο γίνεται σύνδεση της ερυθρομετατόπισης με την απόσταση των ουράνιων σωμάτων.

Ορισμός 4.2.4. Φωτεινότητα

Luminosity (L)

Φωτεινότητα ορίζουμε την ενέργεια ανά μονάδα χρόνου που εκπέμπεται από ένα ουράνιο σώμα και είναι χαρακτηριστικό κάθε αντικειμένου.

Ορισμός 4.2.5. Απόσταση φωτεινότητας

Luminosity Distance (D_L)

Η απόσταση φωτεινότητας στην αστροφυσική εξαρτάται από την κοσμολογία (δηλαδή τη γεωμετρία του χωροχρόνου μέσα στο οποίο το φως ταξιδεύει). Η γεωμετρία του χωροχρόνου (κάτω από κάποιες υποθέσεις) περιγράφεται από 2 παραμέτρους: το ρυθμό διαστολής του σύμπαντος (Hubble constant H_0) και την πυκνότητα μάζας στο σύμπαν (Ω_{mass}). Άρα η πραγματική απόσταση ενός αντικειμένου από τη γη είναι συνάρτηση της ερυθρομετατόπισης (z), Hubble constant (H_0), Ω_{mass} .

Ορισμός 4.2.6. Συνάρτηση φωτεινότητας γαλαξιών

Luminosity function (ϕ)

Η συνάρτηση για την φωτεινότητα των γαλαξιών δίνεται από τη σχέση $\phi(L, z)dL$ και είναι ο αριθμός των πηγών ανά όγκο υπολογισμένο σε comoving συντεταγμένες (δηλαδή σε συντεταγμένες όπου η αριθμητική πυκνότητα των αντικειμένων παραμένει σταθερή) με φωτεινότητες στο διάστημα $L, L + dL$.

Η συνάρτηση φωτεινότητας συνδέεται με την πυκνότητα πιθανότητας των (L, z) μέσω της σχέσης:

$$p(L, z) = \frac{1}{N} \phi(L, z) \frac{dV}{dz}$$

όπου N είναι ο συνολικός αριθμός των πηγών στο παρατηρούμενο σύμπαν.

Επομένως αν ορίσουμε ως $p(L, z)dLdz$ την πιθανότητα να βρούμε πηγή στο διάστημα $L, L + dL$ και $z, z + dz$ θα ισχύει ότι:

$$p(L, z)dLdz = \frac{1}{N} \phi(L, z)dLdV \Rightarrow p(L, z) = \frac{1}{N} \phi(L, z) \frac{dV}{dz}$$

Σημειώνεται ότι $N = \int \int \phi(L, z)dLdV$

Αν υποθέσουμε παραμετρική μορφή για τη συνάρτηση φωτεινότητας με παραμέτρους θ , μπορούμε να εξάγουμε τη συνάρτηση πιθανοφάνειας για τα παρατηρούμενα δεδομένα. Καθώς κάθε έρευνα στην αστροφυσική δίνει κάποιο παρατηρούμενο δείγμα από φωτεινότητες και ερυθρομετατοπίσεις, όπου είναι ορατές μόνο πηγές που έχουν φωτεινότητα πάνω από κάποια ελάχιστη

τιμή και δεδομένης ερυθρομετατόπισης, υπάρχει σφάλμα στην εκτίμηση της συνάρτησης πιθανοφάνειας. Μία άλλη προσέγγιση είναι να θεωρήσουμε ότι ο αριθμός των πηγών στο σύμπαν ακολουθεί διαδικασία Poisson. Οι δύο προσεγγίσεις αυτές παρουσιάζονται από τους B.C.Kelly, X.Fan, M.Vestergaard, (2008). Άλλες εργασίες όπου έχει χρησιμοποιηθεί η Poisson μοντελοποίηση είναι οι: των H.L.Marshall, Y.Anvi, H.Tananbaum, G.Zamorani, (1983) , των J.Aird, K.Nandra, E.S.Laird, A.Georgakakis, M.L.N.Ashby, P.Barmby, A.L.Coil, J-S.Huang, A.M.Koekemoer, C.C.Steidel, C.N.A.Willmer, (2009) και του T.J. Lored, (2004)

Τα δεδομένα στην έρευνα είναι ανεξάρτητα και συνεπώς η συνάρτηση πιθανοφάνειας για όλες τις N πηγές στο σύμπαν παρατηρούμενες και μη θα είναι:

$$p(L, z|\theta) = \prod_{i=1}^N p(L_i, z_i|\theta)$$

Στην πραγματικότητα δεν γνωρίζουμε όλες τις ερυθρομετατοπίσεις και φωτεινότητες για όλες τις πηγές N , ούτε και πραγματικά γνωρίζουμε την τιμή του N , καθώς η έρευνα καλύπτει μόνο ένα μέρος του ουρανού. Επομένως η έρευνα θα περιέχει n πηγές και το N θα είναι άλλη μια παράμετρος στο πρόβλημα που θα πρέπει να εκτιμηθεί. Βάσει αυτής της ιδέας τα δεδομένα μπορούν να θεωρηθούν διωνυμικά με παραμέτρους n και πιθανότητα επιτυχίας η πιθανότητα να ανιχθευθεί μια πηγή.

4.2.2 Poisson δεδομένα

Στην παράγραφο αυτή παρουσιάζουμε τη μοντελοποίηση του προβλήματος με διαδικασίες Poisson, καθώς και τη μπεϋζιανή εκτίμηση αυτού του μοντέλου.

Ο αριθμός των πηγών σε διαφορίσιμες comoving συντεταγμένες όγκου $\left(\frac{dV}{dz}\right) dz$ και διαφορίσιμη φωτεινότητα dL είναι:¹

$$\phi(L, z)dLdV(z) = \phi(L, z)\frac{dV}{dz}dzdL$$

¹H.L.Marshall, Y.Anvi, H.Tananbaum, G.Zamorani, (1983)

Αν $\Omega(z, L) = (p(I = 1|z, L))$ είναι η περιοχή του ουρανού όπου οι πηγές μπορούν να ανιχνευθούν τότε η πιθανοφάνεια είναι το γινόμενο των πιθανοτήτων να παρατηρήσουμε ακριβώς μία πηγή ($dzdL$) για κάθε (z_i, L_i) , $i = 1, \dots, n$ ($i \in A_{obs}$) επί τις πιθανότητες να παρατηρηθούν 0 πηγές στις περιοχές που δεν μπορούμε να ανιχνεύσουμε πηγές. Άρα ο αναμενόμενος αριθμός των πηγών στο $dzdL$ για z, L στο δείγμα χρησιμοποιώντας πιθανότητες Poisson με παραμετρο $\lambda(z, L)dzdL$ είναι:

$$\lambda(z, L)dzdL = \phi(z, L)\Omega(z, L)\frac{dV}{dz}dzdL$$

Η πιθανοφάνεια είναι:

$$p(L, z|\theta, N) = \prod_{i \in A_{obs}} \frac{\lambda(z_i, L_i)^1 dzdL}{1!} e^{-\lambda(z_i, L_i) dzdL} \prod_{j \in A_{mis}} e^{-\lambda(z_j, L_j) dzdL} \frac{(\lambda(z, L) dzdL)^0}{0!}$$

Άρα

$$p(L, z|\theta, N) = \prod_{i \in A_{obs}} \lambda(z_i, L_i) dzdL e^{-\lambda(z_i, L_i) dzdL} \prod_{j \in A_{mis}} e^{-\lambda(z_j, L_j) dzdL} \lambda(z, L) dzdL$$

Η log- πιθανοφάνεια του μοντέλου θα είναι:

$$\begin{aligned} \log(p(L, z|\theta, N)) &= \sum_{i=1}^n \log(\lambda(z_i, L_i)) - \sum_{i=1}^n \lambda(z_i, L_i) dzdL + C = \\ &= \sum_{i=1}^n \log(\phi(z_i, L_i)) - \int \int \phi(z, L) \Omega(z, L) \frac{dV}{dz} dzdL + C \end{aligned}$$

$$\text{όπου } C = - \sum_{j \in A_{mis}} \lambda(z_j, L_j) dL dz$$

Αν αφαιρέσουμε τους όρους που είναι ανεξάρτητοι του μοντέλου τότε:

$$\log(p(L_{obs}, z_{obs}|\theta, N)) = \sum_{i=1}^n \log(\phi(z_i, L_i)) - \int \int p(I = 1|L, z) \phi(z, L|\theta, N) \frac{dV}{dz} dzdL$$

$$\log(L_{obs}, z_{obs}|\theta, N) = \sum_{i=1}^n \log(Np(z_i, L_i|\theta)) - \int \int p(I = 1|L, z) Np(z, L|\theta, N) \frac{dV}{dz} dzdL$$

$$\log(p(L_{obs}, z_{obs}|\theta, N)) = n \log(N) + \sum_{i \in A_{obs}} \log(p(L_i, z_i|\theta, N)) - Np(I = 1|\theta)$$

$$\text{όπου } p(I = 1|\theta) = \int \int p(I = 1|L, z) p(z, L|\theta, N) \frac{dV}{dz} dzdL$$

Η Schechter συνάρτηση φωτεινότητας

Θεωρούμε την ακόλουθη συνάρτηση φωτεινότητας:

$$\phi(L) = \frac{N}{L^* \Gamma(a+1)} \left(\frac{L}{L^*} \right)^a e^{-L/L^*}, \quad \theta = (a, L^*)$$

Η συνάρτηση αυτή είναι ανεξάρτητη της ερυθρομετατόπισης z και ονομάζεται συνάρτηση Schechter. Η συνάρτηση αυτή είναι ανάλογη με τη συνάρτηση πυκνότητας πιθανότητας μιας Γάμμα κατανομής με παραμέτρους $k = a+1$ και L^*

Η πιθανοφάνεια του μοντέλου για την Schechter συνάρτηση φωτεινότητας είναι:

$$p(L_{obs}, z_{obs} | \theta, N) = N^n \left[\prod_{i \in A_{obs}} p(L_i, z_i | \theta) \right] e^{-[N(p(I=1|\theta))]}$$

όπου $p(I=1|\theta) = \int_{L_{min}}^{\infty} \frac{1}{L^* \Gamma(k)} \left(\frac{L}{L^*} \right)^{k-1} e^{-L/L^*} = 1 - F_{(k, L^*)}(L_{min})$ και F : Η συνάρτηση κατανομή πιθανότητας της Γάμμα κατανομής με παραμέτρους (k, L^*) .

Για να απλοποιήσουμε το πρόβλημα, θεωρούμε ότι η συνάρτηση φωτεινότητας είναι ανεξάρτητη της ερυθρομετατόπισης (z) και ότι $k = \alpha + 1$. Η Schechter συνάρτηση φωτεινότητας είναι ένα απλουστευμένο μοντέλο για να περιγράψει τη φωτεινότητα. Χρησιμοποιούμε αυτό το μοντέλο θέλοντας να εφαρμόσουμε τις ιδέες μας σε μία απλή συνάρτηση και έπειτα να υποθέσουμε κάποιο πιο περίπλοκο μοντέλο.

Συνεπώς η από κοινού εκ των υστέρων κατανομή για τις παραμέτρους N, θ θα είναι:

$$p(\theta, N | L_{obs}, z_{obs}) \propto p(\theta, N) p(L_{obs}, z_{obs} | \theta, N) \propto p(\theta) p(N) p(L_{obs}, z_{obs} | \theta, N)$$

αν θεωρήσουμε ότι τα N, θ είναι ανεξάρτητα.

Η περιθώρια εκ των υστέρων κατανομή του θ μπορεί να βρεθεί εάν αθροίσουμε ως προς N και τότε θα ισχύει ότι:

$$p(\theta | L_{obs}, z_{obs}) \propto p(\theta) \prod_{i \in A_{obs}} p(L_i, z_i | \theta) \sum_{N=n}^{\infty} p(N) N^n e^{-[Np(I=1|\theta)]}$$

Η δεσμευμένη εκ των υστέρων κατανομή του θ θα είναι:

$$p(\theta|N, n) \propto p(\theta) \prod_{i \in A_{obs}} p(L_i, z_i|\theta) e^{-Np(I=1|\theta)}$$

Αντίστοιχα, η δεσμευμένη εκ των υστέρων κατανομή για το N θα είναι:

$$p(N|n, \theta) \propto p(N) (N^n e^{[-Np(I=1|\theta)]})$$

Θέλουμε να εκτιμήσουμε τις παραμέτρους της εκ των υστέρων κατανομής. Αν και η Schechter συνάρτηση είναι απλή παρατηρούμε ότι η εκ των υστέρων συνάρτηση είναι περίπλοκη. Ένα από τα πλεονεκτήματα της μπεϋζιανής στατιστικής είναι ότι με τη βοήθεια της μεθόδου MCMC μπορούμε να εκτιμήσουμε τις παραμέτρους ακόμα και σε περίπλοκες συναρτήσεις. Εφαρμόζουμε λοιπόν τον αλγόριθμο Metropolis-Hastings. για να προσομοιώσουμε από τις παραπάνω εκ των υστέρων δεσμευμένες κατανομές.

Αναφέρουμε σε αυτή την παράγραφο συνοπτικά τον αλγόριθμο που χρησιμοποιήσαμε.

- Ο αλγόριθμος αρχίζει με αρχικές τιμές των $(\tilde{k}, \tilde{L}^*, \tilde{N})$. Μια καλή αρχική τιμή είναι η εκτιμήτρια μέγιστης πιθανοφάνειας για τις παραμέτρους. Εμείς πήραμε αυθαίρετες τιμές καθώς είναι δύσκολο να υπολογιστούν αναλυτικά οι εκτιμήτριες μέγιστης πιθανοφάνειας.
- Προσομοιώνουμε προτεινόμενη τιμή του $\log(\hat{L}^*)$ από την κανονική κατανομή $N(\log(\tilde{L}^*), u_1)$.
- Υπολογίζουμε την πιθανότητα αποδοχής της προτεινόμενης τιμής:

$$\begin{aligned}
p_{acc1} &= \min \left[1, \frac{p(\tilde{k}, \hat{L}, \tilde{N} | L_{obs}) \hat{L}^*}{p(\tilde{k}, \tilde{L}^*, \tilde{N} | L_{obs}) \tilde{L}^*} \right] \\
&= \frac{\hat{L}^* \frac{1}{\tilde{L}^*} \left[\prod_{i=1}^n \frac{1}{\tilde{L}^* \Gamma(\tilde{k})} \left(\frac{L_i}{\tilde{L}^*} \right)^{(\tilde{k}-1)} e^{-[L_i/\tilde{L}^*]} \right] e^{-[\tilde{N}(1-F_{(\tilde{k}, \hat{L}^*)}(L_{min}))]}}{\tilde{L}^* \frac{1}{\tilde{L}^*} \left[\prod_{i=1}^n \frac{1}{\tilde{L}^* \Gamma(\tilde{k})} \left(\frac{L_i}{\tilde{L}^*} \right)^{(\tilde{k}-1)} e^{-[L_i/\tilde{L}^*]} \right] e^{-[\tilde{N}(1-F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))]}} \\
&= \left(\frac{\tilde{L}^*}{\hat{L}^*} \right)^{(n\tilde{k})} \left[\prod_{i=1}^n e^{-[L_i(\frac{1}{\tilde{L}^*} - \frac{1}{\hat{L}^*})]} \right] e^{[\tilde{N}(F_{(\tilde{k}, \hat{L}^*)}(L_{min}) - F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))]}
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\begin{aligned}
\log(p_{acc1}) &= \\
nk \log \left(\frac{\tilde{L}^*}{\hat{L}^*} \right) - \sum_{i=1}^n [L_i(\frac{1}{\tilde{L}^*} - \frac{1}{\hat{L}^*})] + \tilde{N}(F_{(\tilde{k}, \hat{L}^*)}(L_{min}) - F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))
\end{aligned}$$

- Ακολουθώντας προσομοιώνουμε έναν τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο αριθμός αυτό είναι μικρότερος από την πιθανότητα αποδοχής που μόλις υπολογίσαμε τότε δεχόμαστε την προτεινόμενη τιμή και τώρα η καινούργια τιμή του L^* θα είναι η προτεινόμενη. Σε αντίθετη περίπτωση η τιμή του L^* παραρμένει ως έχει.

Σημείωση: Στις πιθανότητες αποδοχής υπάρχει ο όρος $\frac{\hat{L}^*}{\tilde{L}^*}$. Αυτό προκύπτει από την αλλαγή μεταβλητής για την προτεινόμενη τιμή (από την παράγωγο $\frac{dLog L^*}{dL^*}$). Το L^* ακολουθεί τη log-normal κατανομή και η $\log(L^*)$ ακολουθεί την κανονική κατανομή.

- Προσομοιώνουμε προτεινόμενη τιμή για το $\log(\hat{k}) = \log(\hat{a} + 1)$ από την κανονική κατανομή $N(\log(\tilde{k}), u_2)$.
- Υπολογίζουμε και πάλι την πιθανότητα αποδοχής της προτεινόμενης τιμής:

$$\begin{aligned}
p_{acc2} &= \frac{\hat{k} \frac{1}{\tilde{L}^*} \left[\prod_{i=1}^n \frac{1}{\tilde{L}^* \Gamma(\hat{k})} \left(\frac{L_i}{\tilde{L}^*} \right)^{(\hat{k}-1)} e^{-[L_i/\tilde{L}^*]} \right] e^{-[N(1-F_{(\hat{k}, \tilde{L}^*)}(L_{min}))]}}{\tilde{k} \frac{1}{\tilde{L}^*} \left[\prod_{i=1}^n \frac{1}{\tilde{L}^* \Gamma(\tilde{k})} \left(\frac{L_i}{\tilde{L}^*} \right)^{(\tilde{k}-1)} e^{-[L_i/\tilde{L}^*]} \right] e^{-[N(1-F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))]}} \\
&= \frac{\hat{k}}{\tilde{k}} \left[\frac{\Gamma(\tilde{k})}{\Gamma(\hat{k})} \right]^n \left[\prod_{i=1}^n \left(\frac{L_i}{\tilde{L}^*} \right)^{(\hat{k}-\tilde{k})} \right] e^{[N(F_{(\hat{k}, \tilde{L})}(L_{min}) - F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))]}
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\log(p_{acc2}) = \log(\hat{k}) - \log(\tilde{k}) + n(\log(\Gamma(\tilde{k})) - \log(\Gamma(\hat{k}))) + N(F_{(\hat{k}, \tilde{L}^*)}(L_{min}) - F_{(\tilde{k}, \tilde{L})}(L_{min})) + (\hat{k} - \tilde{k}) \sum_{i=1}^n \log\left(\frac{L_i}{\tilde{L}^*}\right)$$

- Προσομοιώνουμε έναν τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο αριθμός αυτό είναι μικρότερος από την πιθανότητα αποδοχής που μόλις υπολογίσαμε τότε δεχόμαστε την προτεινόμενη τιμή και τώρα η καινούργια τιμή του k θα είναι η προτεινόμενη. Σε αντίθετη περίπτωση η τιμή του k παραρμένει ως έχει.
- Προτείνουμε τιμή για το N από την αρνητική διωνυμική κατανομή με παραμέτρους n , $p(I = 1|\theta)$
Η πιθανότητα αποδοχής της προτεινόμενης τιμής θα είναι:

$$\begin{aligned}
p_{acc3} &= \min \left[1, \frac{p(\hat{N}|L_{obs}, k, L^*) q(\tilde{N}|L_{obs}, k, L^*)}{p(\tilde{N}|L_{obs}, k, L^*) q(\hat{N}|L_{obs}, k, L^*)} \right] \\
&= \frac{\hat{N}^{n-1} \exp\{-\hat{N}p(I = 1|\theta)\} \binom{\tilde{N}-1}{n-1} p(I = 1|\theta)^n p(I = 0|\theta)^{\tilde{N}-1}}{\tilde{N}^{n-1} \exp\{-\tilde{N}p(I = 1|\theta)\} \binom{\hat{N}-1}{n-1} p(I = 1|\theta)^n p(I = 0|\theta)^{\hat{N}-1}} \\
&= \left(\frac{\hat{N}}{\tilde{N}} \right)^{n-1} \exp\{(\hat{N} - \tilde{N})p(I = 1|\theta)\} \frac{(\tilde{N} - 1)! (\hat{N} - n)!}{(\hat{N} - 1)! (\tilde{N} - n)!} [p(I = 0|\theta)]^{\tilde{N} - \hat{N}}
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\begin{aligned}
\log p_{acc3} &= (n-1) [\log(\hat{N}) - \log(\tilde{N})] + (\tilde{N} - \hat{N})p(I = 1|\theta) + \sum_{i=1}^{\tilde{N}-1} \log(i) - \\
&\quad \sum_{i=1}^{\hat{N}-1} \log(i) + \sum_{i=1}^{\tilde{N}-n} \log(i) - \sum_{i=1}^{\hat{N}-n} \log(i) + (\tilde{N} - \hat{N}) \log(p(I = 0|\theta))
\end{aligned}$$

- Προσομοιώνουμε έναν τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο αριθμός αυτό είναι μικρότερος από την πιθανότητα αποδοχής που μόλις υπολογίσαμε τότε δεχόμαστε την προτεινόμενη τιμή και τώρα η καινούργια τιμή του N θα είναι η προτεινόμενη. Σε αντίθετη περίπτωση η τιμή του N παραμένει ως έχει.

Όταν προσομοιώνουμε τιμές για τον αλγόριθμο Metropolis-Hastings τυχαίου περιπάτου (για την ανανέωση του $\log(\hat{L}^*)$) και του $\log(k)$ παίρνουμε διαπορές u_1, u_2 τέτοιες ώστε το ποσοστό αποδοχής των προτεινόμενων τιμών να είναι μεταξύ 20 – 50%. Το N προσομοιώνεται με ανεξάρτητη γεννήτρια προτεινόμενων τιμών και άρα στόχος είναι η μεγιστοποίηση της πιθανότητας αποδοχής.

4.2.3 Διωνυμικά δεδομένα

Ας εξετάσουμε τώρα την περίπτωση του διωνυμικού μοντέλου.

Ορίζουμε διάνυσμα μεγέθους N που περιέχει σε κάθε συνιστώσα τη δείκτρια συνάρτηση που παίρνει την τιμή 1 εάν η i -οστή πηγή περιέχεται στο δείγμα και μηδέν αλλιώς. Έστω A_{obs} , A_{mis} τα σύνολα που βρίσκονται οι πηγές που έχουν παρατηρηθεί και οι πηγές που δεν παρατηρήθηκαν αντίστοιχα. Σε αυτή την περίπτωση η συνάρτηση πιθανοφάνειας θα είναι

$$p(L, z, I|\theta, N) = C_n^N \prod_{i \in A_{obs}} p(I = 1|L_i, z_i)p(L_i, z_i) \times \prod_{j \in A_{mis}} p(I = 0|L_j, z_j)p(L_j, z_j)$$

όπου $C_n^N = \binom{N}{n}$

Μας ενδιαφέρει η πιθανότητα των παρατηρούμενων δεδομένων, κάτω από την υπόθεση του συγκεκριμένου μοντέλου που έχουμε επιλέξει. Επομένως μας ενδιαφέρει η περιθώρια πιθανοφάνεια των δεδομένων και αυτή μπορεί να βρεθεί ολοκληρώνοντας ως προς την αλλη παράμετρο (μη παρατηρούμενα δεδομένα) την από κοινού πιθανοφάνεια. Έτσι οι υπολογισμοί έχουν ως εξής:

$$p(L_{obs}, z_{obs}, I|\theta, N) = C_n^N \prod_{i \in A_{obs}} p(I = 1|L_i, z_i)p(L_i, z_i|\theta) \times \prod_{j \in A_{mis}} \int_0^\infty \int_0^\infty p(I = 0|L_j, z_j)p(L_j, z_j|\theta) dL_j dz_j$$

Προφανώς η πιθανότητα:

$$p(I = 1|L_i, z_i) = 1 \text{ αν } i \in A_{obs}$$

και αν συμβολίσουμε με

$$p(I = 0|\theta) = \int \int p(I = 1|L, z)p(L, z|\theta)dLdz$$

παίρνουμε ότι:

$$p(L_{obs}, z_{obs}, I|\theta, N) \propto C_n^N [p(I = 0|\theta)]^{N-n} \prod_{i \in A_{obs}} p(L_i, z_i|\theta)$$

Υπολογισμός της εκ των υστέρων κατανομής

Υποθέτοντας ότι οι παράμετροι (θ, N) είναι μεταξύ τους ανεξάρτητες, η από κοινού εκ των προτέρων κατανομή τους θα δίνεται από τη σχέση:

$$p(\theta, N) = p(\theta)p(N)$$

Η εκ των υστέρων κατανομή πιθανότητας των δύο παραμέτρων θα είναι:

$$p(\theta, N|L_{obs}, z_{obs}, I) \propto p(\theta)p(N)p(L_{obs}, z_{obs}, I|\theta, N)$$

Ο συνολικός αριθμός των πηγών N είναι άγνωστος. Υποθέτουμε ομοιόμορφη κατανομή για τον λογάριθμο του N . Προτείνουμε αυτή την εκ των προτέρων κατανομή γιατί θέλουμε μια κατανομή που να μας διευκολύνει στους υπολογισμούς και να είναι μη πληροφοριακή.

$$p(N) \propto \frac{1}{N} \Rightarrow p(\log(N))d\log(N) = p(N)dN$$

Έτσι κάτω από την προαναφερθήσα εκ των προτέρων κατανομή η περιθώρια εκ των υστέρων κατανομή πιθανότητας του θ θα είναι:

$$\begin{aligned} p(\theta|L_{obs}, z_{obs}) &\propto p(\theta) \left[\prod_{i \in A_{obs}} p(L_i, z_i|\theta) \right] \sum_{N=n}^{\infty} N^{-1} C_n^N [p(I = 0|\theta)]^{N-n} \\ &\propto p(\theta) [p(I = 1|\theta)]^{-n} \left[\prod_{i \in A_{obs}} p(L_i, z_i|\theta) \right] \\ &\times \sum_{N=n}^{\infty} \frac{1}{N} C_{n-1}^{N-1} [p(I = 0|\theta)]^{N-n} [p(I = 1|\theta)]^n \end{aligned}$$

Το τελευταίο άθροισμα είναι άθροισμα της Αρνητικής διωνυμικής κατανομής

για όλο το πεδίο ορισμού του και ισούται με μονάδα. Συνεπώς η εκ των υστέρων κατανομή είναι:

$$p(\theta|L_{obs}, z_{obs}) \propto p(\theta) [p(I = 1|\theta)]^{-n} \prod_{i \in A_{obs}} p(L_i, z_i|\theta)$$

Η δεσμευμένη εκ των υστέρων κατανομή του N θα είναι αρνητική διωνυμική κατανομή με παραμέτρους $n, p(I = 1|\theta)$ αφού το N θα είναι ο αριθμός των δοκιμών που χρειάζεται για να έχουμε n επιτυχίες με πιθανότητα επιτυχίας $p(I = 1|\theta)$ (αυτός είναι και ο ορισμός της αρνητικής διωνυμικής κατανομής).

H Schechter συνάρτηση φωτεινότητας

Έστω ότι η συνάρτηση φωτεινότητας έχει τη μορφή

$$\phi(L) = \frac{N}{L^* \Gamma(a+1)} \left(\frac{L}{L^*}\right)^a e^{-L/L^*}, \quad \theta = (a, L^*)$$

Στην ανάλυση που προηγήθηκε καταλήξαμε στον τύπο για την εκ των υστέρων κατανομή

$$p(\theta|L_{obs}, z_{obs}) \propto p(\theta) [p(I = 1|\theta)]^{-n} \prod_{i \in A_{obs}} p(L_i, z_i|\theta)$$

Αντικαθιστούμε λοιπόν την πυκνότητα πιθανότητας της Schechter συνάρτησης και παίρνουμε την εκ των υστέρων κατανομή πιθανότητας:

$$p(k, L^*) \propto \frac{1}{L^*} [p(I = 1|k, L^*)]^{-n} \prod_{i=1}^N \frac{1}{L^* \Gamma(k)} \left(\frac{L_i}{L^*}\right)^{k-1} e^{-L_i/L^*}$$

$$\text{όπου, } [p(I = 1|k, L^*)] = \int_{L_{min}}^{\infty} \frac{1}{L^* \Gamma(k)} \left(\frac{L}{L^*}\right)^{k-1} e^{-L/L^*} = 1 - F_{(k, L^*)}(L_{min})$$

F : Η συνάρτηση κατανομή πιθανότητας της γάμμα κατανομής με παραμέτρους (k, L^*) .

Η δεσμευμένη εκ των υστέρων κατανομή του N είναι αρνητική διωνυμική κατανομή.

$$p(N|n, \theta) = \binom{N-1}{n-1} [p(I = 1|\theta)]^n [p(I = 0|\theta)]^{N-n}$$

Φτιάξαμε λοιπόν αλγόριθμο για την εκτίμηση της συνάρτησης φωτεινότητας όταν τα δεδομένα είναι διωνυμικά.

- Ο αλγόριθμος αρχίζει με αρχικές τιμές των $(\tilde{k}, \tilde{L}^*)$.
- Προσομοιώνουμε προτεινόμενη τιμή του $\log(\hat{L}^*)$ από την κανονική κατανομή $N(\log(\tilde{L}^*), u_1)$.
Εδώ σημειώνεται ότι η διασπορά της κανονικής κατανομής είναι τέτοια ώστε το ποσοστό αποδοχής της προτεινόμενης τιμής να είναι μεταξύ 20 – 50%. Προτείνεται η κατανομή της γεννήτριας τιμών να έχει μέσο το $\log(L^*)$ γιατί ο λογάριθμος ορίζεται σε όλο το διάστημα των πραγματικών αριθμών ενώ το L^* ορίζεται μόνο στους θετικούς πραγματικούς αριθμούς.
- Υπολογίζουμε την πιθανότητα αποδοχής της προτεινόμενης τιμής:

$$\begin{aligned}
p_{acc1} &= \min \left[1, \frac{p(\tilde{k}, \hat{L} | L_{obs}) \hat{L}^*}{p(\tilde{k}, \tilde{L}^* | L_{obs}) \tilde{L}^*} \right] \\
&= \frac{\hat{L}^* \frac{1}{\hat{L}^*} (1 - F_{(\tilde{k}, \hat{L}^*)}(L_{min}))^{-n}}{\tilde{L}^* \frac{1}{\tilde{L}^*} (1 - F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))^{-n}} \prod_{i=1}^n \frac{\frac{1}{\tilde{L}^* \Gamma(k)} (\frac{L_i}{\tilde{L}^*})^{\tilde{k}-1} e^{-(L_i/\tilde{L}^*)}}{\frac{1}{\hat{L}^* \Gamma(k)} (\frac{L_i}{\hat{L}^*})^{\hat{k}-1} e^{-(L_i/\hat{L}^*)}} \\
&= \left(\frac{1 - F_{(\tilde{k}, \tilde{L}^*)}(L_{min})}{1 - F_{(\tilde{k}, \hat{L}^*)}(L_{min})} \right)^n \left(\frac{\tilde{L}^*}{\hat{L}^*} \right)^{n\tilde{k}} \exp \left[- \sum_{i=1}^n L_i \left(\frac{1}{\hat{L}^*} - \frac{1}{\tilde{L}^*} \right) \right]
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\log(p_{acc1}) = n(\log(1 - F_{(\tilde{k}, \tilde{L}^*)}(L_{min})) - \log(1 - F_{(\tilde{k}, \hat{L}^*)}(L_{min}))) + n\tilde{k}(\log(\tilde{L}^*) - \log(\hat{L}^*)) - \sum_{i=1}^n L_i \left(\frac{1}{\hat{L}^*} - \frac{1}{\tilde{L}^*} \right)$$

- Ακολούθως προσομοιώνουμε ένα τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο αριθμός αυτό είναι μικρότερος από την πιθανότητα αποδοχής που μόλις υπολογίσαμε τότε δεχόμαστε την προτεινόμενη τιμή και τώρα η καινούργια τιμή του L^* θα είναι η προτεινόμενη. Σε αντίθετη περίπτωση η τιμή του L^* παραμένει ως έχει.
- Προσομοιώνουμε προτεινόμενη τιμή για το $\log(\hat{k}) = \log(\hat{a} + 1)$ από την κανονική κατανομή $N(\log(\tilde{k}), u_2)$.

- Υπολογίζουμε και πάλι την πιθανότητα αποδοχής της προτεινόμενης τιμής:

$$\begin{aligned}
p_{acc2} &= \min \left[1, \frac{p(\hat{k}\tilde{L}|L_{obs})}{p(\tilde{k}\tilde{L}^*|L_{obs})} \frac{\hat{k}}{\tilde{k}} \right] \\
&= \frac{\hat{k} \frac{1}{\tilde{L}^*} (1 - F_{(\hat{k}, \tilde{L}^*)}(L_{min}))^{-n}}{\tilde{k} \frac{1}{\tilde{L}^*} (1 - F_{(\tilde{k}, \tilde{L}^*)}(L_{min}))^{-n}} \prod_{i=1}^n \frac{\frac{1}{\tilde{L}^* \Gamma(\hat{k})} \left(\frac{L_i}{\tilde{L}^*}\right)^{\hat{k}-1} e^{-L_i/\tilde{L}^*}}{\frac{1}{\tilde{L}^* \Gamma(\tilde{k})} \left(\frac{L_i}{\tilde{L}^*}\right)^{\tilde{k}-1} e^{-L_i/\tilde{L}^*}} \\
&= \frac{\hat{k}}{\tilde{k}} \left(\frac{\Gamma(\tilde{k})}{\Gamma(\hat{k})} \right)^n \left[\frac{1 - F_{(\tilde{k}, \tilde{L}^*)}(L_{min})}{1 - F_{(\hat{k}, \tilde{L}^*)}(L_{min})} \right]^n \prod_{i=1}^n \left(\frac{L_i}{\tilde{L}^*} \right)^{\hat{k}-\tilde{k}}
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\begin{aligned}
\log(p_{acc2}) &= \log(\hat{k}) - \log(\tilde{k}) + n(\log(\Gamma(\tilde{k})) - \log(\Gamma(\hat{k}))) + \\
&+ n \left[\log(1 - F_{(\tilde{k}, \tilde{L}^*)}(L_{min})) - \log(1 - F_{(\hat{k}, \tilde{L}^*)}(L_{min})) \right] + (\hat{k} - \tilde{k}) \sum_{i=1}^n \log\left(\frac{L_i}{\tilde{L}^*}\right)
\end{aligned}$$

- Προσομοιώνουμε ένα τυχαίο αριθμό από την ομοιόμορφη κατανομή στο $(0, 1)$ και όπως και στο προηγούμενο βήμα, εάν αυτός ο αριθμός είναι μικρότερος από την πιθανότητα αποδοχής που υπολογίστηκε πιο πριν τότε δεχόμαστε την προτεινόμενη τιμή και η νέα τιμή για την παράμετρο μας γίνεται η προτεινόμενη. Σε αντίθετη περίπτωση η τιμή μένει σταθερή.
- Τα πιο πάνω βήματα τα τρέχουμε μέχρι να επιτευχθεί σύγκλιση του αλγορίθμου. Αφού πετύχουμε σύγκλιση προσθέτουμε και τα επόμενα βήματα που θα μας δίνουν σε κάθε στάδιο και τιμή για το N από την αρνητική διωνυμική κατανομή ώστε να πάρουμε μια τριάδα των παραμέτρων (k, L^*, N) . Αξίζει να αναφέρουμε ότι στην πράξη χρησιμοποιούμε πολλές φορές το λογάριθμο των πιθανοτήτων αποδοχής για καλύτερους υπολογισμούς και μεγαλύτερη ακρίβεια στον υπολογιστή.
- Προσομοιώνουμε n τυχαίους αριθμούς από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Υπολογίζουμε τον αριθμό m ως εξής:
$$m = \sum_{i=1}^n \left\lfloor \frac{\log u_i}{\log(1-p)} \right\rfloor, \text{ όπου } p = 1 - F_{(k, L^*)}(L_{min})$$

Η τυχαία μεταβλητή $N = n + m$ ακολουθεί την αρνητική διωνυμική κατανομή με τις επιθυμητές παραμέτρους.

- Τρέχουμε τον αλγόριθμο όσες φορές χρειαζόμαστε για να πάρουμε ένα ικανοποιητικό μέγεθος δείγματος για τη στατιστική συμπερασματολογία.

4.2.4 Από κοινού μοντελοποίηση φωτεινότητας και ερυθρομετατόπησης με μίξεις κανονικών κατανομών

Η μίξη κανονικών κατανομών χρησιμοποιείται πολλές φορές όταν η ακριβής μορφή της κατανομής δεν μπορεί να προσδιοριστεί πλήρως και ειδικά όταν είναι πολύπλοκη, πολυκόρυφη ή ασύμμετρη. Στη μελέτη που ακολουθεί υποθέτουμε ότι ο λογάριθμος της φωτεινότητας και ο λογάριθμος της ερυθρομετατόπησης ακολουθούν από κοινού μία μίξη κανονικών κατανομών. Ισοδύναμα τα L, z ακολουθούν τη log-normal κατανομή. Ο λόγος που χρησιμοποιούμε το λογάριθμο των L, z είναι ότι θέλουμε να παίρνουμε τιμές στους θετικούς πραγματικούς αριθμούς μόνο.

Η μίξη των K κανονικών κατανομών θα είναι:

$$p(\log(L_i), \log(z_i) | \pi, \mu, \Sigma) = \sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2}(x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right]$$

όπου $\theta = (\pi, \mu, \Sigma)$, $\sum_{k=1}^K \pi_k = 1$, $x_i = (\log(L_i), \log(z_i))$, $\pi = (\pi_1, \dots, \pi_K)$, $\mu = (\mu_1, \dots, \mu_K)$, $\Sigma = (\Sigma_1, \dots, \Sigma_K)$

Ο πίνακας διασποράς για την k -οστή συνιστώσα κανονικών κατανομών έχει στη διαγώνιο τις διασπορές των L, z και τις συνδιακυμάνσεις στα μη διαγώνια στοιχεία. Λόγω του μετασχηματισμού που εφαρμόσαμε στα L, z η συνάρτηση φωτεινότητας και συνεπώς η συνάρτηση πυκνότητας πιθανότητας των πηγών θα δίνονται από τη σχέση:

$$\phi(L, z | \theta, N) = N p(L, z) \left(\frac{dV}{dz} \right)^{-1} \Rightarrow p(L, z | \theta, N) = p(\log_{10}(L), \log_{10}(z)) \frac{1}{Lz} \frac{1}{(\log(10))^2}$$

(από την αλλαγή των μεταβλητών)

Η συνάρτηση φωτεινότητας θα είναι:

$$\phi(L, z | \theta, N) = \frac{N}{Lz(\log(10))^2} \left(\frac{dV}{dz} \right)^{-1} \sum_{k=1}^K \left[\frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \right] \right]$$

Καθορισμός των εκ των προτέρων κατανομών

Οι γνώσεις μας για τις παράμετρους π, μ, Σ είναι περιορισμένες και επομένως θέλουμε να χρησιμοποιήσουμε μη πληροφοριακές εκ των προτέρων κατανομές. Μία ιδέα θα ήταν να χρησιμοποιούσαμε ομοιόμορφες εκ των προτέρων κατανομές. Όμως αυτή η ιδέα δεν μπορεί να εφαρμοστεί καθώς η εκ των υστέρων κατανομή που παίρνουμε μετά από την ανάλυση μας είναι ακατάλληλη (η από κοινού εκ των υστέρων κατανομή όταν ολοκληρωθεί πρέπει να έχει ολοκλήρωμα ίσο με τη μονάδα).

- Για τα $\pi_1, \dots, \pi_k$ υποθέτουμε ομοιόμορφη κατανομή υπό τον περιορισμό ότι $\sum_{i=1}^K \pi_i = 1$ ισοδύναμα η εκ των προτέρων κατανομή θα είναι $\text{Dirichlet}(1, \dots, 1)$.
- Για τα μ, Σ έχουμε τα εξής:

Υποθέτουμε ότι η συνάρτηση φωτεινότητας έχει μόνο ένα σημείο κορυφής. Για να μπορεί να συμβεί αυτό δίνουμε περισσότερο βάρος πιθανότητας στις κανονικές κατανομές που βρίσκονται κοντά μεταξύ τους (κανονικοποιούμε τις κατανομές με τον πίνακα συνδιακύμανσης).

- Για τα $\mu_1, \dots, \mu_k$ υποθέτουμε ότι είναι ανεξάρτητα και ακολουθούν κατανομή $Cauchy_2$ με παράμετρους μ_0 (διδιάστατο) και πίνακα κλίμακας T . Τα $\mu_1, \dots, \mu_k$ και μ_0 είναι περιορισμένα μεταξύ κάποιων ανώτατων και κατώτατων ορίων για τη φωτεινότητα και την ερυθρομετατόπιση. Ο πίνακας T επιλέγουμε να είναι ο αρμονικός μέσος των πινάκων συνδιακύμανσης των κανονικών κατανομών. Δηλαδή θα είναι τέτοιος ώστε:

$$T = \left(\frac{1}{K} \sum_{k=1}^K \Sigma_k^{-1} \right)^{-1}$$

- Για τον πίνακα διακύμανσης Σ_k υποθέτουμε ότι έχει εκ των προτέρων κατανομή Inverse Wishart με πίνακα A και ένα βαθμό ελευθερίας. Αφήνουμε τον πίνακα A να καθοριστεί από τα δεδομένα.

Συνεπώς η εκ των προτέρων κατανομή του θ θα είναι:

$$\begin{aligned} p(\pi, \mu, \Sigma, \mu_0, A) &\propto \prod_{k=1}^K p(\mu_k | \mu_0, \Sigma) p(\Sigma_k, A) \\ &\propto \prod_{k=1}^K \text{Cauchy}_2(\mu_k | \mu_0, T) \text{InvWishart}(\Sigma_k | A) \end{aligned}$$

Υπολογισμός της εκ των υστέρων κατανομής

Poisson δεδομένα

Έχουμε δει ότι για Poisson δεδομένα η συνάρτηση πιθανοφάνειας είναι

$$p(L_{obs}, z_{obs} | \theta, N) = N^n [\prod_{i \in A_{obs}} p(L_i, z_i | \theta)] e^{-[N(p(I=1|\theta))]}$$

Η εκ των υστέρων κατανομή είναι:

$$\begin{aligned} p(\theta, \mu_0, A, N | \log(L_{obs}), \log(z_{obs})) &\propto p(N | n, \theta) p(\theta, \mu_0, A) \prod_{i \in A_{obs}} [p(\log(L_i), \log(z_i)) | \theta] e^{-Np(I=1|\theta)} \\ &\propto p(N | n, \theta) \left[\prod_{k=1}^K \text{Cauchy}_2(\mu_k | \mu_0, T) \text{InvWishart}(\Sigma_k | A) \right] \\ &\times \prod_{i \in A_{obs}} \sum_{k=1}^K \frac{\pi_k}{2\pi |\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right] [e^{-Np(I=1|\theta)}] \\ &\propto p(N | n, \theta) \prod_{k=1}^K \left[\frac{\Gamma(3/2)}{\Gamma(1/2)\pi} |T|^{-1/2} [1 + (\mu_k - \mu)^t |T|^{-1} (\mu_k - \mu)]^{-3/2} \right] \\ &\times \prod_{k=1}^K \left[\left[2\pi \prod_{i=1}^2 \Gamma\left(\frac{2-i}{2}\right) \right]^{-1} |A|^{1/2} |\Sigma_k|^{-2} \exp \left\{ -\frac{1}{2} (\Sigma_k^{-1} A) \right\} \right] \\ &\times \prod_{i \in A_{obs}} \sum_{k=1}^K \left[\frac{\pi_k}{|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x - \hat{\mu}_k)^T \Sigma_k^{-1} (x - \hat{\mu}_k) \right] \right] \end{aligned}$$

Η δεσμευμένη εκ των υστέρων κατανομή του N θα είναι:

$$p(N | n, \theta) \propto p(N) (N^n e^{-[Np(I=1|\theta)]})$$

όπου $p(I = 1 | \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(I = 1 | \log L, \log z) p(\log L, \log z | \theta) d\log L d\log z$
η πιθανότητα να συμπεριληφθεί πηγή στο δείγμα δεδομένου του μοντέλου που υποθέσαμε.

Στην έρευνα που πραγματοποιούμε θα έχουμε ότι οι πηγές ανιχνεύονται όταν αυτές έχουν φωτεινότητα και ερυθρομετατόπιση μεταξύ κάποιων κατώτατων και ανώτατων ορίων. Λαμβάνοντας αυτό υπόψη και χρησιμοποιώντας τους συμβολισμούς

$$x_i = (\log_{10}(L_i), \log_{10}(z_i))', \quad x = (\log_{10}(N), \log_{10}(z))' \text{ και} \\ x_{high} = (\log_{10}(L_{high}), \log_{10}(z_{high}))', \quad x_{low} = (\log_{10}(L_{low}), \log_{10}(z_{low}))'$$

τότε η πιθανότητα να συμπεριληφθεί πηγή στο δείγμα θα δίνεται από:

$$p(I = 1|\theta) = \int_{x_{low}}^{x_{high}} p(x)dx = \sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)]$$

Με F συμβολίζουμε την αθροιστική συνάρτηση κατανομής της πολυπαραμετρικής Κανονικής κατανομής.

Διωνυμικά δεδομένα

Το N ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους $n, p(I = 1|\theta)$

$$p(N|n, \theta) = \binom{N-1}{n-1} [p(I = 1|\theta)]^n [p(I = 0|\theta)]^{N-n}$$

Αν λοιπόν λάβουμε υπόψη μας την πιθανοφάνεια του μοντέλου κάτω από διωνυμικά δεδομένα και τις εκ των προτέρων κατανομές που υποθέσαμε, η εκ των υστέρων κατανομή για διωνυμικά δεδομένα θα είναι:

$$p(\theta, N, \mu_0, A | \text{Log}L_{obs}, \text{Log}z_{obs}) \propto p(N|\theta, n) p(\theta, \mu_0, A | \text{Log}L_{obs}, \text{Log}z_{obs})$$

$$p(\theta, N, \mu_0, A | \text{Log}L_{obs}, \text{Log}z_{obs}) \propto \binom{N-1}{n-1} [p(I = 1|\theta)]^n [p(I = 0|\theta)]^{N-n} \times \\ \times \prod_{k=1}^K \left[\frac{\Gamma(3/2)}{\Gamma(1/2)\pi} |T|^{-1/2} [1 + (\mu_k - \mu)^t |T|^{-1} (\mu_k - \mu)]^{-3/2} \right] \times \\ \times \prod_{k=1}^K \left[\left[2\pi \prod_{i=1}^2 \Gamma\left(\frac{2-i}{2}\right) \right]^{-1} |A|^{1/2} |\Sigma_k|^{-2} \exp \left\{ -\frac{1}{2} (\Sigma_k^{-1} A) \right\} \right] \times \\ \times \prod_{i \in A_{obs}} \sum_{k=1}^K \left[\frac{\pi_k}{|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x - \hat{\mu}_k)^T \Sigma_k^{-1} (x - \hat{\mu}_k) \right] \right]$$

Λεπτομέρειες για τον αλγόριθμο MCMC δίνονται στο Παράρτημα III.

Κεφάλαιο 5

Εφαρμογή

Στο κεφάλαιο αυτό παρουσιάζουμε μια εφαρμογή των μεθόδων που παρουσιάστηκαν στο κεφάλαιο 4 σε προσομοιωμένα δεδομένα. Η ανάλυση μας ακολουθεί και προεκτείνει την εργασία των B.C.Kelly, X.Fan, M.Vestergaard, (2008).

Οι προσομοιώσεις έγιναν στο πρόγραμμα Matlab. Οι αλγόριθμοι που χρησιμοποιήθηκαν, τα βήματα των αλγορίθμων MCMC και οι πράξεις αναλυτικότερα βρίσκονται στο Παράρτημα II. Σε όλους τους αλγόριθμους δημιουργήσαμε τα αρχικά δεδομένα παίρνοντας δείγματα από τη γάμμα κατανομή με παραμέτρους $L^* = 10^{44}$, $\alpha = 0$ και διαφορετικά όρια για την φωτεινότητα κάτω από την οποία δεν μπορούμε να ανιχνεύσουμε πηγές. Δηλαδή κατά την προσομοίωση των δεδομένων αν είχαμε δεδομένα με φωτεινότητα κάτω από κάποιο L_{min} τότε δεν μπορούμε να τα ανιχνεύσουμε και τα απορρίπτουμε από το δείγμα. Κάναμε προσομοιώσεις με δύο διαφορετικά όρια $L_{min} = 2 \times 10^{43}$, $L_{min} = 2 \times 10^{44}$ και δύο διαφορετικά $N = 1000$, $N = 10000$. Άρα δημιουργήσαμε συνολικά τέσσερα σετ δεδομένων. Καθώς τα αρχικά μας δεδομένα δεν είναι πραγματικά αλλά προσομοιωμένα, γνωρίζουμε τις τιμές των παραμέτρων που θέλουμε να εκτιμήσουμε και συνεπώς θα μπορούμε να ελέγξουμε την ακρίβεια της συμπερασματολογίας.

5.1 Schechter συνάρτηση φωτεινότητας για Poisson δεδομένα

Δημιουργήσαμε έναν αλγόριθμο για να εκτιμήσουμε τις τρεις παραμέτρους που μας ενδιαφέρουν κάτω από την υπόθεση Poisson για τα δεδομένα μας. Εφαρμόσαμε αλγόριθμο Metropolis-Hastings και για τις τρεις παράμετρους. Για τα $\log(L^*)$, $\log(\alpha)$ εφαρμόσαμε βήματα τυχαίου περίπατου με πιθανότητες αποδοχής 20 – 50%, ενώ για το N προτείνουμε τιμές από την αρνητική διωνυμική κατανομή. Παρακάτω παρουσιάζουμε τα αποτελέσματα του αλγορίθμου μας.

Για $L_{min} = 10^{43}$ οι πραγματικές τιμές των παραμέτρων βρίσκονται μέσα σε όλα τα 95% διαστήματα αξιοπιστίας των παραμέτρων που θέλαμε να εκτιμήσουμε. Όταν το $L_{min} = 2 \times 10^{44}$ και $N = 1000$ σε 3 από τα 20 διαστήματα αξιοπιστίας που φτιάξαμε για το α δεν περιείχαν την πραγματική τιμή της παραμέτρου. Οι υπόλοιπες παράμετροι βρίσκονται και στα 20 διαστήματα αξιοπιστίας. Για $N = 10000$ είχαμε επίσης 1 από τα 20 διαστήματα αξιοπιστίας δεν περιείχε την πραγματική τιμή του α ενώ οι υπόλοιπες παράμετροι βρίσκονταν και στα 20 διαστήματα.

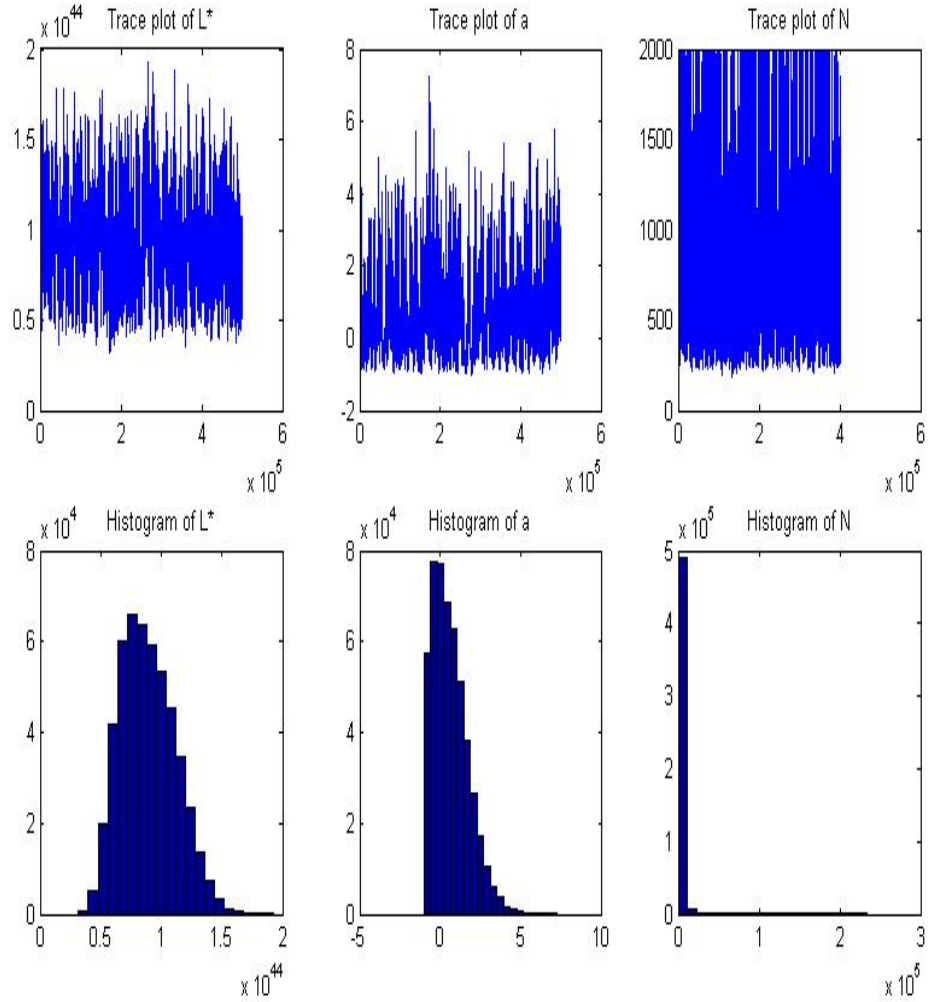
Θέλοντας να δούμε πόσο επηρεάζει η τιμή του N τα αποτελέσματα μας τρέξαμε τον αλγόριθμο για δύο διαφορετικές τιμές του N . Παρατηρήσαμε ότι ο αλγόριθμος με μεγαλύτερο N έδινε καλύτερες εκτιμήσεις. Κατά την άποψη μας ο λόγος που συμβαίνει αυτό είναι ότι για μεγάλα N τα δεδομένα μας προσεγγίζονται από την κατανομή Poisson. Ένας άλλος λόγος που μπορεί να συμβαίνει αυτό είναι ότι όταν έχουμε μεγαλύτερο N παίρνουμε μεγαλύτερο δείγμα κατά τη δημιουργία των αρχικών δεδομένων που χρειαζόμαστε για να τρέξουμε τον αλγόριθμο. Καθώς έχουμε μεγαλύτερο δείγμα αυτό θα συνεπάγεται ότι θα έχουμε και περισσότερη πληροφορία για τις παραμέτρους, άρα και η εκτίμηση για τη συνάρτηση κατανομή τους θα είναι πιο ακριβής. Επίσης για το μικρότερο όριο ελάχιστης φωτεινότητας L_{min} παίρνουμε καλύτερα αποτελέσματα και μικρότερα διαστήματα αξιοπιστίας και γρηγορότερη σύγκλιση. Για μικρότερο L_{min} σημαίνει ότι είμαστε λιγότερο αυστηροί όσον αφορά το ποιες πηγές παρατηρούμε. Θα πάρουμε συνεπώς περισσότερες πηγές και άρα και σε αυτή την περίπτωση θα έχουμε μεγαλύτερο δείγμα.

Εντούτοις, παρόλο που ο αλγόριθμος δίνει ικανοποιητικά αποτελέσματα για

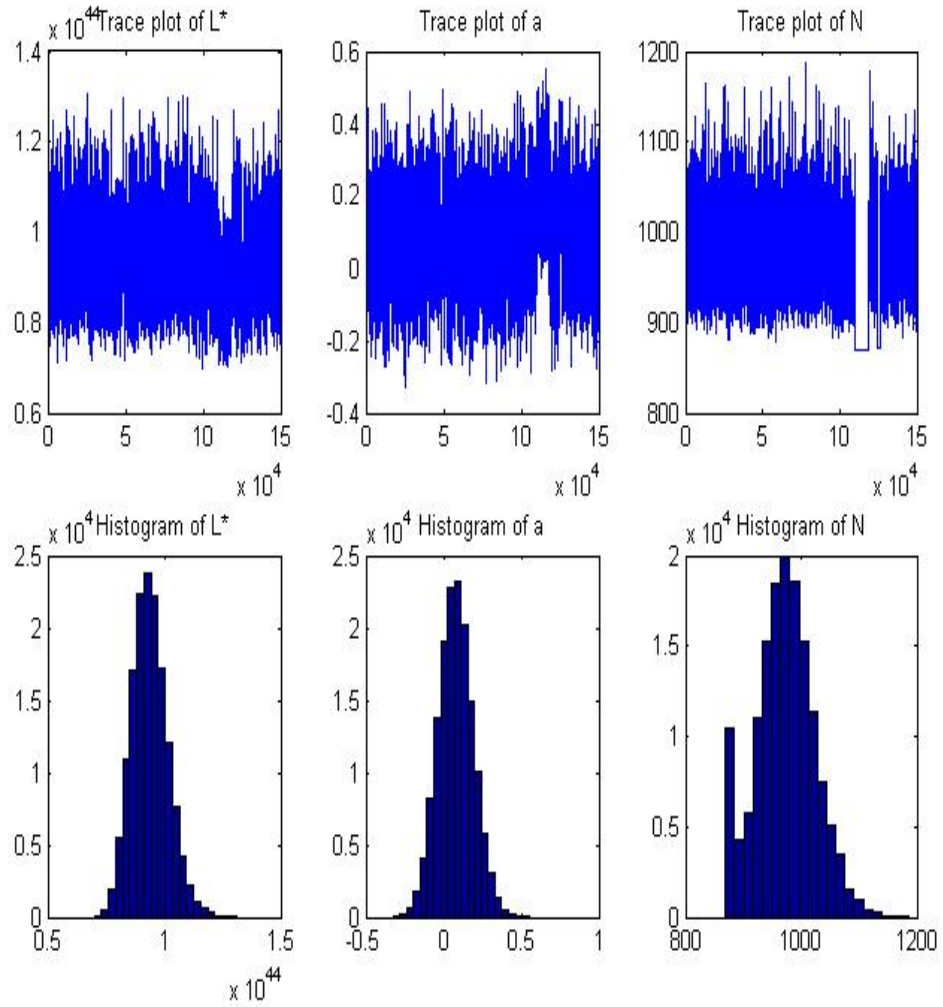
τις κατανομές των παραμέτρων, χρειάστηκε πολύ μεγάλη περίοδος ζεστάματος για να μπορεί να συγκλίνει ο αλγόριθμος (ειδικά όταν το $L_{min} = 10^{44}$). Αυτό μεταφράζεται ως μεγάλο υπολογιστικό κόστος. Χρειάστηκε ακόμα μεγαλύτερη περίοδος ζεστάματος για να συγκλίνει ο αλγόριθμος όταν το $L_{min} = 2 \times 10^{44}$ και μεγαλύτερο δείγμα από την εκ των υστέρων κατανομή για να μπορούμε να πάρουμε ικανοποιητικά αποτελέσματα για τις εκ των υστέρων κατανομές των παραμέτρων. Αξίζει να σημειωθεί ότι έχουμε απλοποιήσει αρκετά το πρόβλημα που αντιμετωπίζουμε. Πήραμε μια απλή συνάρτηση φωτεινότητας, θεωρήσαμε ότι αυτή εξαρτάται από τη φωτεινότητα μόνο, ενώ στην πραγματικότητα εξαρτάται και από την ερυθρομετατόπιση. Ακόμα και σε αυτή την περίπτωση ο αλγόριθμος χρειάστηκε να τρέξει πολλές φορές και χρειάστηκε πολύς χρόνος για να είναι η συμπερασματολογία μας ικανοποιητικά ακριβής. Αντίστοιχη υπολογιστική δυσκολία για το συγκεκριμένο πρόβλημα αντιμετωπίζεται και από την κλασσική στατιστική. Κατά τη μελέτη μας και τη σύγκριση των αποτελεσμάτων μας με δημοσιευμένες εργασίες παρατηρήσαμε ότι η μέθοδος της συμπερασματολογίας μας υπερτερεί υπολογιστικά σε σχέση με τη μέθοδο μέγιστης πιθανοφάνειας. Οι πράξεις για να εφαρμόσουμε τη μέθοδο της μέγιστης πιθανοφάνειας είναι περίπλοκες και υπολογιστικά πολύ δύσκολες ακόμα και στην απλή περίπτωση της Schechter συνάρτησης φωτεινότητας. Επίσης, ο αλγόριθμος μας έτρεξε πιο γρήγορα από τον αλγόριθμο για εκτιμήσεις με τη μέθοδο μέγιστης πιθανοφάνειας.

Τώρα ως προς την ακρίβεια της συμπερασματολογίας, η μπεϋζιανή προσέγγιση επίσης υπερτερεί. Οι κατανομές των παραμέτρων που μας ενδιαφέρουν είναι μονοκόρυφες, αλλά ασύμμετρες. Ομοίως μη συμμετρικά γύρω από την κορυφή είναι και τα διαστήματα αξιοπιστίας, γι' αυτό και παράγουν πιο ακριβή συμπερασματολογία από τα προσεγγιστικά κλασσικά συμμετρικά διαστήματα εμπιστοσύνης. Θεωρούμε λοιπόν ότι η μπεϋζιανή προσέγγιση υπερτερεί από τη μέθοδο μέγιστης πιθανοφάνειας.

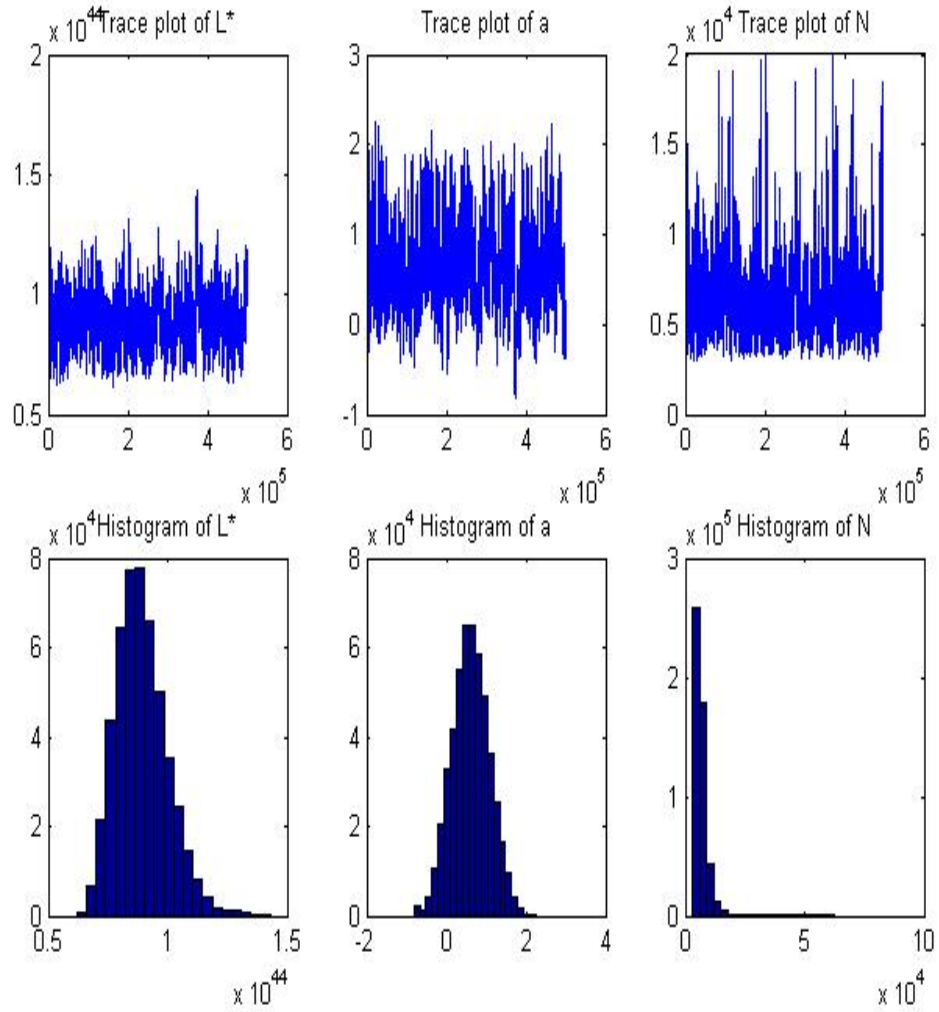
Διαγράμματα προσομοιώσεων για δεδομένα Poisson



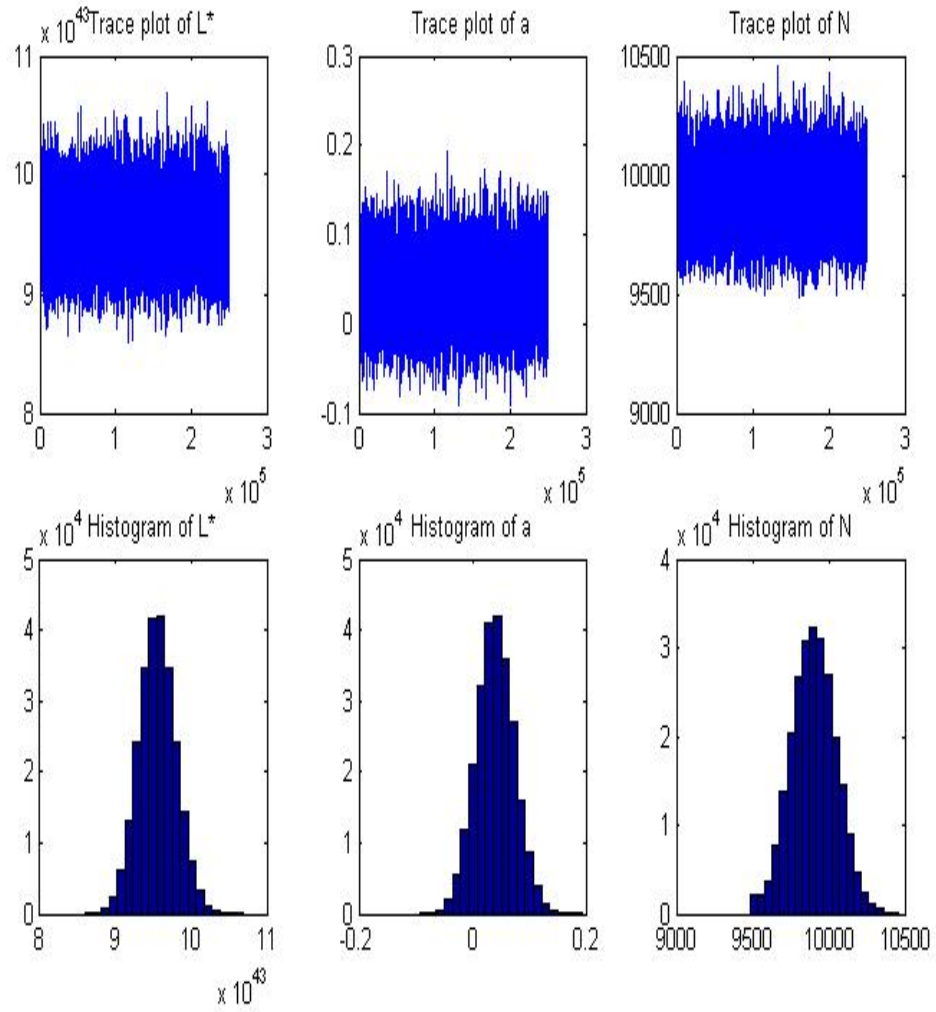
Σχήμα 5.1: Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και Poisson δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$.



Σχήμα 5.2: Διαγράμματα προσομοίωσης για *Schechter* συνάρτηση φωτεινότητας και *Poisson* δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$.



Σχήμα 5.3: Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$.



Σχήμα 5.4: Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και Poisson δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$.

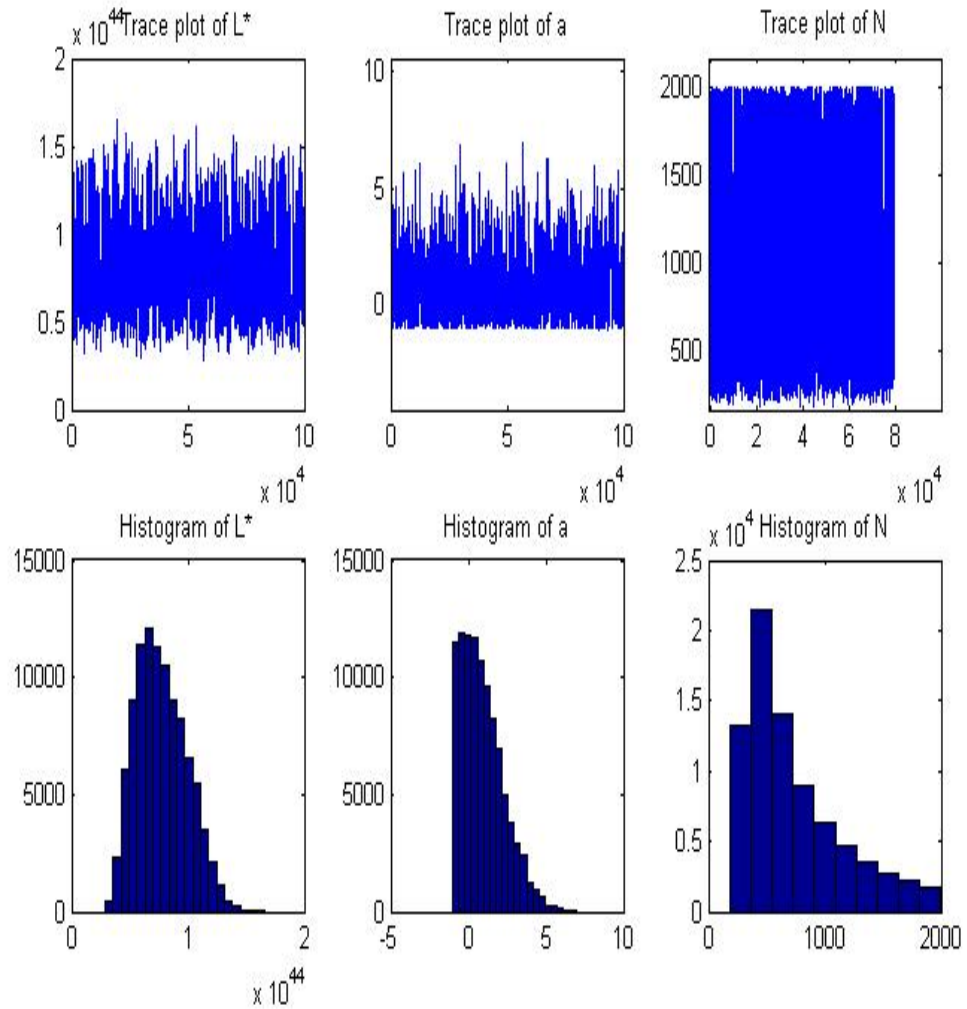
5.2 Schechter συνάρτηση φωτεινότητας για διωνυμικά δεδομένα

Κάναμε συνολικά τέσσερις προσομοιώσεις για τον αλγόριθμο που υποθέτει διωνυμικά δεδομένα όπως και για τα δεδομένα Poisson. Πήραμε δύο διαφορετικά N και δύο διαφορετικά L_{min} . Κάναμε δύο προσομοιώσεις για την εκτίμηση των παραμέτρων με ελάχιστη φωτεινότητα $L_{min}^* = 2 \times 10^{44}$ και $N = 1000$, $N = 10000$ και άλλες δύο με ελάχιστη φωτεινότητα $L_{min}^* = 2 \times 10^{43}$ και $N = 1000$, $N = 10000$.

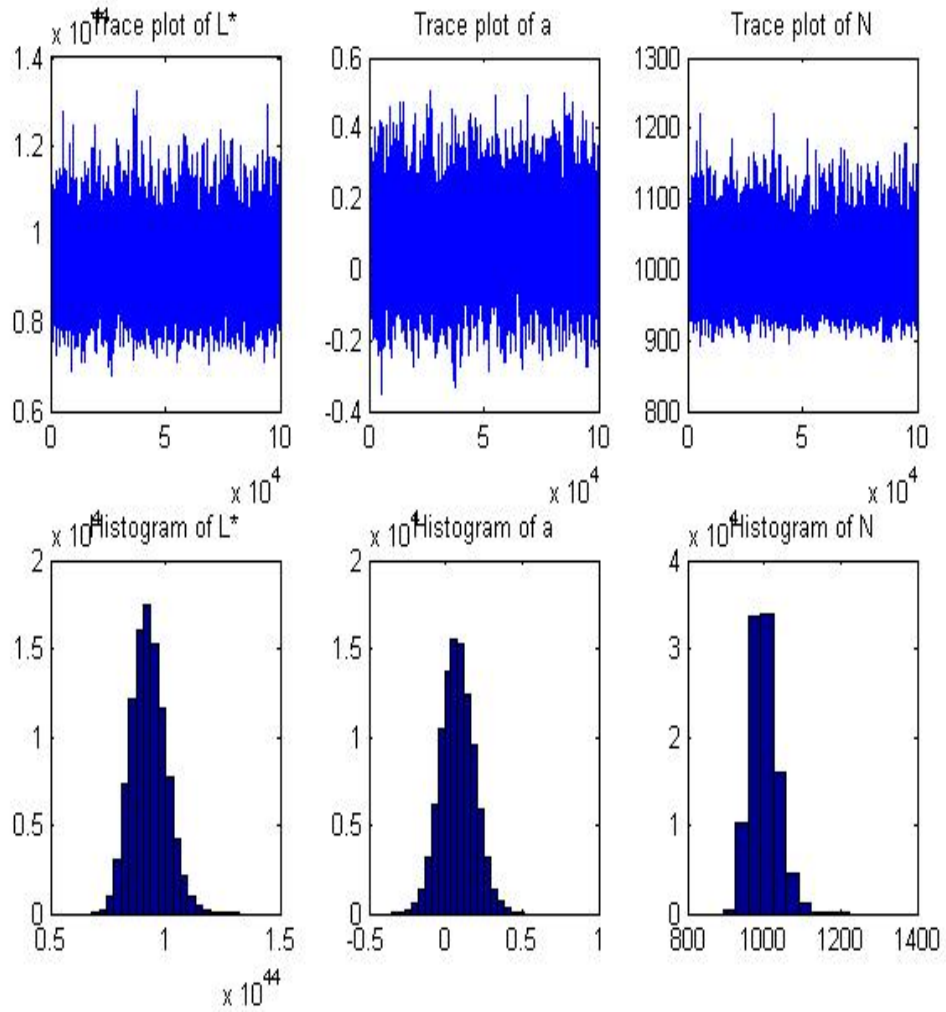
Σε όλες τις περιπτώσεις τα αποτελέσματά μας ήταν πολύ καλά. Οι πραγματικές τιμές των παραμέτρων βρίσκονται μέσα και στα 20 διαστήματα αξιοπιστίας που κατασκέυάσαμε, στατιστικής σημαντικότητας 95%. Όπως και στα Poisson δεδομένα ο αλγόριθμος έδωσε καλύτερα αποτελέσματα (πήραμε δηλαδή μικρότερα διαστήματα αξιοπιστίας) και είχε πιο γρήγορη σύγκλιση όταν το ελάχιστο όριο φωτεινότητας ήταν $L_{min}^* = 2 \times 10^{43}$ και ο αριθμός των πηγών ήταν $N = 10000$. Παρατηρήσαμε ότι κάτω από την υπόθεση για διωνυμικά δεδομένα ο αλγόριθμος συγκλίνει γρηγορότερα και δίνει λίγο καλύτερα αποτελέσματα, ως προς την ακρίβεια των εκτιμήσεων σε σχέση με τον αλγόριθμο που τρέξαμε κάτω από την υπόθεση της Poisson.

Στις επόμενες σελίδες παρουσιάζουμε τα ιστογράμματα και τα ιχνογράμματα των παραμέτρων που θέλουμε να εκτιμήσουμε για τις τέσσερις προαναφερθείσες περιπτώσεις.

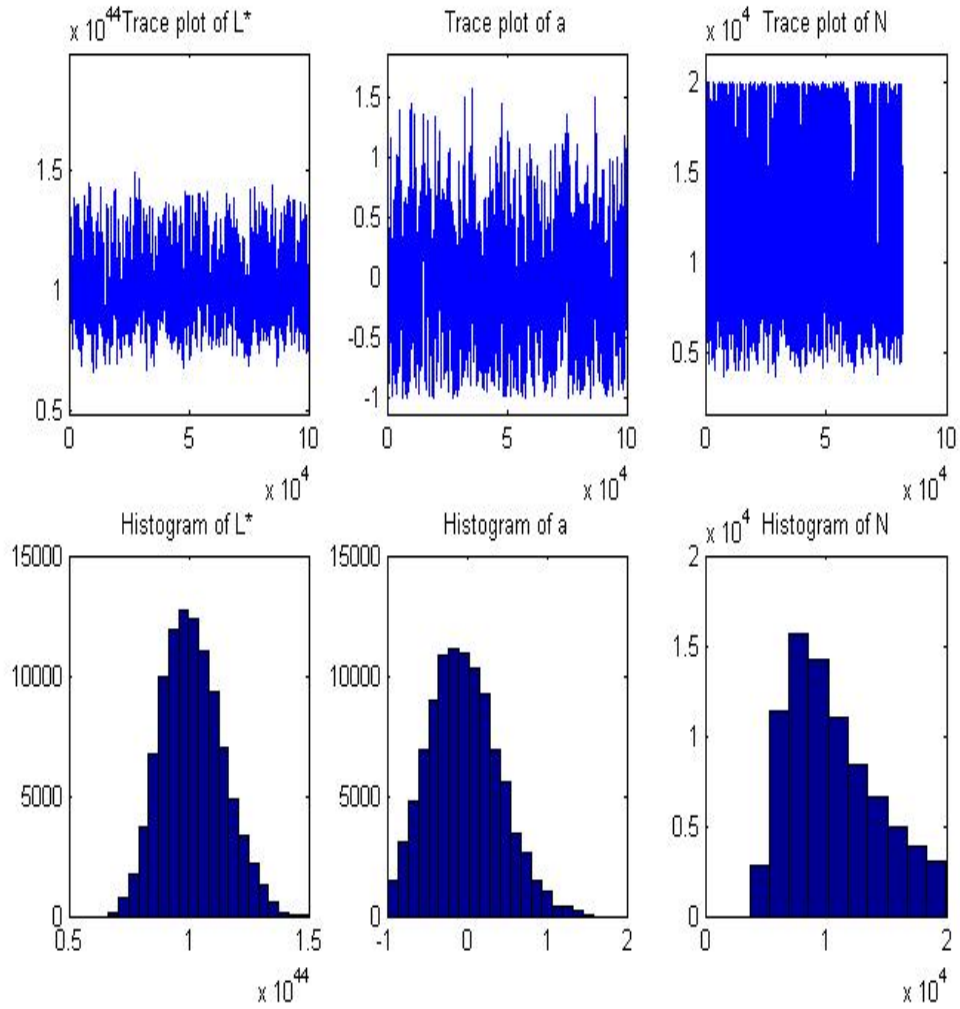
Διαγράμματα προσομοιώσεων για διωνυμικά δεδομένα



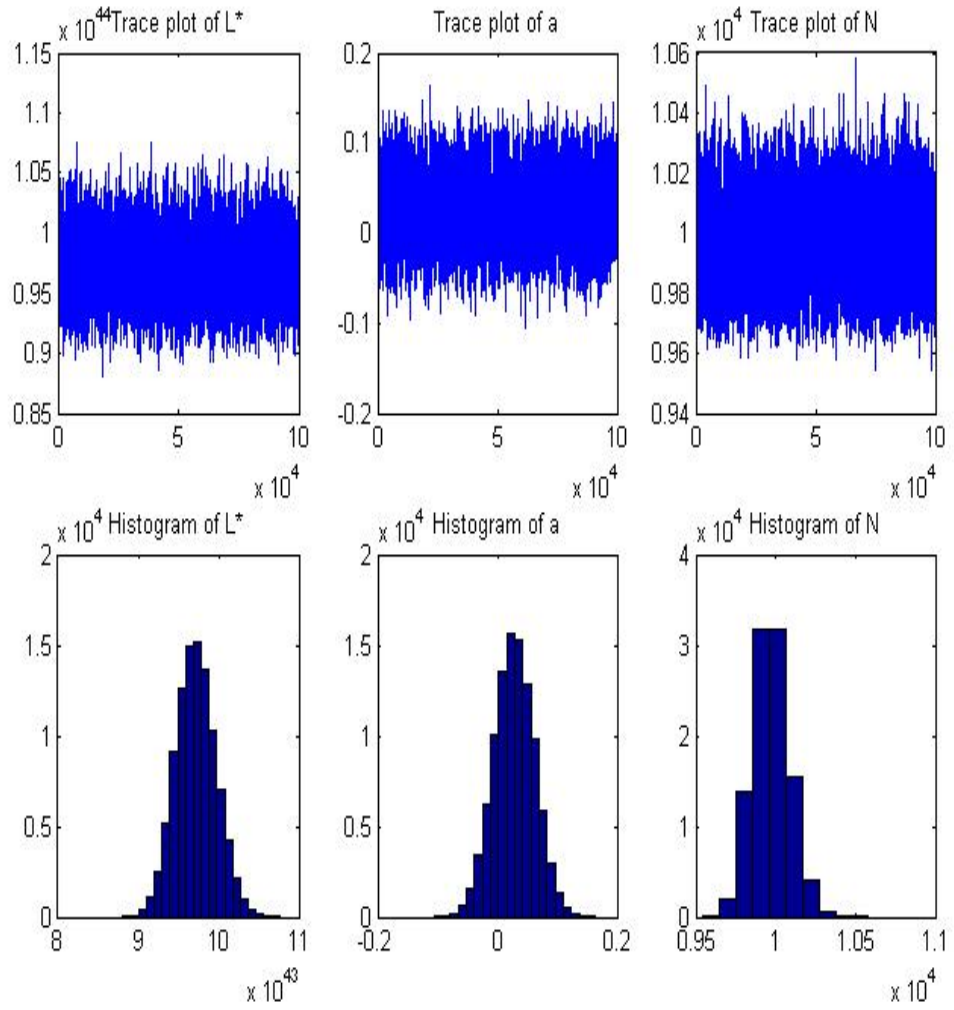
Σχήμα 5.5: Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$.



Σχήμα 5.6: Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 1000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$.



Σχήμα 5.7: Διαγράμματα προσομοίωσης για Schechter συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{44}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{44}$.



Σχήμα 5.8: Διαγράμματα προσομοίωσης για *Schechter* συνάρτηση φωτεινότητας και διωνυμικά δεδομένα. Τα αρχικά δεδομένα προήλθαν από την κατανομή φωτεινότητας με παραμέτρους $N = 10000$, $L^* = 10^{43}$, $\alpha = 0$ και $L_{min} = 2 \times 10^{43}$.

5.3 Συμπεράσματα

Στην παράγραφο αυτή επιχειρούμε να παρουσιάσουμε τα συμπεράσματα μας από τη σύγκριση των δύο τρόπων μοντελοποίησης του προβλήματος που εξετάσαμε σε αυτή τη διπλωματική εργασία. Γενικότερα παρατηρήσαμε ότι όταν στο πρόβλημα που αντιμετωπίζουμε ο συνολικός αριθμός των πηγών (N) ήταν μικρός, τότε η υπόθεση για τα διωνυμικά δεδομένα έδινε ελαφρώς καλύτερη συμπερασματολογία. Ο αλγόριθμος έτρεχε πιο γρήγορα και χρειαζόταν μικρότερη περίοδο ζεστάματος. Αυτό βέβαια οφείλεται και στο γεγονός ότι η τυχαία μεταβλητή N προσομοιώνεται με αλγόριθμο Gibbs που είναι σαφώς πιο γρήγορη προσομοίωση παρά με τον αλγόριθμο Metropolis-Hastings. Καθώς το N μεγαλώνει, η πιθανοφάνεια του μοντέλου κάτω από την υπόθεση ότι τα δεδομένα είναι Poisson και η πιθανοφάνεια κάτω από την υπόθεση των διωνυμικών δεδομένων δίνουν την ίδια συμπερασματολογία. Αυτό επιβεβαιώνει και τη θεωρία, καθώς όσο μεγαλώνει το N η διωνυμική κατανομή προσεγγίζεται από την κατανομή Poisson. Η διαδικασία Poisson έχει το πλεονέκτημα ότι έχει πολύ και καλά μελετημένες καλές ιδιότητες, ενώ διάφορα θεωρήματα μας επιτρέπουν να μετασχηματίζουμε τα δεδομένα και ακόμα να ισχύουν οι καλές ιδιότητες. Αυτός είναι και ο λόγος που στην επιστημονική κοινότητα των αστροφυσικών οι διαδικασίες Poisson χρησιμοποιούνται ευρύτατα, ενώ οι διωνυμικές διαδικασίες μόνο τελευταία άρχισαν χρησιμοποιούνται.

Παράρτημα Ι

Χρήσιμες κατανομές πιθανότητας

Εκθετική Κατανομή

Η συνεχής τυχαία μεταβλητή X ακολουθεί την εκθετική κατανομή αν έχει τη συνάρτηση πυκνότητας πιθανότητας:

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \lambda > 0 \\ 0, & x < 0 \end{cases}$$

Αρνητική διωνυμική Κατανομή

Η διακριτή τυχαία μεταβλητή X ακολουθεί την αρνητική διωνυμική κατανομή με παραμέτρους n, p αν έχει συνάρτηση πιθανότητας:

$$f(k) = \binom{k-1}{n-1} (1-p)^{k-n} p^n, \quad k = n, n+1, \dots$$

Gamma Κατανομή

Η τυχαία μεταβλητή X ακολουθεί την κατανομή *Gamma* με παραμέτρους α, β , $(\Gamma(\alpha, \beta))$ αν έχει την συνάρτηση πυκνότητας πιθανότητας

$$f(x; \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0, \alpha > 0, \beta > 0 \\ 0 & \end{cases}$$

Beta Κατανομή

Η τυχαία μεταβλητή X ακολουθεί την κατανομή *Beta*(α, β) αν έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x; \alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & 0 \leq x \leq 1, \alpha > 0, \beta > 0 \\ 0 & \end{cases}$$

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, & 0 \leq x \leq 1, \alpha > 0, \beta > 0 \\ 0 & \end{cases}$$

Dirichlet Κατανομή

Η κατανομή Dirichlet είναι πολυπαραμετρική γενίκευση της Beta κατανομής. Η πολυδιάστατη τυχαία μεταβλητή $\underline{X} = (X_1, \dots, X_k)$ με παραμέτρους $\alpha_1, \dots, \alpha_k$ αν έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x_1, \dots, x_k; \alpha_1, \dots, \alpha_k) = \begin{cases} \frac{\Gamma(\alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)} \prod_{i=1}^k x_i^{\alpha_i - 1}, \\ x_1, \dots, x_k \geq 0, \alpha_1, \dots, \alpha_k > 0, \sum_{i=1}^k x_i = 1 \\ 0, \text{ αλλιού} \end{cases}$$

Πολυδιάστατη t-Student Κατανομή

Η p -διάστατη τυχαία μεταβλητή ακολουθεί την p -διάστατη t-student κατανομή με ν βαθμούς ελευθερίας p - διάστατο διάνυσμα μέσων μ και θετικά ορισμένο πίνακα κλίμακας $\Sigma \in \mathbb{R}^{p \times p}$ αν έχει κατανομή πυκνότητα πιθανότητας:

$$f(\underline{X}; \mu, \Sigma, \nu) = \frac{\Gamma(\nu + p)/2}{\Gamma(\nu/2) \nu^{p/2} \pi^{p/2}} |\Sigma|^{-1/2} \left[1 + \frac{1}{\nu} (\underline{X} - \mu)^T \Sigma^{-1} (\underline{X} - \mu) \right]^{-(\nu+p)/2}$$

Στην περίπτωση που έχουμε 1 βαθμό ελευθερίας τότε ονομάζουμε την κατανομή Cauchy ή Lorentz κατανομή.

Κανονική Κατανομή

Η συνεχής τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή ή την Gaussian κατανομή με μέσο μ και διακύμανση σ^2 αν έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

X^2 Κατανομή

Η συνεχής τυχαία μεταβλητή ακολουθεί την X^2 κατανομή με k βαθμούς ελευθερίας αν έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} \frac{1}{2^{k/2} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, & x \geq 0 \\ 0, & \text{αλλιού} \end{cases}$$

Η χ^2 κατανομή με k βαθμούς ελευθερίας είναι το άθροισμα k ανεξάρτητων τυποποιημένων κατανομών.

Wishart Κατανομή

Η κατανομή Wishart είναι πολυδιάστατη επέκταση της χ^2 κατανομής.

Η τυχαία μεταβλητή $X \in \mathbb{R}^{p \times p}$ ακολουθεί την Wishart κατανομή με $\nu > p-1$ βαθμούς ελευθερίας και πίνακα κλίμακας $\Sigma \in \mathbb{R}^{p \times p}$, θετικά ορισμένο, αν έχει συνάρτηση πιθανότητας

$$f(X; \Sigma, \nu) = \left[2^{\nu p/2} \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma\left(\frac{\nu+1-i}{2}\right) \right]^{-1} |\Sigma|^{-\nu/2} |X|^{(\nu-p-1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} X) \right\}$$

Inverse-Wishart Κατανομή

Η τυχαία μεταβλητή $X \in \mathbb{R}^{p \times p}$ έχει την κατανομή Inverse-Wishart με $\nu > p-1$ βαθμούς ελευθερίας και πίνακα κλίμακας $S \in \mathbb{R}^{p \times p}$, θετικά ορισμένο, αν έχει συνάρτηση πιθανότητας

$$f(X; S, \nu) = \left[2^{\nu p/2} \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma\left(\frac{\nu+1-i}{2}\right) \right]^{-1} |S|^{\nu/2} |X|^{-(\nu+p+1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(X^{-1} S) \right\}$$

Παράρτημα II

Κώδικες Matlab για Schechter συνάρτηση φωτεινότητας

Αλγόριθμος για Poisson δεδομένα και Schechter συνάρτηση φωτεινότητας

Αλγόριθμος σε Matlab

```
function [draws,ac1,ac2,ac3]=mh2b(data,Lmin,v1,v2,nburn,ndraw)

%metropolis-hastings for Schechter luminosity function and poisson data
%The luminosity function is
%f(L)=(N/Lstar*Gamma(a+1))*(L/Lstar)^a*exp{(-L/Lstar)}
%the luminosity function is independent from
%Assume that N is discrete

%start with initial value (a,Lstar,N)
L=data;
n=length(data);
draws=zeros(ndraw,3);

%draw initial values
a=0;
Lstar=0.9*10^44;
lLstar=log(Lstar);
k=a+1;
N=10000;

accept1=0; %counter for acceptance probabilities
accept2=0;
accept3=0;
mu1=lLstar;
mu2=log(k);

%draw a proposal value for Log(Lstar) from a normal distribution
%MCMC follows:
it=-nburn;
while it<ndraw;
    it=it+1;
    %draw lLstar candidate and kcandidate
```

```

lLstarcan=normrnd(mu1,v1);
Lstarcan=exp(lLstarcan); %
%acceptance probability for Lstar and for a
logpacc1=n*k*(log(Lstar)-log(Lstarcan))-(1/Lstarcan-1/Lstar)*sum(L)-...
    N*(1-gamcdf(Lmin,k,Lstarcan)-(1-gamcdf(Lmin,k,Lstar)));

u1=rand;
if log(u1)<logpacc1
    accept1=accept1+1;
    lLstar=lLstarcan;
end
Lstar=exp(lLstar);

lkcan=normrnd(mu2,v2);
kcan=exp(lkcan);
logpacc2=log(kcan)-log(k)+n*gammaLn(k)-n*gammaLn(kcan)+...
    N*((1-gamcdf(Lmin,k,Lstar))-(1-gamcdf(Lmin,kcan,Lstar)))+...
    (kcan-k)*sum(log(L/Lstar));

u2=rand;
if log(u2)<logpacc2
    accept2=accept2+1;
    k=kcan;
end

a=k-1;
mu1=lLstar;
mu2=log(k);

%draw values for N
p=1-gamcdf(Lmin,k,Lstar);
Ncan=nbirnd(n,p)+n;

%calculate acceptance probability for N
logpacc3=(n-1)*log(Ncan)-(n-1)*log(N)-p*(Ncan-N)+sum(log(1:(N-1)))-...
    sum(log(1:(Ncan-1)))+sum(log(1:(Ncan-n)))-sum(log(1:(N-n)))+...
    (N-Ncan)*log((1-p));

u3=rand;
if log(u3)<logpacc3
    accept3=accept3+1;
    N=Ncan;
end

%after burn-in record Lstar, a, N
if it>0;
    draws(it,1)=Lstar;
    draws(it,2)=a;

```

```

        draws(it,3)=N;
    end

end
    ac1=accept1/(nburn+ndraw); %acceptance ratio for Lstar
    ac2=accept2/(nburn+ndraw); %acceptacne ratio for k
    ac3=accept3/(nburn+ndraw); %acceptance ratio for N
end

```

Αλγόριθμος για διωνυμικά δεδομένα και Schechter συνάρτηση φωτεινότητας

Αλγόριθμος σε Matlab

```
function [draws,ac1,ac2,N]=mh3(data,Lmin,v1,v2,nburn,ndraw)

%metropolis-hastings for Schechter luminosity function
%L is the luminosity, z is the redshift, a,Lstar parameters to be estimated
%Binomial data

%start with initial value (a,Lstar,N)
L=data;
n=length(data);
draws=zeros(ndraw,3);

%draw initial values !!!!!!! check for MLE
a=0; %
Lstar=0.9*10^44; %%%
lLstar=log(Lstar);
k=a+1;

accept1=0; %counter for acceptance probability for Lstar
accept2=0;
mu1=lLstar;
mu2=log(k);

%draw a proposal value for Log(Lstar) from a normal distribution
%MCMC follows:
it=-nburn;
while it<ndraw;
    it=it+1;
    %draw lLstar candidate and kcandidate

    lLstarcan=normrnd(mu1,v1);
    Lstarcan=exp(lLstarcan); %
    %acceptance probability for Lstar and for a
    logpacc1=n*log(1-gamcdf(Lmin,k,Lstar))-n*log(1-gamcdf(Lmin,k,Lstarcan))+...
        n*k*(log(Lstar)-log(Lstarcan))-sum(L)*(1/Lstarcan-1/Lstar);
    %draw random numbers from Uniform(1,0) for accepting or rejecting the
    %proposal values

    u1=rand;
    if log(u1)<logpacc1
        accept1=accept1+1;
        lLstar=lLstarcan;
    end
    Lstar=exp(lLstar);
```

```

lkcan=normrnd(mu2,v2);
kcan=exp(lkcan);
logpacc2=log(kcan)-log(k)+n*gammaIn(k)-n*gammaIn(kcan)+...
    n*(log(1-gamcdf(Lmin,k,Lstar))-log(1-gamcdf(Lmin,kcan,Lstar)))+...
    (kcan-k)*sum(log(L/Lstar));
u2=rand;
if log(u2)<logpacc2
    accept2=accept2+1;
    k=kcan;
end

a=k-1;
mu1=lLstar;
mu2=log(k);

%after burn-in record
if it>0;
    draws(it,1)=Lstar;
    draws(it,2)=a;
end
end
ac1=accept1/(nburn+ndraw); %acceptance ratio for Lstar
ac2=accept2/(nburn+ndraw); %acceptacne ration for k

%Sample for N from negative binomial distribution
for i=1:ndraw;
    j=rand(1,n);
    p=1-gamcdf(Lmin,draws(i,2)+1,draws(i,1));
    m=sum(floor((log(j))/log(1-p)));
    draws(i,3)=n+m;
    %N(i,1)=nbinrnd(n,p);
end
p=1-gamcdf(Lmin,draws(:,2)+1,draws(:,1));
N=nbinrnd(n,p)+n;
end

%generate data and run MH example

%First scenario : Lmin=2*10^44 ,N=100000
clear all
Lmin=2*10^44; N=10000;
D=exprnd(10^44,N,1);
data=D(D>Lmin);
n=length(data);
%for Lmin=2*10^44 use v1=0.25, v2=1
v1=0.1;v2=0.5;
nburn=50000;
ndraw=100000;

```

```

[ draws , ac1 , ac2 , Ndraws]=mh3( data , Lmin , v1 , v2 , nburn , ndraw );

mean( draws (: , 1))
mean( draws (: , 2))
mean( draws (: , 3))

subplot(2,3,1)
plot( draws (: , 1))

subplot(2,3,2)
plot( draws (: , 2))

subplot(2,3,3)
%plot( draws (: , 3)
plot( draws( draws (: , 3) < 2*N, 3))

subplot(2,3,4)
hist( draws (: , 1) , 20)

subplot(2,3,5)
hist( draws (: , 2) , 20)

subplot(2,3,6)
hist( draws (: , 3) , 20)
%hist( draws( draws (: , 3) < 2*N, 3))

P=zeros(20,6);
M=zeros(20,3);

for i=1:20;
    [ draws]=mh3( data , Lmin , v1 , v2 , nburn , ndraw );
    P(i,1:2)= prctile( draws (: , 1) , [5 95]);
    P(i,3:4)= prctile( draws (: , 2) , [5 95]);
    P(i,5:6)= prctile( draws (: , 3) , [5 95]);
    M(i,1)=mean( draws (: , 1));
    M(i,2)=mean( draws (: , 2));
    M(i,3)=mean( draws (: , 3));
end

```

Παράρτημα III

Αλγόριθμος MCMC για τη συνάρτηση φωτεινότητας που ακολουθεί μίξη κανονικών κατανομών και κώδικας Matlab

Διωνυμικά Δεδομένα

- Αρχίζουμε με αρχικές τιμές των $\pi, \mu, \Sigma, \mu_0, A, T$.
- Προτείνουμε τιμή για π από $Dirichlet(\tilde{g}_1, \dots, \tilde{g}_K)$, όπου $\tilde{g}_k = c_\pi n \tilde{\pi}_k + 1$ όπου n είναι ο αριθμός των πηγών στην έρευνα, c_π είναι σταθερά που καθορίζει τη σταθερά της κατανομής Dirichlet.

Μικρή τιμή του c_π παράγει τιμή $\hat{\pi}$ μεγαλύτερη από το $\tilde{\pi}$

Επιλέγουμε σταθερά τέτοια ώστε οι προτεινόμενες τιμές θα γίνονται αποδεκτές με πιθανότητα 15-40%.

Η πιθανότητα αποδοχής της προτεινόμενης τιμής θα είναι:

$$\begin{aligned}
 p_{acc1} &= \frac{Dirichlet(\tilde{\pi}|\hat{g})p(\hat{\theta}|L_{obs}, z_{obs})}{Dirichlet(\hat{\pi}|\tilde{g})p(\tilde{\theta}|L_{obs}, z_{obs})} \\
 &= \frac{\frac{\Gamma(\tilde{g}_1 + \dots + \tilde{g}_K)}{\Gamma(\tilde{g}_1) \dots \Gamma(\tilde{g}_K)} \prod_{k=1}^K \tilde{\pi}_k^{\tilde{g}_k - 1} \prod_{k=1}^K Cauchy_2(\mu_k | \mu_0, T) InvWishart_1(\Sigma_k | A)}{\frac{\Gamma(\hat{g}_1 + \dots + \hat{g}_K)}{\Gamma(\hat{g}_1) \dots \Gamma(\hat{g}_K)} \prod_{k=1}^K \hat{\pi}_k^{\hat{g}_k - 1} \prod_{k=1}^K Cauchy_2(\mu_k | \mu_0, T) InvWishart_1(\Sigma_k | A)} \\
 &\times \left[\frac{\sum_{k=1}^K \frac{\hat{\pi}_k}{2\pi|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)]}{\sum_{k=1}^K \frac{\tilde{\pi}_k}{2\pi|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)]} \right]^{-n} \\
 &\times \prod_{i \in A_{obs}} \frac{\sum_{k=1}^K \left[\frac{\hat{\pi}_k}{2\pi|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \right] \right]}{\sum_{k=1}^K \left[\frac{\tilde{\pi}_k}{2\pi|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \right] \right]}
 \end{aligned}$$

$\Rightarrow$

$$\begin{aligned}
p_{acc1} &= \frac{\Gamma(\tilde{g}_1 + \dots + \tilde{g}_K)}{\Gamma(\tilde{g}_1) \dots \Gamma(\tilde{g}_K)} \frac{\prod_{k=1}^K \tilde{\pi}_k^{\tilde{g}_k - 1}}{\frac{\Gamma(\hat{g}_1 + \dots + \hat{g}_K)}{\Gamma(\hat{g}_1) \dots \Gamma(\hat{g}_K)} \prod_{k=1}^K \hat{\pi}_k^{\hat{g}_k - 1}} \\
&\times \left[\frac{\sum_{k=1}^K \frac{\hat{\pi}_k}{|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)]}{\sum_{k=1}^K \frac{\tilde{\pi}_k}{|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)]} \right]^{-n} \\
&\times \prod_{i \in A_{obs}} \frac{\sum_{k=1}^K \left[\frac{\hat{\pi}_k}{|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right] \right]}{\sum_{k=1}^K \left[\frac{\tilde{\pi}_k}{|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right] \right]}
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\begin{aligned}
\log p_{acc1} &= \log(\Gamma(\tilde{g}_1 + \dots + \tilde{g}_K)) - \log(\Gamma(\tilde{g}_1) + \dots + \Gamma(\tilde{g}_K)) + \log(\Gamma(\hat{g}_1) + \dots + \Gamma(\hat{g}_K)) - \log(\Gamma(\hat{g}_1 + \dots + \hat{g}_K)) \\
&+ \sum_{k=1}^K ((\hat{g}_k - 1) \log(\tilde{\pi}_k)) - \sum_{k=1}^K ((\tilde{g}_k - 1) \log(\hat{\pi}_k)) - n \log \left(\sum_{k=1}^K \frac{\hat{\pi}_k}{|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)] \right) \\
&+ n \log \left(\sum_{k=1}^K \frac{\tilde{\pi}_k}{|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)] \right) \\
&+ \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log(\hat{\pi}_k) - \frac{1}{2} \log(|\Sigma_k|) - \frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right] - \\
&- \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log(\tilde{\pi}_k) - \frac{1}{2} \log(|\Sigma_k|) - \frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right]
\end{aligned}$$

- Προσομοιώνουμε ένα τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο τυχαίος αριθμός είναι μικρότερος από την πιθανότητα αποδοχής τότε αποδεχόμαστε την προτεινόμενη τιμή. Σε αντίθετη περίπτωση η τιμή του π μένει ως έχει και συνεχίζουμε στο επόμενο βήμα του αλγόριθμου.
- Για την προσομοίωση τιμών για το μέσο εφαρμόζουμε Metropolis-Hastings αλγόριθμο τυχαίου περιπάτου. Προτείνουμε τιμή για το μ_k από κανονική κατανομή $N_k(\tilde{\mu}_k, V_k)$. Ο πίνακας διακύμανσης της κ-διάστατης κανονικής κατανομής V_k θα είναι τέτοιος ώστε η πιθανότητα αποδοχής της προτεινόμενης τιμής θα είναι περίπου 20-50%.

Η πιθανότητα αποδοχής σε αυτό το βήμα θα είναι:

p_{acc2}

$$\begin{aligned}
&= \frac{p(\hat{\theta}|L_{obs}, z_{obs})}{p(\tilde{\theta}|L_{obs}, z_{obs})} \\
&= \frac{\prod_{k=1}^K Cauchy_2(\hat{\mu}_k|\mu_0, T) Inv - Wishart_1(\Sigma_k|A)p(I=1|\hat{\theta})^{-n}}{\prod_{k=1}^K Cauchy_2(\tilde{\mu}_k|\mu_0, T) Inv - Wishart_1(\Sigma_k|A)p(I=1|\tilde{\theta})^{-n}} \\
&\times \frac{\prod_{i \in A_{obs}} p(\log L_i, \log z_i|\hat{\theta})}{\prod_{i \in A_{obs}} p(\log L_i, \log z_i|\tilde{\theta})} \\
&= \frac{\prod_{k=1}^K [1 + (\hat{\mu}_k - \mu_0)^T T^{-1}(\hat{\mu}_k - \mu_0)]^{-\frac{(1+p)}{2}}}{\prod_{k=1}^K [1 + (\tilde{\mu}_k - \mu_0)^T T^{-1}(\tilde{\mu}_k - \mu_0)]^{-\frac{(1+p)}{2}}} \\
&\times \left[\frac{\sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} F(x_{high}\hat{\mu}_k, \Sigma_k) - F(x_{low}, \hat{\mu}_k, \Sigma_k)}{\sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} F(x_{high}\tilde{\mu}_k, \Sigma_k) - F(x_{low}, \tilde{\mu}_k, \Sigma_k)} \right]^{-n} \\
&\times \frac{\prod_{i \in A_{obs}} \sum_{k=1}^K \left[\frac{\pi_k}{|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2}(x - \hat{\mu}_k)^T \Sigma_k^{-1}(x - \hat{\mu}_k) \right] \right]}{\prod_{i \in A_{obs}} \sum_{k=1}^K \left[\frac{\pi_k}{|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2}(x - \tilde{\mu}_k)^T \Sigma_k^{-1}(x - \tilde{\mu}_k) \right] \right]}
\end{aligned}$$

(Στην περίπτωση μας $p = 2$)

Ο αλγόριθμος της πιθανότητας αποδοχής είναι:

$$\begin{aligned}
\log p_{acc2} &= -\frac{3}{2} \log(1 + (\hat{\mu}_k - \mu_0)^T T^{-1}(\hat{\mu}_k - \mu_0)) + \\
&+ \frac{3}{2} \log((1 + (\tilde{\mu}_k - \mu_0)^T T^{-1}(\tilde{\mu}_k - \mu_0)) - \\
&- n \left[\log \left(\sum_{k=1}^K \frac{\pi_k}{|\Sigma_k|^{1/2}} (F(x_{high}, \hat{\mu}_k, \Sigma_k)) - F(x_{low}, \hat{\mu}_k, \Sigma_k) \right) \right] + \\
&+ n \left[\log \left(\sum_{k=1}^K \frac{\pi_k}{|\Sigma_k|^{1/2}} F(x_{high}, \tilde{\mu}_k, \Sigma_k) - F(x_{low}, \tilde{\mu}_k, \Sigma_k) \right) \right] + \\
&+ \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log \left(\frac{\pi_k}{|\Sigma_k|^{1/2}} \right) - \frac{1}{2} (x_i - \hat{\mu}_k)^T \Sigma_k^{-1} (x_i - \hat{\mu}_k) \right] - \\
&- \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log \left(\frac{\pi_k}{|\Sigma_k|^{1/2}} \right) - \frac{1}{2} (x_i - \tilde{\mu}_k)^T \Sigma_k^{-1} (x_i - \tilde{\mu}_k) \right] = \\
&= -\frac{3}{2} \log(1 + (\hat{\mu}_k - \mu_0)^T T^{-1}(\hat{\mu}_k - \mu_0)) + \frac{3}{2} \log((1 + (\tilde{\mu}_k - \mu_0)^T T^{-1}(\tilde{\mu}_k - \mu_0)) - \\
&- n \left[\log \left(\sum_{k=1}^K \frac{\pi_k}{|\Sigma_k|^{1/2}} (F(x_{high}, \hat{\mu}_k, \Sigma_k)) - F(x_{low}, \hat{\mu}_k, \Sigma_k) \right) \right] + \\
&+ n \left[\log \left(\sum_{k=1}^K \frac{\pi_k}{|\Sigma_k|^{1/2}} F(x_{high}, \tilde{\mu}_k, \Sigma_k) - F(x_{low}, \tilde{\mu}_k, \Sigma_k) \right) \right] +
\end{aligned}$$

$$+ \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log(\pi_k) - \frac{1}{2} \log(|\Sigma_k|) - \frac{1}{2} (x_i - \hat{\mu}_k)^T \Sigma_k^{-1} (x_i - \hat{\mu}_k) \right] - \\ - \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log(\pi_k) - \frac{1}{2} \log(|\Sigma_k|) - \frac{1}{2} (x_i - \tilde{\mu}_k)^T \Sigma_k^{-1} (x_i - \tilde{\mu}_k) \right]$$

Κατά την κατασκευή της εκ των προτέρων κατανομής υποθέσαμε ότι για τους μέσους θα ισχύει ότι: $\log L_{low} \leq \mu_{l,k} \leq \log(L_{high}), \log(z_{low}) \leq \mu_{z,k} \leq \log(z_{high})$. Επομένως για όλα τα k ελέγχουμε εάν οι προτεινόμενες τιμές είναι μεταξύ των προτεινόμενων ορίων. Εάν δεν ισχύει αυτό τότε απορρίπτουμε τις τιμές αυτές και προτείνονται άλλες.

- Προσομοιώνουμε ένα τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο τυχαίος αριθμός είναι μικρότερος από την πιθανότητα αποδοχής τότε αποδεχόμαστε την προτεινόμενη τιμή. Σε αντίθετη περίπτωση η τιμή του μ_k μένει ως έχει και συνεχίζουμε στο επόμενο βήμα του αλγόριθμου.
- Προτείνουμε τιμή για τον πίνακα διασπράς Σ_k από την κατανομή $Wishart_{\nu_k}(\tilde{\Sigma}_k/\nu_k)$ όπου ν_k ένα σετ βαθμών ελευθερίας και $\tilde{\Sigma}_k$ ο πίνακας κλίμακας. Μεγαλύτερες τιμές ν_k παράγουν προτεινόμενες τιμές $\hat{\Sigma}_k$ πιο κοντά στη τιμή του $\tilde{\Sigma}_k$. Επιλέγουμε βαθμούς ελευθερίας τέτοιους ώστε να γίνονται αποδεκτές οι προτεινόμενες τιμές με ποσοστό 15 – 40%
Η πιθανότητα αποδοχής σε αυτό το βήμα θα είναι η εξής:

$$\begin{aligned}
p_{acc3} &= \frac{Wishart_{\nu_k}(\tilde{\Sigma}_k|\hat{\Sigma}_k/\nu_k) p(\hat{\theta}|\log(L_{obs}), \log(z_{obs}))}{Wishart_{\nu_k}(\hat{\Sigma}_k|\tilde{\Sigma}_k/\nu_k) p(\tilde{\theta}|\log(L_{obs}), \log(z_{obs}))} \\
&= \frac{[2^{(\nu p/2)} \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma(\frac{\nu+1-i}{2})]^{-1} |\hat{\Sigma}_k|^{-\nu/2}}{[2^{(\nu p/2)} \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma(\frac{\nu+1-i}{2})]^{-1} |\tilde{\Sigma}_k|^{-\nu/2}} \\
&\times \frac{|\tilde{\Sigma}_k|^{(\nu-p-1)/2} \exp\{-\frac{1}{2}tr((\hat{\Sigma}_k)^{-1}\tilde{\Sigma}_k)\}}{|\hat{\Sigma}_k|^{(\nu-p-1)/2} \exp\{-\frac{1}{2}tr((\tilde{\Sigma}_k)^{-1}\hat{\Sigma}_k)\}} \\
&\times \frac{\prod_{k=1}^K Cauchy_2(\mu_k|\mu_0, T)Inv - Wishart_1(\hat{\Sigma}_k|A)p(I=1|\hat{\theta})^{-n}}{\prod_{k=1}^K Cauchy_2(\mu_k|\mu_0, T)Inv - Wishart_1(\tilde{\Sigma}_k|A)p(I=1|\tilde{\theta})^{-n}} \\
&\times \frac{\prod_{i \in A_{obs}} p(\log(L_i), \log(z_i)|\hat{\theta})}{\prod_{i \in A_{obs}} p(\log(L_i), \log(z_i)|\tilde{\theta})} \\
(p=2) \\
&= \left[\frac{|\tilde{\Sigma}_k|}{|\hat{\Sigma}_k|} \right]^{(\nu_k-3/2)} \exp\left\{-\frac{\nu_k}{2}tr(\hat{\Sigma}_k)^{-1}\tilde{\Sigma}_k - (\tilde{\Sigma}_k)^{-1}\hat{\Sigma}_k\right\} \\
&\times \left[\frac{\sum_{k=1}^K \frac{\pi_k}{|\tilde{\Sigma}_k|^{1/2}} \left[F(x_{high}, \mu_k, \hat{\Sigma}_k) - F(x_{low}, \mu_k, \hat{\Sigma}_k) \right]}{\sum_{k=1}^K \frac{\pi_k}{|\tilde{\Sigma}_k|^{1/2}} \left[F(x_{high}, \mu_k, \tilde{\Sigma}_k) - F(x_{low}, \mu_k, \tilde{\Sigma}_k) \right]} \right]^{-n} \\
&\times \prod_{i \in A_{obs}} \frac{\sum_{k=1}^K \left[\frac{\pi_k}{|\hat{\Sigma}_k|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu_k)^T \hat{\Sigma}_k^{-1}(x - \mu_k)\right] \right]}{\sum_{k=1}^K \left[\frac{\pi_k}{|\tilde{\Sigma}_k|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu_k)^T \tilde{\Sigma}_k^{-1}(x - \mu_k)\right] \right]}
\end{aligned}$$

Ο λογάριθμος της πιθανότητας αποδοχής θα είναι:

$$\begin{aligned}
\log p_{acc3} &= (\nu_k - \frac{3}{2}) \left[\log(|\tilde{\Sigma}_k|) - \log(|\hat{\Sigma}_k|) \right] - \frac{\nu_k}{2} tr \left[(\hat{\Sigma}_k)^{-1}\tilde{\Sigma}_k - (\tilde{\Sigma}_k)^{-1}\hat{\Sigma}_k \right] - \\
&-n \left[\log\left(\sum_{k=1}^K \frac{\pi_k}{|\tilde{\Sigma}_k|^{1/2}} (F(x_{high}, \mu_k, \hat{\Sigma}_k)) - F(x_{low}, \mu, \hat{\Sigma}_k)\right) \right] + \\
&+n \left[\log\left(\sum_{k=1}^K \frac{\pi_k}{|\tilde{\Sigma}_k|^{1/2}} F(x_{high}, \mu_k, \tilde{\Sigma}_k) - F(x_{low}, \mu_k, \tilde{\Sigma}_k)\right) \right] + \\
&+ \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log\left(\frac{\pi_k}{|\hat{\Sigma}_k|^{1/2}}\right) - \frac{1}{2}(x_i - \mu_k)^T \hat{\Sigma}_k^{-1}(x_i - \mu_k) \right] - \\
&- \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log\left(\frac{\pi_k}{|\tilde{\Sigma}_k|^{1/2}}\right) - \frac{1}{2}(x_i - \mu_k)^T \tilde{\Sigma}_k^{-1}(x_i - \mu_k) \right] =
\end{aligned}$$

$$\begin{aligned}
&= (\nu_k - \frac{3}{2}) \left[\log(|\hat{\Sigma}_k|) - \log(|\tilde{\Sigma}_k|) \right] - \frac{\nu_k}{2} \text{tr} \left[(\hat{\Sigma}_k)^{-1} \tilde{\Sigma}_k - (\tilde{\Sigma}_k)^{-1} \hat{\Sigma}_k \right] - \\
&-n \left[\log \left(\sum_{k=1}^K \frac{\pi_k}{|\Sigma_k|^{1/2}} (F(x_{high}, \mu_k, \hat{\Sigma}_k)) - F(x_{low}, \mu, \hat{\Sigma}_k) \right) \right] + \\
&+n \left[\log \left(\sum_{k=1}^K \frac{\pi_k}{|\Sigma_k|^{1/2}} F(x_{high}, \mu_k, \tilde{\Sigma}_k) - F(x_{low}, \mu_k, \tilde{\Sigma}_k) \right) \right] + \\
&+ \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log(\pi_k) - \frac{1}{2} \log(|\hat{\Sigma}_k|) - \frac{1}{2} (x_i - \mu_k)^T \hat{\Sigma}_k^{-1} (x_i - \mu_k) \right] - \\
&- \sum_{i \in A_{obs}} \sum_{k=1}^K \left[\log(\pi_k) - \frac{1}{2} \log(|\tilde{\Sigma}_k|) - \frac{1}{2} (x_i - \mu_k)^T \tilde{\Sigma}_k^{-1} (x_i - \mu_k) \right]
\end{aligned}$$

- Προσομοιώνουμε ένα τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο τυχαίος αριθμός είναι μικρότερος από την πιθανότητα αποδοχής τότε αποδεχόμαστε την προτεινόμενη τιμή. Σε αντίθετη περίπτωση η τιμή του Σ_k μένει ως έχει και συνεχίζουμε στο επόμενο βήμα του αλγόριθμου.
- Προτείνουμε τιμή για την παράμετρο μ_0 από την κανονική κατανομή $N_2(\tilde{\mu}_0, V_0)$ με διακύμανση τέτοια ώστε να γίνονται αποδεκτές τιμές σε ποσοστό 20-50%.

$$\begin{aligned}
p_{acc4} &= \prod_{k=1}^K \frac{Cauchy_2(\mu_k | \hat{\mu}_0, T)}{Cauchy_2(\mu_k | \tilde{\mu}_0, T)} \left[\frac{\int_{\log(L_{low})}^{\log(L_{high})} \int_{\log(z_{low})}^{\log(z_{high})} Cauchy_2(\mu_k | \tilde{\mu}_0, T) d\mu_k}{\int_{\log(L_{low})}^{\log(L_{high})} \int_{\log(z_{low})}^{\log(z_{high})} Cauchy_2(\mu_k | \hat{\mu}_0, T) d\mu_k} \right]^K \\
&= \prod_{k=1}^K \frac{[1 + (\mu_k - \hat{\mu}_0)^T T^{-1} (\mu_k - \hat{\mu}_0)]^{-\frac{(1+p)}{2}}}{[1 + (\mu_k - \tilde{\mu}_0)^T T^{-1} (\mu_k - \tilde{\mu}_0)]^{-\frac{(1+p)}{2}}} \left[\frac{\int_{\log(L_{low})}^{\log(L_{high})} \int_{\log(z_{low})}^{\log(z_{high})} [1 + (\mu_k - \tilde{\mu}_0)^T T^{-1} (\mu_k - \tilde{\mu}_0)]^{-\frac{(1+p)}{2}} d\mu_k}{\int_{\log(L_{low})}^{\log(L_{high})} \int_{\log(z_{low})}^{\log(z_{high})} [1 + (\mu_k - \hat{\mu}_0)^T T^{-1} (\mu_k - \hat{\mu}_0)]^{-\frac{(1+p)}{2}} d\mu_k} \right]^K
\end{aligned}$$

Το γινόμενο με τα ολοκληρώματα οφείλεται στους περιορισμούς για το μ_0 .

Ο λογάριθμος της πιθανότητας αποδοχής θα δίνεται από:

$$\begin{aligned}
\log p_{acc4} &= \\
&\sum_{k=1}^K \left[-\frac{3}{2} \log(1 + (\mu_k - \hat{\mu}_0)^T T^{-1} (\mu_k - \hat{\mu}_0)) + \frac{3}{2} \log(1 + (\mu_k - \tilde{\mu}_0)^T T^{-1} (\mu_k - \tilde{\mu}_0)) \right] + \\
&+ K \log \left[\frac{\int_{\log(L_{low})}^{\log(L_{high})} \int_{\log(z_{low})}^{\log(z_{high})} [1 + (\mu_k - \tilde{\mu}_0)^T T^{-1} (\mu_k - \tilde{\mu}_0)]^{-\frac{3}{2}} d\mu_k}{\int_{\log(L_{low})}^{\log(L_{high})} \int_{\log(z_{low})}^{\log(z_{high})} [1 + (\mu_k - \hat{\mu}_0)^T T^{-1} (\mu_k - \hat{\mu}_0)]^{-\frac{3}{2}} d\mu_k} \right]
\end{aligned}$$

Μια καλή επιλογή για το V_0 είναι ο δειγματικός πίνακας συνδιακύμανσης.

- Όπως και σε κάθε βήμα προσομοιώνουμε ένα τυχαίο αριθμό από την ομοιόμορφη κατανομή στο διάστημα $(0, 1)$. Αν ο τυχαίος αριθμός είναι μικρότερος από την πιθανότητα αποδοχής τότε αποδεχόμαστε την προτεινόμενη τιμή. Σε αντίθετη περίπτωση η τιμή του μ_0 μένει ως έχει

και συνεχίζουμε στο επόμενο βήμα του αλγόριθμου.

- Προσομοιώνουμε τιμή για τον πίνακα κλίμακας A με αλγόριθμο Gibbs Sampler.

$$\hat{A} \sim Wishart_{\nu_A}(S)$$

Με βαθμούς ελευθερίας $\nu_A = K + 3$ και $S = \left(\sum_{k=1}^K \tilde{\Sigma}_k^{-1} \right)$

- Τέλος προσομοιώνουμε με αλγόριθμο Gibbs Sampler τιμές για την παράμετρο N από αρνητική διωνυμική κατανομή με παράμετρο η και πιθανότητα επιτυχίας

$$p(I = 1|\theta) = \int_{x_{low}}^{x_{high}} p(x)dx = \sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} [F(x_{high}, \mu_k, \Sigma_k) - F(x_{low}, \mu_k, \Sigma_k)]$$

Βιβλιογραφία

- [1] Poisson Processes, John Frank Charles Kingman (1993). Oxford University Press, USA.
- [2] Poisson Point Processes, Imaging, Tracking and Sensing, Roy L. Streit (2010). Springer Editions,USA.
- [3] Modeling and Analysis of stochastic systems, Vidyadhar G. Kulkarni (1995),1st edition Chapman and Hall,USA.
- [4] The Bayesian Choice, Christian P. Robert (2007), 2nd edition.Springer Editions, New York.
- [5] Bayesian Inference, Lecture Notes, Loukia Meligkotsidou (2008)
- [6] Probability and Random Processes, Geoffrey R. Grimmet, David R. Stirzaker (2001), 3rd edition. Oxford Univeristy Press.
- [7] Θεωρία Πιθανοτήτων, Νικόλαος Δ. Παπαδάτος, Αθήνα (2006), Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.
- [8] Εισαγωγή στις στοχαστικές ανελίξεις, Ουρανία Χρυσάφινου (2004), Εκδόσεις Σοφία, Θεσσαλονίκη.
- [9] Simulation, Sheldon M. Ross (2006), 4th edition. Elsevier Academic Press.
- [10] Markov Chains and Mixing Times, David A. Levin, Yuval Peres, Elizabeth L. Wilmer, (2002). American Mathematical Society.
- [11] Finite Markov Chains and Algorithmic Applications, Olle Häggström,(2002), Cambridge University Press
- [12] Spatial Statistics and Computational Methods, Jesper Moller (2003), Springer-Verlag New York,Inc.
- [13] A flexible method of estimating luminosity functions, B.C.Kelly, X.Fan, M.Vestergaard, August 2008, AJ 682:874-895

- [14] Analysis of complete quasar samples to obtain parameters of luminosity and evolution functions, H.L.Marshall,Y.Anvi, H.Tananbaum, G.Zamorani, 1983, *Astrophysical Journal*,269:35-41
- [15] The evolution of the hard X-ray luminosity function of AGN, J.Aird, K.Nandra, E.S.Laird, A.Georgakakis, M.L.N.Ashby, P.Barmby, A.L.Coil, J-S.Huang, A.M.Koekemoer, C.C.Steidel, C.N.A.Willmer, 2009, *Astron.Soc.* 000,1-24.
- [16] Accounting for Source Uncertainties in Analyses of Astronomical Survey Data, T.J.Loredo, 2004, *Astrophysical Journal* 0409387v1