ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων

Στοχαστικές Διαδικασίες και Μοντελοποίηση

Σημειώσεις Μαθήματος

Ελισαβετ Κωνσταντινου

Σάμος, Οκτώβριος 2006

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή

1.1 Στοιχεία Πιθανοτήτων

Έστω ότι θέλουμε να πραγματοποιήσουμε ένα πείραμα του οποίου την έξοδο δεν μπορούμε να την ξέρουμε από πριν. Ωστόσο, παρόλο που δεν γνωρίζουμε την έξοδο του πειράματος, ξέρουμε όλες τις πιθανές του εξόδους. Το σύνολο όλων των πιθανών αποτελεσμάτων ενός πειράματος ονομάζεται *δειγματικός χώρος* και θα τον συμβολίζουμε με S. Για παράδειγμα, αν το πείραμά μας αποτελείται από την ρίψη ενός νομίσματος, τότε $S = \{K, \Gamma\}$ όπου K σημαίνει ότι η έξοδος του πειράματος είναι κορώνα και Γ ότι είναι γράμματα. Αν το πείραμα αποτελείται από τη ρίψη δύο νομισμάτων τότε $S = \{(K, K), (K, \Gamma), (\Gamma, K), (\Gamma, \Gamma)\}$. Ο δειγματικός χώρος μπορεί να είναι και συνεχής. Για παράδειγμα, έστω ότι στο πείραμά μας μετράμε το χρόνο λειτουργίας μιας μηχανής. Τότε $S = [0, \infty)$.

Κάθε υποσύνολο E του S ονομάζεται *γεγονός*. Για οποιάδηποτε γεγονότα E και F του δειγματικού χώρου S μπορεί να οριστεί ένα νέο γεγονός $E \cup F$ το οποίο ονομάζεται *ένωση* των E και F και αποτελείται από όλα τα αποτελέσματα του πειράματος που ανήκουν είτε στο E ή στο F ή και στο E και στο F. Ομοίως μπορεί να οριστεί το γεγονός EF ή $E \cap F$ το οποίο ονομάζεται *τομή* των E και F και αποτελείται από τα αποτελέσματα του πειράματος που ανήκουν και στο E και στο F. Αν $EF = \emptyset$ τότε τα γεγονότα E και F ονομάζονται *αμοιβαίως αποκλειόμενα*. Με τον ίδιο τρόπο που ορίζεται η ένωση και η τομή δύο γεγονότων μπορεί να οριστεί και για περισσότερα. Αν $E_1, E_2, \dots$ είναι ξεχωριστά γεγονότα, τότε με $\bigcup_{n=1}^{\infty} E_n$ και $\bigcap_{n=1}^{\infty} E_n$ θα συμβολίζεται η ένωση και η τομή τους αντίστοιχα. Τέλος, με E^c θα συμβολίζουμε το *συμπλήρωμα* του E, δηλαδή το γεγονός που αποτελείται από όλα τα αποτελέσματα του πειράματος που δεν ανήκουν στο E.

Για κάθε γεγονός E του δειγματικού χώρου S ενός πειράματος ορίζεται ένας αριθμός $P(E)$ ο οποίος ικανοποιεί τις εξής τρεις συνθήκες:

1. $0 \leq P(E) \leq 1$

2. $P(S) = 1$

3. Για κάθε ακολουθία γεγονότων $E_1, E_2, \ldots$ που είναι αμοιβαίως αποκλειόμενα, δηλαδή $E_n E_m = \emptyset$ αν $n \neq m$, τότε $P(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n)$.

Ο αριθμός αυτός θα λέμε ότι αποτελεί την πιθανότητα του γεγονότος E. Ισχύουν οι ακόλουθες σχέσεις:

$$P(E) = 1 - P(E^c)$$

$$P(E \cup F) = P(E) + P(F) - P(EF)$$

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)$$

$$P(E_1 \cup E_2 \cup \ldots \cup E_n) = \sum_i P(E_i) - \sum_{i<j} P(E_i E_j) + \sum_{i<j<k} P(E_i E_k E_j) -$$

$$\sum_{i<j<k<l} P(EiE_j E_k E_l) + \ldots + (-1)^{n+1} P(E_1 E_2 \ldots E_n)$$

1.2 Δεσμευμένες Πιθανότητες

Έστω ότι ρίχνουμε δύο ζάρια και παρατηρούμε ότι το πρώτο έφερε 4. Με δεδομένη αυτή την πληροφορία, ποια είναι η πιθανότητα το άθροισμα των δύο ζαριών να είναι έξι;Προφανώς χωρίς αυτή την πληροφορία θα έπρεπε να εξετάσουμε και τα 36 πιθανά αποτελέσματα του πειράματος. Τώρα οι πιθανές έξοδοι του πειράματός μας είναι τα ζευγάρια (4,1), (4,2), (4,3), (4,4), (4,5) και (4,6) και η πιθανότητα που ψάχνουμε είναι ίση με $\frac{1}{6}$. Αν θεωρήσουμε ότι E είναι το γεγονός το άθροισμα των ζαριών να είναι έξι και F το γεγονός το πρώτο ζάρι να φέρει 4, τότε η πιθανότητα που μόλις υπολογίσαμε ονομάζεται *δεσμευμένη πιθανότητα* να πραγματοποιηθεί το E δοθέντος ότι έχει ήδη πραγματοποιηθεί το F και συμβολίζεται με $P(E|F)$.

Ο γενικός τύπος για τον υπολογισμό της δεσμευμένης πιθανότητας $P(E|F)$ για κάθε ζευγάρι γεγονότων E και F μπορεί να προκύψει με τον ίδιο τρόπο που έγινε ο υπολογισμός της πιθανότητας στο προηγούμενο παράδειγμα. Συγκεκριμένα, αν το γεγονός F ισχύει, τότε για να πραγματοποιηθεί και το E θα πρέπει η έξοδος του πειράματος να ανήκει στο EF. Επειδή όμως γνωρίζουμε ότι το F έχει ήδη πραγματοποιηθεί, προκύπτει ότι το F είναι ο νέος μας δειγματικός χώρος και έτσι η πιθανότητα να πραγματοποιηθεί το γεγονός $E|F$ είναι

$$P(E|F) = \frac{P(EF)}{P(F)}.$$

Προφανώς η εξίσωση αυτή είναι καλά ορισμένη αν $P(F) > 0$.

Παράδειγμα 1.2.1 *Μια οικογένεια έχει δύο παιδιά. Ποια είναι η πιθανότητα και τα δύο παιδιά να είναι αγόρια δοθέντος ότι το ένα τουλάχιστον είναι;*

Λύση. Έστω ότι $S = \{(\alpha,\alpha),(\alpha,\kappa),(\kappa,\alpha),(\kappa,\kappa)\}$ είναι ο δειγματικός χώρος (α σημαίνει αγόρι, κ κορίτσι), B το γεγονός και τα δύο παιδιά να είναι αγόρια και A το γεγονός τουλάχιστον το ένα παιδί να είναι αγόρι. Τότε η πιθανότητα που ζητείται είναι ίση με

$$P(B|A) = \frac{P(BA)}{P(A)} = \frac{P((\alpha,\alpha))}{P((\alpha,\alpha),(\alpha,\kappa),(\kappa,\alpha))} = \frac{1/4}{3/4} = \frac{1}{3}.$$

∎

Δύο γεγονότα E και F θα καλούνται *ανεξάρτητα* αν $P(EF) = P(E)P(F)$. Επομένως, αν τα E και F είναι ανεξάρτητα τότε $P(E|F) = P(E)$ και $P(F|E) = P(F)$. Δηλαδή η γνώση ενός εκ των δύο γεγονότων δεν επηρεάζει την πιθανότητα πραγματοποίησης του άλλου.

Τύπος του Bayes: Έστω E και F δύο γεγονότα. Μπορούμε να εκφράσουμε το E ως $E = EF \cup EF^c$. Επειδή τα EF και EF^c είναι αμοιβαίως αποκλειόμενα ισχύει:

$$P(E) = P(EF) + P(EF^c) = P(E|F)P(F) + P(E|F^c)P(F^c).$$

Η παραπάνω εξίσωση μπορεί να γενικευτεί με τον ακόλουθο τρόπο. Έστω ότι $F_1, F_2, \ldots, F_n$ είναι αμοιβαίως αποκλειόμενα γεγονότα τέτοια ώστε $\bigcup_{i=1}^n F_i = S$. Κάθε γεγονός E είναι τότε ίσο με $E = \bigcup_{i=1}^n EF_i$, όπου τα γεγονότα EF_i είναι αμοιβαίως αποκλειόμενα. Ισχύει:

$$P(E) = \sum_{i=1}^n P(EF_i) = \sum_{i=1}^n P(E|F_i)P(F_i).$$

Από την παραπάνω εξίσωση προκύπτει ότι:

$$P(F_j|E) = \frac{P(EF_j)}{P(E)} = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}.$$

Η εξίσωση αυτή αποτελεί τον τύπο του Bayes.

Κεφάλαιο 2

Τυχαίες Μεταβλητές

2.1 Εισαγωγή

Συχνά αυτό που μας ενδιαφέρει κατά την διεξαγωγή ενός πειράματος είναι κάποιες συναρτήσεις των αποτελεσμάτων του και όχι τα ίδια τα αποτελέσματα. Για παράδειγμα, αν το πείραμα αποτελείται από την ρίψη δύο ζαριών συχνά ενδιαφερόμαστε για το άθροισμα των ζαριών και όχι για το τί θα φέρει το καθένα. Οι συναρτήσεις που ορίζονται πάνω στον δειγματικό χώρο (δηλαδή έχουν πεδίο ορισμού τα στοιχεία του δειγματικού χώρου) ονομάζονται *τυχαίες μεταβλητές*. Μια τυχαία μεταβλητή μπορεί να πάρει διάφορες τιμές, σε κάθε μια από τις οποίες μπορεί να τεθεί μια συγκεκριμένη πιθανότητα.

Παράδειγμα 2.1.1 *Έστω X μια τυχαία μεταβλητή που ορίζεται ως το άθροισμα δύο ζαριών. Τότε:*

$$P\{X=2\} = P\{(1,1)\} = \frac{1}{36}$$

$$P\{X=3\} = P\{(1,2),(2,1)\} = \frac{2}{36}$$

$$P\{X=4\} = P\{(1,3),(2,2),(3,1)\} = \frac{3}{36}$$

$$P\{X=5\} = P\{(1,4),(2,3),(3,2),(4,1)\} = \frac{4}{36}$$

$$P\{X=6\} = P\{(1,5),(2,4),(3,3),(4,2),(5,1)\} = \frac{5}{36}$$

$$P\{X=7\} = P\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} = \frac{6}{36}$$

$$P\{X=8\} = P\{(2,6),(3,5),(4,4),(5,3),(6,2)\} = \frac{5}{36}$$

$$P\{X=9\} = P\{(3,6),(4,5),(5,4),(6,3)\} = \frac{4}{36}$$

$$P\{X=10\}=P\{(4,6),(5,5),(6,4)\}=\frac{3}{36}$$

$$P\{X=11\}=P\{(5,6),(6,5)\}=\frac{2}{36}$$

$$P\{X=12\}=P\{(6,6)\}=\frac{1}{36}$$

Με άλλα λόγια η τυχαία μεταβλητή X μπορεί να πάρει οποιαδήποτε τιμή μεταξύ του 2 και του 12, και η πιθανότητα να πάρει μία από αυτές τις τιμές δίνεται από τις παραπάνω ισότητες. Ισχύει επίσης ότι

$$P\left(\bigcup_{n=2}^{12}\{X=n\}\right)=\sum_{n=2}^{12}P\{X=n\}=1.$$

Παράδειγμα 2.1.2 *Έστω ότι ρίχνουμε συνεχώς ένα νόμισμα μέχρι να εμφανιστεί κορώνα. Σε κάθε ρίψη η πιθανότητα να εμφανιστεί κορώνα είναι ίση με p. Αν N είναι μια τυχαία μεταβλητή που προσδιορίζει των αριθμό των πιθανών ρίψεων μέχρι να έρθει κορώνα, ποιο είναι το πεδίο ορισμού της και ποιες οι αντίστοιχες πιθανότητες;*

Λύση. Προφανώς η τυχαία μεταβλητή N μπορεί να πάρει τις τιμές $1,2,3,\ldots$ και οι αντίστοιχες πιθανότητες για αυτές τις τιμές είναι:

$$P\{N=1\}=P\{K\}=p$$

$$P\{N=2\}=P\{\Gamma,K\}=(1-p)p$$

$$P\{N=3\}=P\{\Gamma,\Gamma,K\}=(1-p)^2p$$

$$\vdots$$

$$P\{N=n\}=P\{\Gamma,\Gamma,\ldots,K\}=(1-p)^{n-1}p$$

Παρατηρείστε ότι

$$P\left(\bigcup_{n=1}^{\infty}\{N=n\}\right)=\sum_{n=1}^{\infty}P\{N=n\}$$

$$=p\sum_{n=1}^{\infty}(1-p)^{n-1}=\frac{p}{1-(1-p)}=1$$

∎

Οι τυχαίες μεταβλητές που παίρνουν είτε ένα πεπερασμένο ή ένα μετρήσιμο αριθμό πιθανών τιμών ονομάζονται *διακριτές*. Τέτοιες ήταν για παράδειγμα οι τυχαίες μεταβλητές που χρησιμοποιήθηκαν στα προηγούμενα παραδείγματα. Ωστόσο υπάρχουν τυχαίες μεταβλητές που μπορούν να πάρουν οποιαδήποτε τιμή σε κάποιο διάστημα (a,b). Αυτές καλούνται *συνεχείς* τυχαίες μεταβλητές.

Η *συνάρτηση κατανομής* $F()$ μιας τυχαίας μεταβλητής ορίζεται για κάθε πραγματικό αριθμό b, με $-\infty < b < \infty$, από την σχέση

$$F(b) = P\{X \leq b\}.$$

Δηλαδή, $F(b)$ είναι η πιθανότητα η τυχαία μεταβλητή X να πάρει τιμή μικρότερη ή ίση από το b. Για την συνάρτηση κατανομής ισχύουν οι παρακάτω ιδιότητες:

- Η $F(b)$ είναι μια μη φθίνουσα συνάρτηση του b
- $\lim_{b\to\infty} F(b) = F(\infty) = 1$ και
- $\lim_{b\to-\infty} F(b) = F(-\infty) = 0$.

Ισχύει επίσης ότι

$$P\{a < X \leq b\} = F(b) - F(a)$$

για κάθε $a < b$ και

$$P\{X < b\} = \lim_{h\to 0^+} P\{X \leq b - h\} = \lim_{h\to 0^+} F(b-h).$$

2.2 Διακριτές Τυχαίες Μεταβλητές

Για κάθε διακριτή τυχαία μεταβλητή X, ορίζεται η *συνάρτηση πιθανότητας* $p(\alpha)$ του X ως εξής:

$$p(\alpha) = P\{X = \alpha\}.$$

Αν η τυχαία μεταβλητή X παίρνει τις τιμές $x_1, x_2, \ldots$ τότε:

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

Η συνάρτηση κατανομής $F()$ μπορεί να εκφραστεί με τη βοήθεια της συνάρτησης πιθανότητας $p()$ ως:

$$F(\alpha) = \sum_{\forall x_i \leq \alpha} p(x_i)$$

Για παράδειγμα, αν η τυχαία μεταβλητή X μπορεί να πάρει τις τιμές 1, 2 και 3 με $p(1) = \frac{1}{2}$, $p(2) = \frac{1}{3}$, $p(3) = \frac{1}{6}$, τότε η συνάρτηση κατανομής F του X είναι ίση με:

$$F(\alpha) = \begin{cases} 0 & \alpha < 1 \\ \frac{1}{2} & 1 \leq \alpha < 2 \\ \frac{5}{6} & 2 \leq \alpha < 3 \\ 1 & 3 \leq \alpha \end{cases}$$

Στις επόμενες ενότητες ορίζονται μερικές βασικές κατηγορίες διακριτών τυχαίων μεταβλητών.

2.2.1 Η Τυχαία Μεταβλητή Bernoulli

Ας υποθέσουμε ότι πραγματοποιούμε ένα πείραμα του οποίου το αποτέλεσμα είναι είτε "επιτυχία" ή "αποτυχία". Αν θεωρήσουμε μία τυχαία μεταβλητή X η οποία θα παίρνει την τιμή 1 αν έχουμε επιτυχία και την τιμή 0 αν έχουμε αποτυχία, τότε η συνάρτηση πιθανότητας της X θα είναι:

$$p(0) = P\{X = 0\} = 1 - p$$

$$p(1) = P\{X = 1\} = p \tag{2.1}$$

όπου p, με $0 \leq p \leq 1$, είναι η πιθανότητα να έχουμε επιτυχία. Επομένως, μία τυχαία μεταβλητή X θα λέμε ότι είναι *τυχαία μεταβλητή Bernoulli* αν η συνάρτηση πιθανότητάς της δίνεται από την εξίσωση (2.1).

2.2.2 Η Διωνυμική Τυχαία Μεταβλητή

Έστω ότι θέλουμε να πραγματοποιήσουμε n ανεξάρτητες δοκιμές, κάθε μία από τις οποίες μπορεί να οδηγήσει σε "επιτυχία" με πιθανότητα p και σε "αποτυχία" με πιθανότητα $1-p$. Αν η τυχαία μεταβλητή X αντιπροσωπεύει τον αριθμό των επιτυχιών που προέκυψαν στις n δοκιμές, τότε λέμε ότι η X είναι *διωνυμική τυχαία μεταβλητή* με παραμέτρους (n, p).

Η συνάρτηση πιθανότητας της διωνυμικής τυχαίας μεταβλητής δίνεται από τη σχέση:

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}$$

όπου $i = 0, 1, \ldots, n$ και $\binom{n}{i} = \frac{n!}{(n-i)!i!}$.

Παράδειγμα 2.2.1 *Γνωρίζουμε ότι σε ένα εργοστάσιο παραγωγής λαμπτήρων κάθε λαμπτήρας που παράγεται έχει πιθανότητα ίση με 0.1 να είναι ελαττωματικός. Ποια η πιθανότητα σε ένα δείγμα τριών λαμπτήρων, το πολύ ένας να είναι ελαττωματικός;*

Λύση. Αν X είναι ο αριθμός των ελαττωματικών λαμπτήρων σε ένα δείγμα, τότε X είναι μια διωνυμική τυχαία μεταβλητή με παραμέτρους $(3, 0.1)$. Επομένως, η πιθανότητα που ζητείται είναι ίση με

$$P\{X = 0\} + P\{X = 1\} = \binom{3}{0}(0.1)^0(0.9)^3 + \binom{3}{1}(0.1)^1(0.9)^2 = 0.972$$

■

2.2.3 Η Γεωμετρική Τυχαία Μεταβλητή

Ας υποθέσουμε ότι σε ένα πείραμα διεξάγουμε ανεξάρτητες δοκιμές έως ότου έχουμε ''επιτυχία'' και γνωρίζουμε ότι κάθε μία από τις δοκιμές έχει πιθανότητα p να οδηγήσει σε ''επιτυχία''. Αν η τυχαία μεταβλητή X δηλώνει τον αριθμό των δοκιμών που απαιτούνται μέχρι να έχουμε την πρώτη επιτυχία, τότε η X ονομάζεται *γεωμετρική τυχαία μεταβλητή* με παράμετρο p. Η συνάρτηση πιθανότητας μιας γεωμετρικής τυχαίας μεταβλητής είναι ίση με:

$$p(n) = P\{X = n\} = (1-p)^{n-1}p$$

όπου $n = 1, 2, \ldots$.

2.2.4 Η Τυχαία Μεταβλητή Poisson

Μια τυχαία μεταβλητή X η οποία μπορεί να πάρει τις τιμές $0, 1, 2, \ldots$, θα ονομάζεται *τυχαία μεταβλητή Poisson* με παράμετρο λ αν για κάποιο $\lambda > 0$ ισχύει

$$p(i) = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!}$$

με $i = 0, 1, \ldots$.

Μια σημαντική ιδιότητα των τυχαίων μεταβλητών Poisson είναι ότι μπορούν να χρησιμοποιηθούν για την προσέγγιση διωνυμικών τυχαίων μεταβλητών όταν το n είναι πολύ μεγάλο και το p πολύ μικρό. Για να γίνει αυτό κατανοητό, ας υποθέσουμε ότι X είναι μια διωνυμική τυχαία μεταβλητή με παραμέτρους (n, p) και έστω $\lambda = np$. Τότε:

$$\begin{aligned} P\{X = i\} &= \frac{n!}{(n-i)!i!}p^i(1-p)^{n-i} \\ &= \frac{n!}{(n-i)!i!}\left(\frac{\lambda}{n}\right)^i\left(1-\frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)\ldots(n-i+1)}{n^i}\frac{\lambda^i}{i!}\frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}. \end{aligned}$$

Όταν το n είναι μεγάλο και το p μικρό ισχύει ότι $\left(1-\frac{\lambda}{n}\right)^n \approx e^{-\lambda}$, $\left(1-\frac{\lambda}{n}\right)^i \approx 1$ και $\frac{n(n-1)\ldots(n-i+1)}{n^i} \approx 1$. Επομένως, για n μεγάλο και p μικρό ισχύει $P\{X = i\} \approx e^{-\lambda}\frac{\lambda^i}{i!}$.

Παράδειγμα 2.2.2 *Έστω ότι ο αριθμός των τυπογραφικών λαθών σε μία σελίδα ενός βιβλίου ακολουθεί κατανομή Poisson με παράμετρο $\lambda = 1$. Υπολογίστε την πιθανότητα να υπάρχει τουλάχιστον ένα λάθος σε αυτή τη σελίδα.*

Λύση. Η πιθανότητα που ζητάμε είναι η $P\{X \geq 1\} = 1 - P\{X = 0\} = 1 - e^{-1} \approx 0.633$. ∎

2.3 Συνεχείς Τυχαίες Μεταβλητές

Στην ενότητα αυτή θα μελετήσουμε τυχαίες μεταβλητές οι οποίες παίρνουν τιμές από ένα μη μετρήσιμο σύνολο. Πιο συγκεκριμένα, μια τυχαία μεταβλητή X θα λέμε ότι είναι συνεχής αν υπάρχει μια μη αρνητική συνάρτηση $f(x)$, ορισμένη για κάθε πραγματικό αριθμό $x \in (-\infty, \infty)$, που έχει την ιδιότητα

$$P\{X \in B\} = \int_B f(x)dx \tag{2.2}$$

για κάθε υποσύνολο B των πραγματικών αριθμών. Η συνάρτηση $f(x)$ ονομάζεται *συνάρτηση πυκνότητας πιθανότητας* της τυχαίας μεταβλητής X.

Προφανώς, η συνάρτηση $f(x)$ θα πρέπει να ικανοποιεί τη σχέση

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx.$$

Κάθε ''ερώτηση'' που μπορεί να γίνει για την πιθανότητα η τυχαία μεταβλητή X να ανήκει σε ένα συγκεκριμένο διάστημα των πραγματικών αριθμών, μπορεί να απαντηθεί μέσω της συνάρτησης $f(x)$. Για παράδειγμα, αν $B = [a, b]$ τότε από την εξίσωση (2.2) προκύπτει ότι:

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx. \tag{2.3}$$

Αν θέσουμε $a = b$ στην προηγούμενη εξίσωση έχουμε ότι:

$$P\{X = a\} = \int_a^a f(x)dx = 0.$$

Δηλαδή η πιθανότητα μια συνεχής τυχαία μεταβλητή X να πάρει μια συγκεκριμένη τιμή είναι ίση με το μηδέν.

Η συνάρτηση κατανομής $F()$ μιας συνεχούς τυχαίας μεταβλητής X ορίζεται ως εξής:

$$F(a) = P\{X \in (-\infty, a]\} = \int_{-\infty}^{a} f(x)dx.$$

Παραγωγίζοντας και τα δύο μέλη της εξίσωσης έχουμε:

$$\frac{d}{da}F(a) = f(a).$$

Δηλαδή, η συνάρτηση πυκνότητας πιθανότητας είναι η παράγωγος της συνάρτησης κατανομής.

Στη συνέχεια της ενότητας παρουσιάζονται οι σημαντικότερες κατηγορίες συνεχών τυχαίων μεταβλητών.

2.3.1 Η Ομοιόμορφη Τυχαία Μεταβλητή

Μια τυχαία μεταβλητή θα λέμε ότι είναι ομοιόμορφα κατανεμημένη στο διάστημα (0,1) αν η συνάρτηση πυκνότητας πιθανότητάς της είναι ίση με:

$$f(x) = \begin{cases} 1 & \text{αν } 0 < x < 1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Παρατηρείστε ότι η παραπάνω συνάρτηση είναι μια συνάρτηση πυκνότητας πιθανότητας αφού $f(x) \geq 0$ και

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 dx = 1.$$

Δεδομένου ότι η συνάρτηση $f(x)$ είναι θετική μόνο όταν $x \in (0,1)$, προκύπτει ότι μια τυχαία μεταβλητή X που ακολουθεί την ομοιόμορφη κατανομή μπορεί να πάρει τιμή στο διάστημα $(0,1)$ μόνο. Επίσης, επειδή η $f(x)$ είναι σταθερή για $x \in (0,1)$, η τυχαία μεταβλητή X είναι εξίσου πιθανό να βρίσκεται ''κοντά'' σε οποιαδήποτε τιμή στο $(0,1)$. Για να γίνει αυτό πιο κατανοητό, παρατηρήστε ότι για οποιοδήποτε $0 < a < b < 1$,

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx = b - a.$$

Γενικότερα, μια τυχαία μεταβλητή X καλείται ομοιόμορφη στο διάστημα (a,b) αν η συνάρτηση πυκνότητας πιθανότητάς της είναι ίση με:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{αν } a < x < b \\ 0 & \text{διαφορετικά} \end{cases}$$

Παράδειγμα 2.3.1 *Υπολογίστε την συνάρτηση κατανομής μιας ομοιόμορφα κατανεμημένης τυχαίας μεταβλητής X στο διάστημα (a,b).*

Λύση. Γνωρίζουμε ότι $F(c) = \int_{-\infty}^{c} f(x)dx$. Επομένως

$$F(c) = \begin{cases} 0 & \text{αν } c \leq a \\ \frac{c-a}{b-a} & \text{αν } a < c < b \\ 1 & \text{αν } c \geq b \end{cases}$$

∎

2.3.2 Η Εκθετική Τυχαία Μεταβλητή

Μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας ίση με:

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

όπου $\lambda > 0$, ονομάζεται *εκθετική τυχαία μεταβλητή* με παράμετρο λ. Η συνάρτηση κατανομής για την εκθετική τυχαία μεταβλητή είναι:

$$F(a) = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}$$

για $a \geq 0$. Περισσότερες λεπτομέρειες για την εκθετική κατανομή θα δοθούν σε επόμενο κεφάλαιο.

2.3.3 Η Τυχαία Μεταβλητή Γάμμα

Μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας ίση με:

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{a-1}}{\Gamma(a)} & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases}$$

όπου $\lambda > 0$, $a > 0$ ονομάζεται *τυχαία μεταβλητή γάμμα* με παραμέτρους λ, a. Ο αριθμός $\Gamma(a)$ ορίζεται από τη σχέση $\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx$. Αποδεικνύεται εύκολα με επαγωγή ότι για ακέραιο a ισχύει $\Gamma(a) = (a-1)!$.

2.3.4 Η Κανονική Τυχαία Μεταβλητή

Μία τυχαία μεταβλητή X θα λέμε ότι είναι *κανονική τυχαία μεταβλητή* ή ότι ακολουθεί *κανονική κατανομή* με παραμέτρους μ και σ^2 αν η συνάρτηση πυκνότητας πιθανότητας της X είναι ίση με:

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

όπου $-\infty < x < \infty$. Η συνάρτηση αυτή είναι συμμετρική γύρω από το μ.

Αν X είναι μια τυχαία μεταβλητή με παραμέτρους μ και σ^2, τότε η τυχαία μεταβλητή $Y = aX + b$ είναι επίσης κανονική με παραμέτρους $a\mu + b$ και $a^2\sigma^2$. Για να αποδείξουμε αυτόν τον ισχυρισμό, ας υποθέσουμε αρχικά ότι $a > 0$ και ότι $F_Y()$ είναι η συνάρτηση κατανομής της τυχαίας μεταβλητής Y. Τότε:

$$F_Y(c) = P\{Y \leq c\} = P\{aX + b \leq c\} = P\left\{X \leq \frac{c-b}{a}\right\}$$

$$= F_X\left(\frac{c-b}{a}\right) = \int_{-\infty}^{\frac{c-b}{a}} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} dx$$

$$= \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}a\sigma} exp\left\{\frac{-(u-(a\mu+b))^2}{2(a\sigma)^2}\right\} du$$

όπου η τελευταία εξίσωση προέκυψε με μία αλλαγή στις μεταβλητές (συγκεκριμένα θέτοντας $u = ax + b$). Επειδή όμως ισχύει ότι $F_Y(a) = \int_{-\infty}^a f_Y(u)du$, προκύπτει ότι η συνάρτηση πιθανότητας $f_Y()$ της τυχαίας μεταβλητής Y είναι:

$$f_Y(u) = \frac{1}{\sqrt{2\pi}a\sigma} exp\left\{\frac{-(u-(a\mu+b))^2}{2(a\sigma)^2}\right\}$$

όπου $-\infty < u < \infty$. Επομένως, η τυχαία μεταβλητή Y είναι κανονική με παραμέτρους $a\mu + b$ και $a^2\sigma^2$. Παρόμοια είναι η απόδειξη και για την περίπτωση που $a < 0$.

Συνέπεια του παραπάνω αποτελέσματος είναι ότι αν X είναι μια κανονική τυχαία μεταβλητή με παραμέτρους μ και σ^2 τότε η $Y = \frac{(X-\mu)}{\sigma}$ είναι επίσης κανονική με παραμέτρους 0 και 1. Μια τέτοια τυχαία μεταβλητή καλείται *τυπική ή μοναδιαία* κανονική μεταβλητή.

2.4 Αναμενόμενη Τιμή μιας Τυχαίας Μεταβλητής

2.4.1 Διακριτές Τυχαίες Μεταβλητές

Αν X είναι μια διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας $p(x)$, τότε η αναμενόμενη ή μέση τιμή της X ορίζεται ως εξής:

$$E[X] = \sum_{x:p(x)>0} xp(x).$$

Με άλλα λόγια, η αναμενόμενη τιμή της X είναι ο μέσος όρος των πιθανών τιμών που μπορεί να πάρει η μεταβλητή, όπου όμως κάθε τέτοια τιμή συνεισφέρει στον μέσο όρο ανάλογα με το πόσο πιθανό είναι η X να πάρει αυτή τη τιμή.

Παράδειγμα 2.4.1 *Υπολογίστε την αναμενόμενη τιμή $E[X]$ όταν η X είναι μια τυχαία μεταβλητή Bernoulli με παράμετρο p.*

Λύση. Ισχύει ότι $E[X] = \sum_{x:p(x)>0} xp(x) = 0p(0) + 1p(1)$. Αλλά $p(0) = 1-p$ και $p(1) = p$. Επομένως $E[X] = p$. ∎

Παράδειγμα 2.4.2 *Υπολογίστε την αναμενόμενη τιμή $E[X]$ όταν η X είναι μια τυχαία μεταβλητή Poisson με παράμετρο λ.*

Λύση. Έχουμε ότι $E[X] = \sum_{i=0}^{\infty} \frac{ie^{-\lambda}\lambda^i}{i!} = \sum_{i=1}^{\infty} \frac{e^{-\lambda}\lambda^i}{(i-1)!} = \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!}$ $= \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$. ∎

2.4.2 Συνεχείς Τυχαίες Μεταβλητές

Αν X είναι μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$, τότε η αναμενόμενη ή μέση τιμή της X ορίζεται από την σχέση:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx.$$

Παράδειγμα 2.4.3 *Υπολογίστε την αναμενόμενη τιμή μιας τυχαίας μεταβλητής ομοιόμορφα κατανεμημένης στο διάστημα (a, b).*

Λύση. Ισχύει ότι $E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$. ∎

Παράδειγμα 2.4.4 *Υπολογίστε την αναμενόμενη τιμή $E[X]$ όταν η X είναι μία κανονική τυχαία μεταβλητή με παραμέτρους μ και σ^2.*

Λύση. Ισχύει ότι $E[X] = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x e^{-(x-\mu)^2/2\sigma^2}dx$. Θέτοντας $x = (x-\mu)+\mu$ και στη συνέχεια $y = x-\mu$ έχουμε ότι:

$$E[X] = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-\mu)e^{-(x-\mu)^2/2\sigma^2}dx + \mu\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} y e^{-y^2/2\sigma^2}dy + \mu\int_{-\infty}^{\infty} f(x)dx.$$

Λόγω συμμετρίας, το πρώτο ολοκλήρωμα είναι ίσο με το μηδέν. Επομένως, $E[X] = \mu\int_{-\infty}^{\infty} f(x)dx = \mu$.

∎

2.4.3 Αναμενόμενη Τιμή για Συναρτήσεις Τυχαίων Μεταβλητών

Έστω ότι μας δίνεται μια τυχαία μεταβλητή X και η κατανομή πιθανότητάς της. Ας υποθέσουμε επίσης ότι θέλουμε να υπολογίσουμε την αναμενόμενη τιμή μιας συνάρτησης της X, έστω $g(X)$. Ένας απλός τρόπος για τον υπολογισμό της τιμής αυτής είναι να βρούμε την κατανομή της τυχαίας μεταβλητής $g(X)$ (μέσω της γνώσης μας για την κατανομή της X) και στη συνέχεια να υπολογίσουμε την τιμή $E[g(X)]$ βάση των ορισμών που δόθηκαν στις προηγούμενες ενότητες.

Παράδειγμα 2.4.5 *Έστω ότι X είναι μια διακριτή τυχαία μεταβλητή με την ακόλουθη συνάρτηση πιθανότητας:$p(0) = 0.2$, $p(1) = 0.5$, $p(2) = 0.3$. Υπολογίστε την αναμενόμενη τιμή $E\left[X^2\right]$.*

Λύση. Θέτοντας $Y = X^2$ έχουμε ότι η Y είναι μια τυχαία μεταβλητή που μπορεί να πάρει τις τιμές 0^2, 1^2 και 2^2 με αντίστοιχες πιθανότητες:

$$p_Y(0) = P\left\{Y = 0^2\right\} = 0.2$$

$$p_Y(1) = P\left\{Y = 1^2\right\} = 0.5$$

$$p_Y(4) = P\left\{Y = 2^2\right\} = 0.3.$$

Άρα, $E\left[X^2\right] = E[Y] = 0(0.2) + 1(0.5) + 4(0.3) = 1.7$. Παρατηρείστε ότι $1.7 = E\left[X^2\right] \neq (E[X])^2 = 1.21$. ∎

Παράδειγμα 2.4.6 *Έστω X μια τυχαία μεταβλητή ομοιόμορφα κατανεμημένη στο διάστημα $(0,1)$. Υπολογίστε το $E\left[X^3\right]$.*

Λύση. Έστω $Y = X^3$. Για $0 < a < 1$ έχουμε ότι $F_Y(a) = P\{Y \leq a\} = P\{X^3 \leq a\} = P\{X \leq a^{1/3}\} = a^{1/3}$. Παραγωγίζοντας την $F_Y(a)$ υπολογίζουμε τη συνάρτηση πυκνότητας πιθανότητας $f_Y(a)$. Συγκεκριμένα, $f_Y(a) = \frac{1}{3}a^{-2/3}$, $0 < a < 1$. Επομένως, $E\left[X^3\right] = E[Y] = \int_{-\infty}^{\infty} a f_Y(a)da = \frac{1}{3}\int_0^1 a^{1/3}da = \frac{1}{3}\frac{3}{4} = \frac{1}{4}$. ∎

Ο τρόπος που περιγράφηκε παραπάνω για τον υπολογισμό της αναμενόμενης τιμής μιας συνάρτησης της X δεν είναι ιδιαίτερα αποδοτικός στις περισσότερες περιπτώσεις. Η ακόλουθη Πρόταση δείχνει πως μπορούμε να υπολογίσουμε την αναμενόμενη τιμή μιας συνάρτησης $g(X)$ της τυχαίας μεταβλητής X χωρίς να χρειάζεται να προσδιορίσουμε πρώτα την κατανομή της.

Πρόταση 2.4.1 *(α) Αν X είναι μια διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας $p(x)$, τότε για κάθε συνάρτηση g ισχύει:*

$$E[g(X)] = \sum_{x:p(x)>0} g(x)p(x)$$

(β) Αν X είναι μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$, τότε για κάθε συνάρτηση g ισχύει:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Παράδειγμα 2.4.7 *Εφαρμόζοντας την πρόταση αυτή στο παράδειγμα 2.4.6 έχουμε κατευθείαν ότι $E\left[X^3\right] = \int_0^1 x^3 dx = \frac{1}{4}$.*

Πόρισμα 2.4.1 *Αν a και b είναι σταθερές, τότε*

$$E[aX + b] = aE[X] + b.$$

Απόδειξη. Αφήνεται ως άσκηση. ∎

Εκτός από την αναμενόμενη τιμή, μια άλλη ποσότητα που μας ενδιαφέρει συχνά είναι η *διασπορά* μιας τυχαίας μεταβλητής X. Αυτή συμβολίζεται ως $Var(X)$ και είναι ίση με:

$$Var(X) = E\left[(X - E[X])^2\right].$$

Δηλαδή, η διασπορά της X μετράει τη μέση τιμή του τετραγώνου της απόκλισης της X από την αναμενόμενη τιμή της.

Ας υποθέσουμε ότι η X είναι μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f()$ και αναμενόμενη τιμή $E[X] = \mu$. Τότε:

$$Var(X) = E\left[(X-\mu)^2\right] = E\left[X^2 - 2\mu X + \mu^2\right] = \int_{-\infty}^{\infty}(x^2 - 2\mu x + \mu^2)f(x)dx$$

$$= \int_{-\infty}^{\infty} x^2 f(x)dx - 2\mu\int_{-\infty}^{\infty} x f(x)dx + \mu^2\int_{-\infty}^{\infty} f(x)dx$$

$$= E\left[X^2\right] - 2\mu\mu + \mu^2 = E\left[X^2\right] - \mu^2.$$

Μια παρόμοια απόδειξη ισχύει και για την περίπτωση των διακριτών τυχαίων μεταβλητών. Επομένως, προκύπτει ότι για όλες τις τυχαίες μεταβλητές ισχύει:

$$Var(X) = E\left[X^2\right] - \left(E\left[X\right]\right)^2.$$

2.5 Από Κοινού Κατανομή Τυχαίων Μεταβλητών

2.5.1 Από Κοινού Συναρτήσεις Κατανομής

Μέχρι στιγμής έχουμε ασχοληθεί με την κατανομή μίας μόνο τυχαίας μεταβλητής. Ωστόσο, πολλές φορές χρειάζεται ο υπολογισμός κάποιας πιθανότητας που αφορά σε δύο ή περισσότερες τυχαίες μεταβλητές. Για το λόγο αυτό, ορίζουμε για οποιεσδήποτε τυχαίες μεταβλητές X και Y την *από κοινού συνάρτηση κατανομής* ως εξής:

$$F(a, b) = P\left\{X \leq a, Y \leq b\right\}$$

με $-\infty < a, b < \infty$. Η συνάρτηση κατανομής της X προκύπτει από την κοινού συνάρτηση κατανομής ως εξής:

$$F_X(a) = P\left\{X \leq a\right\} = P\left\{X \leq a, Y < \infty\right\} = F(a, \infty).$$

Ομοίως

$$F_Y(b) = P\left\{Y \leq b\right\} = P\left\{X < \infty, Y \leq b\right\} = F(\infty, b).$$

Στην περίπτωση που οι X και Y είναι διακριτές τυχαίες μεταβλητές, ορίζεται η *από κοινού συνάρτηση πιθανότητας* των X και Y ως:

$$p(x, y) = P\left\{X = x, Y = y\right\}.$$

Οι συναρτήσεις πιθανότητας των X και Y προκύπτουν από την $p(x, y)$ από τις σχέσεις:

$$p_X(x) = \sum_{y:p(x,y)>0} p(x, y)$$

και

$$p_Y(y) = \sum_{x:p(x,y)>0} p(x, y).$$

Θα λέμε ότι δύο τυχαίες μεταβλητές X και Y είναι *από κοινού συνεχείς* αν υπάρχει μια συνάρτηση $f(x, y)$, ορισμένη για κάθε πραγματικό αριθμό x και y, που έχει την παρακάτω ιδιότητα για όλα τα σύνολα A και B των πραγματικών αριθμών:

$$P\left\{X \in A, Y \in B\right\} = \int_B \int_A f(x, y)dxdy.$$

Η συνάρτηση $f(x,y)$ καλείται *από κοινού συνάρτηση πυκνότητας πιθανότητας* των X και Y. Η συνάρτηση πυκνότητας πιθανότητας της X προκύπτει εύκολα αν είναι γνωστή η $f(x,y)$ αφού:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy.$$

Ομοίως για την Y ισχύει:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

Ισχύουν επίσης τα εξής:

$$P\{X \in A\} = P\{X \in A, Y \in (-\infty,\infty)\} = \int_{-\infty}^{\infty}\int_A f(x,y)dxdy = \int_A f_X(x)dx$$

και

$$P\{Y \in B\} = P\{X \in (-\infty,\infty), Y \in B\} = \int_B\int_{-\infty}^{\infty} f(x,y)dxdy = \int_B f_Y(y)dy.$$

Χρησιμοποιώντας την πρόταση 2.4.1 για την περίπτωση δύο τυχαίων μεταβλητών αποδεικνύεται ότι:

$$E\left[g(X,Y)\right] = \sum_y\sum_x g(x,y)p(x,y)$$

αν οι τυχαίες μεταβλητές X, Y είναι διακριτές και

$$E\left[g(X,Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy$$

αν είναι συνεχείς. Για παράδειγμα, αν $g(X,Y) = X+Y$ και X, Y συνεχείς τότε:

$$E\left[X+Y\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy$$

$$= E\left[X\right] + E\left[Y\right].$$

Το πρώτο ολοκλήρωμα προκύπτει θέτοντας $g(x,y) = x$ (δηλαδή $g(X,Y) = X$) και το δεύτερο αν $g(x,y) = y$. Το παραπάνω αποτέλεσμα ισχύει και για την περίπτωση που έχουμε διακριτές τυχαίες μεταβλητές. Γενικότερα, αν $X_1, X_2, \ldots, X_n$ είναι n τυχαίες μεταβλητές, τότε για κάθε n σταθερές $a_1, a_2, \ldots, a_n$ ισχύει:

$$E\left[a_1X_1 + a_2X_2 + \ldots a_nX_n\right] = a_1E\left[X_1\right] + \ldots + a_nE\left[X_n\right].$$

2.5.2 Ανεξάρτητες Τυχαίες Μεταβλητές

Δύο τυχαίες μεταβλητές X και Y καλούνται *ανεξάρτητες* αν για όλα τα a, b ισχύει:

$$P\{X \leq a, Y \leq b\} = P\{X \leq a\} P\{Y \leq b\}.$$

Με άλλα λόγια οι X και Y είναι ανεξάρτητες αν για όλα τα a, b, τα γεγονότα $E_a = \{X \leq a\}$ και $F_b = \{Y \leq b\}$ είναι ανεξάρτητα. Επίσης, όταν οι μεταβλητές X, Y είναι ανεξάρτητες, η από κοινού συνάρτηση κατανομής τους είναι ίση με:

$$F(a, b) = F_X(a) F_Y(b)$$

για όλα τα a, b. Στην περίπτωση που οι X και Y είναι διακριτές τυχαίες μεταβλητές ισχύει:

$$p(x, y) = p_X(x) p_Y(y)$$

ενώ αν είναι από κοινού συνεχείς

$$f(x, y) = f_X(x) f_Y(y).$$

Πρόταση 2.5.1 *Αν X και Y είναι ανεξάρτητες τυχαίες μεταβλητές, τότε για οποιεσδήποτε συναρτήσεις h και g ισχύει:*

$$E[g(X)h(Y)] = E[g(X)] E[h(Y)]$$

2.5.3 Συνδιασπορά και Διασπορά Αθροίσματος Τυχαίων Μεταβλητών

Η *συνδιασπορά* δύο τυχαίων μεταβλητών X και Y συμβολίζεται με $Cov(X, Y)$ και είναι ίση με:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

$$= E[XY - YE[X] - XE[Y] + E[X] E[Y]]$$

$$= E[XY] - E[Y] E[X] - E[X] E[Y] + E[X] E[Y]$$

$$= E[XY] - E[X] E[Y].$$

Παρατηρήστε ότι αν οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες, τότε από την Πρόταση 2.5.1 προκύπτει ότι $Cov(X, Y) = 0$.

Η συνδιασπορά έχει τις ακόλουθες ιδιότητες:

1. $Cov(X, X) = Var(X)$
2. $Cov(X, Y) = Cov(Y, X)$
3. $Cov(cX, Y) = cCov(X, Y)$
4. $Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$

Η τέταρτη ιδιότητα γενικεύεται ως εξής:

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov\left(X_i, Y_j\right). \tag{2.4}$$

Η διασπορά του αθροίσματος πολλών τυχαίων μεταβλητών μπορεί να προκύψει χρησιμοποιώντας την εξίσωση (2.4):

$$\begin{aligned} Var(\sum_{i=1}^{n} X_i) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov\left(X_i, X_j\right) \\ &= \sum_{i=1}^{n} Cov\left(X_i, X_i\right) + \sum_{i=1}^{n}\sum_{j\neq i} Cov\left(X_i, X_j\right) \\ &= \sum_{i=1}^{n} Var\left(X_i\right) + 2\sum_{i=1}^{n}\sum_{j<i} Cov\left(X_i, X_j\right). \end{aligned}$$

Αν οι τυχαίες μεταβλητές X_i, $i = 1, \ldots, n$, είναι ανεξάρτητες τότε:

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var\left(X_i\right).$$

Παράδειγμα 2.5.1 *Υπολογίστε τη διασπορά μιας διωνυμικής τυχαίας μεταβλητής X με παραμέτρους n και p.*

Λύση. Επειδή η διωνυμική τυχαία μεταβλητή αντιπροσωπεύει τον αριθμό των επιτυχιών σε n ανεξάρτητες δοκιμές, όπου κάθε δοκιμή έχει πιθανότητα p να επιτύχει, μπορούμε να γράψουμε την τυχαία μεταβλητή X ως $X = X_1 + \ldots + X_n$ όπου οι X_i είναι ανεξάρτητες τυχαίες μεταβλητές Bernoulli τέτοιες ώστε

$$X_i = \begin{cases} 1 & \text{αν } i\text{-οστή δοκιμή ήταν επιτυχής} \\ 0 & \text{διαφορετικά} \end{cases}$$

Επομένως, $Var(X) = Var(X_1) + \ldots + Var(X_n)$. Αλλά $Var(X_i) = E\left[X_i^2\right] - (E\left[X_i\right])^2 = E\left[X_i\right] - (E\left[X_i\right])^2$ αφού $X_i^2 = X_i$. Άρα, $Var(X_i) = p - p^2$ και $Var(X) = np(1-p)$. ∎

2.6 Ροπογεννήτριες Συναρτήσεις

Η *ροπογεννήτρια συνάρτηση* $\phi(t)$ μιας τυχαίας μεταβλητής X ορίζεται για όλες τις τιμές t ως:

$$\phi(t) = E\left[e^{tX}\right] = \begin{cases} \sum_x e^{tx} p(x) & \text{αν η } X \text{ είναι διακριτή} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{αν η } X \text{ είναι συνεχής} \end{cases}$$

Η αναμενόμενη τιμή και η διασπορά της X μπορούν να προκύψουν μέσω της συνάρτησης $\phi(t)$. Για παράδειγμα,

$$\phi'(t) = \frac{d}{dt} E\left[e^{tX}\right] = E\left[\frac{d}{dt}(e^{tX})\right] = E\left[Xe^{tX}\right].$$

Επομένως, $\phi'(0) = E[X]$. Όμοια, $\phi''(t) = \frac{d}{dt}\phi'(t) = E\left[X^2 e^{tX}\right]$ και $\phi''(0) = E\left[X^2\right]$. Δηλαδή, $Var(X) = E\left[X^2\right] - (E[X])^2 = \phi''(0) - (\phi'(0))^2$. Γενικότερα, ισχύει ότι $\phi^n(0) = E[X^n]$, με $n \geq 1$.

Παράδειγμα 2.6.1 *Υπολογίστε τη ροπογεννήτρια συνάρτηση μιας τυχαίας μεταβλητής X που ακολουθεί κατανομή Poisson με παράμετρο λ. Στη συνέχεια βρείτε την αναμενόμενη τιμή και τη διασπορά της.*

Λύση. $\phi(t) = E\left[e^{tX}\right] = \sum_{n=0}^{\infty} \frac{e^{tn}e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda e^t)^n}{n!} = e^{-\lambda}e^{\lambda e^t}$.
Παραγωγίζοντας προκύπτει ότι $\phi'(t) = \lambda e^t exp\left\{\lambda(e^t-1)\right\}$ και
$\phi''(t) = (\lambda e^t)^2 exp\left\{\lambda(e^t-1)\right\} + \lambda e^t exp\left\{\lambda(e^t-1)\right\}$. Επομένως, $E[X] = \phi'(0) = \lambda$, $E\left[X^2\right] = \phi''(0) = \lambda^2 + \lambda$ και $Var(X) = E\left[X^2\right] - (E[X])^2 = \lambda$. ■

Μια σημαντική ιδιότητα των ροπογεννητριών συναρτήσεων είναι ότι η ροπογεννήτρια συνάρτηση του αθροίσματος πολλών ανεξάρτητων τυχαίων μεταβλητών είναι ίση με το γινόμενο των ροπογεννητριών συναρτήσεων των επιμέρους τυχαίων μεταβλητών. Αν για παράδειγμα X και Y είναι ανεξάρτητες τυχαίες μεταβλητές και οι αντίστοιχες ροπογεννήτριες συναρτήσεις τους είναι οι $\phi_X(t)$ και $\phi_Y(t)$, τότε η ροπογεννήτρια συνάρτηση του αθροίσματος $X+Y$ είναι ίση με:

$$\phi_{X+Y}(t) = E\left[e^{t(X+Y)}\right] = E\left[e^{tX}\right]E\left[e^{tY}\right] = \phi_X(t)\phi_Y(t).$$

Επίσης, η ροπογεννήτρια συνάρτηση μιας τυχαίας μεταβλητής X καθορίζει μοναδικά και την κατανομή της. Με άλλα λόγια, υπάρχει μία 1-1 αντιστοιχία ανάμεσα στη ροπογεννήτρια συνάρτηση και τη συνάρτηση κατανομής μιας τυχαίας μεταβλητής.

Παράδειγμα 2.6.2 *Υπολογίστε την κατανομή του αθροίσματος $X+Y$ όταν οι X και Y είναι ανεξάρτητες τυχαίες μεταβλητές Poisson με παραμέτρους λ_1 και λ_2 αντίστοιχα.*

Λύση. Από το προηγούμενο παράδειγμα έχουμε ότι $\phi_X(t) = e^{\lambda_1(e^t-1)}$ και $\phi_Y(t) = e^{\lambda_2(e^t-1)}$. Άρα $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)}$. Δηλαδή το άθροισμα $X+Y$ ακολουθεί επίσης κατανομή Poisson με παράμετρο $\lambda_1+\lambda_2$. ■

2.7 Οριακά Θεωρήματα

Στην ενότητα αυτή παρουσιάζονται ορισμένες ανισότητες και θεωρήματα πολύ σημαντικά στη θεωρία πιθανοτήτων. Αρχικά παρουσιάζεται η ανισότητα Markov:

Πρόταση 2.7.1 *Αν X είναι μία τυχαία μεταβλητή που παίρνει μόνο θετικές τιμές, τότε για κάθε $a > 0$ ισχύει ότι*

$$P\{X \geq a\} \leq \frac{E[X]}{a}.$$

Απόδειξη. Παρακάτω δίνεται η απόδειξη για την περίπτωση που η τυχαία μεταβλητή X είναι συνεχής με συνάρτηση πυκνότητας πιθανότητας $f()$. Παρόμοια είναι η απόδειξη και στην περίπτωση που η X είναι διακριτή τυχαία μεταβλητή. Ισχύει ότι:

$$E[X] = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \geq \int_a^\infty xf(x)dx$$
$$\geq a\int_a^\infty f(x)dx = aP\{X \geq a\}.$$

∎

Πρόταση 2.7.2 *(Ανισότητα Chebyshev): Αν X είναι μια τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2, τότε για κάθε $k > 0$ ισχύει*

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}.$$

Απόδειξη. Επειδή η $(X - \mu)^2$ είναι μια τυχαία μεταβλητή που δεν παίρνει αρνητικές τιμές, μπορούμε να εφαρμόσουμε την ανισότητα Markov με $a = k^2$. Δηλαδή:

$$P\left\{(X - \mu)^2 \geq k^2\right\} \leq \frac{E\left[(X - \mu)^2\right]}{k^2}.$$

Αλλά $(X - \mu)^2 \geq k^2$ αν και μόνο αν $|X - \mu| \geq k$. Επομένως η παραπάνω ανισότητα είναι ισοδύναμη με:

$$P\{|X - \mu| \geq k\} \leq \frac{E\left[(X - \mu)^2\right]}{k^2} = \frac{\sigma^2}{k^2}.$$

∎

Οι ανισότητες Markov και Chebyshev είναι πολύ σημαντικές αφού μας επιτρέπουν να βρίσκουμε όρια στις πιθανότητες όταν δεν γνωρίζουμε την κατανομή, αλλά μόνο τη μέση τιμή και τη διασπορά.

Παράδειγμα 2.7.1 *Ο αριθμός των μηχανών που παράγονται σε ένα εργοστάσιο κάθε βδομάδα αποτελεί μια τυχαία μεταβλητή με μέση τιμή 500.*

1. *Τί μπορούμε να πούμε για την πιθανότητα η παραγωγή αυτής της βδομάδας να είναι τουλάχιστον 1000;*

2. *Αν η διασπορά είναι ίση με 100, τότε τί μπορούμε να πούμε για την πιθανότητα η παραγωγή αυτής της βδομάδας να είναι μεταξύ 400 και 600;*

Λύση. Έστω X ο αριθμός των μηχανών που παράγονται σε μια βδομάδα.

1. Σύμφωνα με την ανισότητα Markov έχουμε ότι $P\{X \geq 1000\} \leq \frac{E[X]}{1000} = \frac{1}{2}$.

2. Από την ανισότητα Chebyshev προκύπτει ότι $P\{|X-500| \geq 100\} \leq \frac{\sigma^2}{100^2} = \frac{1}{100}$. Άρα $P\{|X-500| < 100\} \geq 1 - \frac{1}{100} = \frac{99}{100}$.

■

Το επόμενο Θεώρημα, γνωστό ως Νόμος των Μεγάλων Αριθμών, είναι πιθανώς το πιο γνωστό αποτέλεσμα στη θεωρία πιθανοτήτων. Ο νόμος των μεγάλων αριθμών δηλώνει ότι ο μέσος όρος μιας σειράς ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την ίδια κατανομή συγκλίνει στη μέση τιμή αυτής της κατανομής.

Θεώρημα 2.7.1 *(Νόμος των Μεγάλων Αριθμών) Έστω $X_1, X_2, \ldots$ μια σειρά από ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κοινή κατανομή με $E[X_i] = \mu$. Τότε, με πιθανότητα 1 ισχύει:*

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \to \mu$$

καθώς $n \to \infty$.

Ένα ακόμα πολύ σημαντικό Θεώρημα της θεωρίας πιθανοτήτων είναι το Κεντρικό Οριακό Θεώρημα. Το θεώρημα αυτό παρέχει μια απλή μέθοδο για τον υπολογισμό προσεγγίσεων πιθανότητας για αθροίσματα ανεξάρτητων τυχαίων μεταβλητών.

Θεώρημα 2.7.2 *(Κεντρικό Οριακό Θεώρημα) Έστω $X_1, X_2, \ldots$ μια σειρά από ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κοινή κατανομή με αναμενόμενη τιμή μ και διασπορά σ^2. Τότε η κατανομή της τυχαίας μεταβλητής*

$$\frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}$$

τείνει στην τυπική (μοναδιαία) κανονική κατανομή καθώς $n \to \infty$. Δηλαδή,

$$P\left\{\frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx.$$

Κεφάλαιο 3

Δεσμευμένες Πιθανότητες

3.1 Εισαγωγή

Η μελέτη των δεσμευμένων πιθανοτήτων είναι πολύ χρήσιμη στη θεωρία πιθανοτήτων γενικότερα. Υπάρχουν δύο κύριοι λόγοι για αυτό. Πρώτον, πολύ συχνά θέλουμε να υπολογίσουμε πιθανότητες όταν μας δίνεται από πριν κάποια επιπλέον πληροφορία. Κατά δεύτερον, για τον υπολογισμό κάποιας πιθανότητας αποδεικνύεται ότι πολλές φορές είναι χρήσιμο να ορίσουμε μια νέα τυχαία μεταβλητή και μέσω των δεσμευμένων πιθανοτήτων να απαντήσουμε στο αρχικό μας ερώτημα. Στο κεφάλαιο αυτό θα ορίσουμε βασικές έννοιες των δεσμευμένων πιθανοτήτων και θα δούμε την χρησιμότητά τους στην επίλυση διαφόρων προβλημάτων.

3.2 Διακριτές Τυχαίες Μεταβλητές

Γνωρίζουμε ότι για οποιαδήποτε γεγονότα E και F, η δεσμευμένη πιθανότητα να συμβεί το E δοθέντος του F είναι ίση με $P(E|F) = \frac{P(EF)}{P(F)}$. Επομένως, αν X και Y είναι δύο διακριτές τυχαίες μεταβλητές, μπορούμε να ορίσουμε την *δεσμευμένη συνάρτηση πιθανότητας* της X δοθέντος ότι $Y = y$ ως εξής:

$$p_{X|Y}(x|y) = P\{X = x|Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}} = \frac{p(x,y)}{p_Y(y)}.$$

Η παραπάνω εξίσωση ισχύει για όλες τις τιμές του y για τις οποίες $P\{Y = y\} > 0$. Ομοίως ορίζεται και η *δεσμευμένη συνάρτηση κατανομής* της X δοθέντος ότι $Y = y$:

$$F_{X|Y}(x|y) = P\{X \leq x|Y = y\} = \sum_{a \leq x} p_{X|Y}(a|y).$$

Τέλος, η *δεσμευμένη αναμενόμενη τιμή* της X δοθέντος ότι $Y = y$ ορίζεται ως:

$$E[X|Y = y] = \sum_x xP\{X = x|Y = y\} = \sum_x xp_{X|Y}(x|y).$$

Προφανώς, αν η τυχαία μεταβλητή X είναι ανεξάρτητη της Y, τότε η δεσμευμένη συνάρτηση πιθανότητας, κατανομής και η αναμενόμενη τιμή είναι ίδιες με τις αντίστοιχες που θα είχαμε αν δεν υπήρχε καμία συνθήκη. Δηλαδή για παράδειγμα, $p_{X|Y}(x|y) = P\{X = x|Y = y\} = P\{X = x\} = p_X(x)$.

Παράδειγμα 3.2.1 *Αν X και Y είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομή Poisson με αναμενόμενες τιμές λ_1 και λ_2 αντίστοιχα, υπολογίστε την δεσμευμένη συνάρτηση πιθανότητας της X δοθέντος ότι $X + Y = n$.*

Λύση. Αν $Z = X + Y$ έχουμε ότι

$$p_{X|Z}(x|n) = P\{X = x|X + Y = n\} = \frac{P\{X = x, X + Y = n\}}{P\{X + Y = n\}}$$

$$= \frac{P\{X = x, Y = n - x\}}{P\{X + Y = n\}} = \frac{P\{X = x\}P\{Y = n - x\}}{P\{X + Y = n\}}$$

$$= \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2}\lambda_2^{n-x}}{(n-x)!}\left[\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}\right]^{-1} = \frac{n!}{(n-x)!x!}\frac{\lambda_1^x\lambda_2^{n-x}}{(\lambda_1+\lambda_2)^n}$$

$$= \begin{pmatrix} n \\ x \end{pmatrix}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-x}.$$

Με άλλα λόγια, η δεσμευμένη κατανομή της X δοθέντος ότι $X + Y = n$ είναι η διωνυμική κατανομή με παραμέτρους n και $\frac{\lambda_1}{\lambda_1+\lambda_2}$.

∎

3.3 Συνεχείς Τυχαίες Μεταβλητές

Αν X και Y είναι συνεχείς τυχαίες μεταβλητές που έχουν από κοινού συνάρτηση πυκνότητας πιθανότητας $f(x, y)$, τότε η *δεσμευμένη συνάρτηση πυκνότητας πιθανότητας* της X δοθέντος ότι $Y = y$ ορίζεται για όλα τα y για τα οποία $f_Y(y) > 0$ και είναι ίση με:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

Η *δεσμευμένη συνάρτηση κατανομής* και η *δεσμευμένη αναμενόμενη τιμή* της X δοθέντος ότι $Y = y$ είναι ίσες αντίστοιχα με:

$$F_{X|Y}(x|y) = P\{X \leq x|Y = y\} = \int_{-\infty}^{x} f_{X|Y}(x|y)dx$$

και

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx.$$

Παράδειγμα 3.3.1 *Υποθέστε ότι η από κοινού συνάρτηση πυκνότητας πιθανότητας των τυχαίων μεταβλητών X και Y δίνεται από την σχέση:*

$$f(x,y) = \begin{cases} 6xy(2-x-y) & \text{αν } 0<x<1,\ 0<y<1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Υπολογίστε την δεσμευμένη αναμενόμενη τιμή της X δοθέντος ότι $Y=y$.

Λύση. Αρχικά πρέπει να υπολογίσουμε την δεσμευμένη συνάρτηση πυκνότητας πιθανότητας. Ισχύει ότι:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{6xy(2-x-y)}{\int_0^1 6xy(2-x-y)dx} = \frac{6xy(2-x-y)}{y(4-3y)} = \frac{6x(2-x-y)}{4-3y}.$$

Επομένως,

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx = \int_0^1 \frac{6x^2(2-x-y)}{4-3y}dx = \frac{2(2-y)-6/4}{4-3y} = \frac{5-4y}{8-6y}.$$

∎

Παράδειγμα 3.3.2 *Έστω ότι η από κοινού συνάρτηση πυκνότητας πιθανότητας δύο τυχαίων μεταβλητών X και Y δίνεται από την σχέση:*

$$f(x,y) = \begin{cases} \frac{1}{2}ye^{-xy} & \text{αν } 0<x<\infty,\ 0<y<2 \\ 0 & \text{διαφορετικά} \end{cases}$$

Ποια είναι η αναμενόμενη τιμή $E\left[e^{X/2}|Y=1\right]$;

Λύση. Ισχύει ότι:

$$f_{X|Y}(x|1) = \frac{f(x,1)}{f_Y(1)} = \frac{\frac{1}{2}e^{-x}}{\int_0^\infty \frac{1}{2}e^{-x}dx} = e^{-x}.$$

Επομένως,

$$E\left[e^{X/2}|Y=1\right] = \int_0^\infty e^{x/2} f_{X|Y}(x|1)dx = \int_0^\infty e^{x/2}e^{-x}dx = 2.$$

∎

3.4 Υπολογισμός Αναμενόμενων Τιμών

Ορίζουμε ως $E[X|Y]$ τη συνάρτηση της τυχαίας μεταβλητής Y η οποία για $Y=y$ παίρνει την τιμή $E[X|Y=y]$. Παρατηρήστε ότι και η $E[X|Y]$ είναι μια τυχαία μεταβλητή. Μια πολύ σημαντική ιδιότητα των δεσμευμένων αναμενόμενων τιμών είναι ότι για όλες τις τυχαίες μεταβλητές X και Y ισχύει:

$$E[X] = E[E[X|Y]]. \tag{3.1}$$

Αν Y είναι μια διακριτή τυχαία μεταβλητή, τότε η εξίσωση (3.1) δηλώνει ότι:

$$E[X] = \sum_y E[X|Y=y] P\{Y=y\} \tag{3.2}$$

ενώ αν Y είναι μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$

$$E[X] = \int_{-\infty}^{\infty} E[X|Y=y] f_Y(y) dy. \tag{3.3}$$

Παρακάτω δίνεται η απόδειξη για την εξίσωση (3.2). Παρόμοια είναι και η απόδειξη της εξίσωσης (3.3).
Απόδειξη.

$$\sum_y E[X|Y=y] P\{Y=y\} = \sum_y \sum_x xP\{X=x|Y=y\} P\{Y=y\}$$

$$= \sum_y \sum_x x \frac{P\{X=x, Y=y\}}{P\{Y=y\}} P\{Y=y\} = \sum_x x \sum_y P\{X=x, Y=y\}$$

$$= \sum_x xP\{X=x\} = E[X].$$

∎

Παράδειγμα 3.4.1 *Ο Παύλος πρόκειται να διαβάσει ένα κεφάλαιο είτε από το βιβλίο των πιθανοτήτων ή από το βιβλίο της ιστορίας. Αν ο αριθμός των τυπογραφικών λαθών σε ένα κεφάλαιο του βιβλίου των πιθανοτήτων ακολουθεί κατανομή Poisson με μέση τιμή 2 και αν ο αντίστοιχος αριθμός λαθών σε ένα κεφάλαιο του βιβλίου της ιστορίας ακολουθεί την ίδια κατανομή με μέση τιμή 5, ποιος είναι ο αναμενόμενος αριθμός λαθών που θα συναντήσει ο Παύλος αν υποθέσουμε ότι είναι εξίσου πιθανό να διαλέξει ένα από τα δύο βιβλία;*

Λύση. Ορίζουμε δύο τυχαίες μεταβλητές X και Y, όπου X είναι ο αριθμός των λαθών και $Y=1$ αν ο Παύλος επιλέξει το βιβλίο της ιστορίας και $Y=2$ αν επιλέξει το βιβλίο των πιθανοτήτων. Τότε $E[X] = E[X|Y=1] P\{Y=1\} + E[X|Y=2] P\{Y=2\} = 5\frac{1}{2} + 2\frac{1}{2} = \frac{7}{2}$. ∎

Παράδειγμα 3.4.2 *Ένας ανθρακωρύχος έχει παγιδευτεί σε ένα ορυχείο που έχει τρεις πόρτες. Η πρώτη πόρτα οδηγεί σε ένα τούνελ που θα τον οδηγήσει στην έξοδο μετά από δύο ώρες. Η δεύτερη πόρτα οδηγεί σε ένα τούνελ που θα τον επιστρέψει στο ίδιο σημείο μετά από τρεις ώρες και η τρίτη θα τον οδηγήσει πίσω στο σημείο που ήταν μετά από πέντε ώρες. Αν υποθέσουμε ότι κάθε φορά ο ανθρακωρύχος είναι το ίδιο πιθανό να επιλέξει οποιαδήποτε από τις τρεις πόρτες, ποιος είναι ο αναμενόμενος χρόνος για να βγει από το ορυχείο;*

Λύση. Έστω X ο χρόνος που χρειάζεται ο ανθρακωρύχος για να βγει από το ορυχείο και έστω Y μια τυχαία μεταβλητή που προσδιορίζει την πόρτα που διαλέγει. Τότε $E[X] = E[X|Y=1]P\{Y=1\} + E[X|Y=2]P\{Y=2\} + E[X|Y=3]P\{Y=3\} = \frac{1}{3}(E[X|Y=1] + E[X|Y=2] + E[X|Y=3])$. Αλλά $E[X|Y=1] = 2$, $E[X|Y=2] = 3 + E[X]$ και $E[X|Y=3] = 5 + E[X]$. Επομένως, $E[X] = \frac{1}{3}(2 + 3 + E[X] + 5 + E[X]) \Rightarrow E[X] = 10$. ∎

Δεσμευμένες αναμενόμενες τιμές μπορούν να οριστούν και για δύο τυχαίες μεταβλητές W, Y. Δηλαδή, το ανάλογο των εξισώσεων (3.2) και (3.3), υπό την συνθήκη ότι $Y = y$, είναι:

$$E[X|Y=y] = \sum_{w} E[X|W=w, Y=y]\, P\{W=w|Y=y\}$$

αν W είναι μια διακριτή τυχαία μεταβλητή και

$$E[X|Y=y] = \int_{-\infty}^{\infty} E[X|W=w, Y=y]\, f_{W|Y}(w|y)dw$$

αν W είναι μια συνεχής τυχαία μεταβλητή. Ισχύει επίσης ότι $E[X|Y] = E[E[X|Y,W]|Y]$.

3.5 Υπολογισμός Διασποράς

Η δεσμευμένη διασπορά της τυχαίας μεταβλητής X δοθέντος ότι $Y = y$ ορίζεται ως εξής:

$$Var(X|Y=y) = E\left[(X - E[X|Y=y])^2 |Y=y\right].$$

Αναλύοντας το δεξιό μέρος της εξίσωσης προκύπτει ότι:

$$Var(X|Y=y) = E\left[X^2|Y=y\right] - (E[X|Y=y])^2.$$

Ισχύει δηλαδή κάτι ανάλογο της γνωστής εξίσωσης $Var(X) = E\left[X^2\right] - (E[X])^2$.

Ορίζοντας τώρα ως $Var(X|Y)$ τη συνάρτηση της τυχαίας μεταβλητής Y η οποία για $Y = y$ παίρνει την τιμή $Var(X|Y=y)$, προκύπτει το παρακάτω αποτέλεσμα.

Πρόταση 3.5.1 *Ισχύει ότι:*

$$Var(X) = E[Var(X|Y)] + Var(E[X|Y])$$

Απόδειξη. Έχουμε ότι

$$E[Var(X|Y)] = E\left[E\left[X^2|Y\right] - (E[X|Y])^2\right]$$

$$= E\left[E\left[X^2|Y\right]\right] - E\left[\left(E\left[X|Y\right]\right)^2\right]$$

$$= E\left[X^2\right] - E\left[\left(E\left[X|Y\right]\right)^2\right]$$

και

$$Var\left(E\left[X|Y\right]\right) = E\left[\left(E\left[X|Y\right]\right)^2\right] - \left(E\left[E\left[X|Y\right]\right]\right)^2$$

$$= E\left[\left(E\left[X|Y\right]\right)^2\right] - \left(E\left[X\right]\right)^2.$$

Επομένως,

$$E\left[Var(X|Y)\right] + Var\left(E\left[X|Y\right]\right) = E\left[X^2\right] - \left(E\left[X\right]\right)^2 = Var(X).$$

∎

3.6 Υπολογισμός Πιθανοτήτων

Έστω E ένα τυχαίο γεγονός και X μια τυχαία μεταβλητή τέτοια ώστε $X = 1$ αν το γεγονός E συμβεί και $X = 0$ αν το E δεν συμβεί. Προκύπτει από τον ορισμό της X ότι $E\left[X\right] = P\left\{E\right\}$ και $E\left[X|Y = y\right] = P\left\{E|Y = y\right\}$ για οποιαδήποτε τυχαία μεταβλητή Y. Επομένως από τις εξισώσεις (3.2) και (3.3) έχουμε ότι:

$$P\left\{E\right\} = \sum_y P\left\{E|Y = y\right\} p(y)$$

αν η τυχαία μεταβλητή Y είναι διακριτή και

$$P\left\{E\right\} = \int_{-\infty}^{\infty} P\left\{E|Y = y\right\} f_Y(y)dy$$

αν η τυχαία μεταβλητή Y είναι συνεχής.

Παράδειγμα 3.6.1 *Έστω X και Y δύο ανεξάρτητες συνεχείς τυχαίες μεταβλητές με συναρτήσεις πυκνότητας πιθανότητας f_X και f_Y αντίστοιχα. Υπολογίστε την πιθανότητα $P\left\{X < Y\right\}$.*

Λύση.

$$P\left\{X < Y\right\} = \int_{-\infty}^{\infty} P\left\{X < Y|Y = y\right\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} P\left\{X < y|Y = y\right\} f_Y(y)dy = \int_{-\infty}^{\infty} P\left\{X < y\right\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} F_X(y) f_Y(y)dy$$

όπου $F_X(y) = \int_{-\infty}^{y} f_X(x)dx$.

∎

Παράδειγμα 3.6.2 *Υποθέστε ότι ο αριθμός των ανθρώπων που επισκέπτονται ένα μουσείο κάθε μέρα αποτελεί μια τυχαία μεταβλητή Poisson με μέση τιμή λ. Υποθέστε επίσης ότι κάθε επισκέπτης είναι με πιθανότητα p γυναίκα και με πιθανότητα $1-p$ άνδρας. Υπολογίστε την από κοινού πιθανότητα ακριβώς n γυναίκες και m άνδρες να επισκέφτηκαν σήμερα το μουσείο.*

Λύση. Έστω N_1 ο αριθμός των γυναικών, N_2 ο αριθμός των ανδρών και $N = N_1 + N_2$ ο συνολικός αριθμός των ανθρώπων που επισκέφτηκαν σήμερα το μουσείο. Τότε:

$$P\{N_1 = n, N_2 = m\} = \sum_{i=0}^{\infty} P\{N_1 = n, N_2 = m | N = i\} P\{N = i\}.$$

Αλλά $P\{N_1 = n, N_2 = m | N = i\} = 0$ όταν $i \neq n + m$. Επομένως,

$$\begin{aligned} P\{N_1 = n, N_2 = m\} &= P\{N_1 = n, N_2 = m | N = n + m\} P\{N = n + m\} \\ &= P\{N_1 = n, N_2 = m | N = n + m\} e^{-\lambda} \frac{\lambda^{n+m}}{(n+m)!}. \end{aligned}$$

Καθένας από τους $n+m$ επισκέπτες του μουσείου είναι γυναίκα με πιθανότητα p και άνδρας με πιθανότητα $1-p$. Αν θεωρήσουμε ως επιτυχία το γεγονός ο επισκέπτης να είναι γυναίκα και αποτυχία το γεγονός να είναι άνδρας, τότε η πιθανότητα $P\{N_1 = n, N_2 = m | N = n + m\}$ είναι η πιθανότητα να έχουμε n επιτυχίες σε $n + m$ δοκιμές (διωνυμική κατανομή). Άρα

$$P\{N_1 = n, N_2 = m\} = \binom{n+m}{n} p^n (1-p)^m e^{-\lambda} \frac{\lambda^{n+m}}{(n+m)!}.$$

∎

Κεφάλαιο 4

Εκθετική Κατανομή και Διαδικασία Poisson

4.1 Εισαγωγή

Συχνά για την περιγραφή ενός πραγματικού φαινομένου με ένα μαθηματικό μοντέλο είναι απαραίτητο να κάνουμε διάφορες απλοποιήσεις. Μια τέτοια απλοποίηση είναι να θεωρήσουμε ότι ορισμένες τυχαίες μεταβλητές είναι εκθετικά κατανεμημένες. Ο κύριος λόγος είναι ότι η εκθετική κατανομή αποτελεί σε πολλές περιπτώσεις μια καλή προσέγγιση της πραγματικής κατανομής.

Στο κεφάλαιο αυτό θα μελετήσουμε ιδιότητες της εκθετικής κατανομής και θα δούμε πώς συνδέεται με μια κατηγορία διαδικασιών καταμέτρησης, γνωστές ως διαδικασίες Poisson. Το δεύτερο μέρος του κεφαλαίου ασχολείται αποκλειστικά με αυτές τις διαδικασίες, τονίζοντας μεταξύ άλλων την χρησιμότητά τους στην περιγραφή πραγματικών φαινομένων.

4.2 Η Εκθετική Κατανομή

4.2.1 Ορισμοί

Μια συνεχής τυχαία μεταβλητή ακολουθεί την εκθετική κατανομή με παράμετρο $\lambda > 0$, αν η συνάρτηση πυκνότητας πιθανότητάς της είναι ίση με:

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

ή ισοδύναμα αν η συνάρτηση κατανομής της είναι η

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 1 - e^{-\lambda x} & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases}$$

Η μέση τιμή της εκθετικής κατανομής είναι:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{\infty} \lambda x e^{-\lambda x} dx$$

$$= -xe^{-\lambda x}|_0^\infty + \int_0^\infty e^{-\lambda x}dx = \frac{1}{\lambda}\int_0^\infty \lambda e^{-\lambda x}dx = \frac{1}{\lambda}.$$

Η ροπογεννήτρια συνάρτηση $\phi(t)$ της εκθετικής κατανομής δίνεται από τη σχέση:

$$\phi(t) = E\left[e^{tX}\right] = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \frac{\lambda}{\lambda - t}$$

για $t < \lambda$. Επομένως,

$$E\left[X^2\right] = \frac{d^2}{dt^2}\phi(t)|_{t=0} = \frac{2\lambda}{(\lambda-t)^3}|_{t=0} = \frac{2}{\lambda^2}$$

και

$$Var(X) = E\left[X^2\right] - (E\left[X\right])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

4.2.2 Ιδιότητες της Εκθετικής Κατανομής

Μια τυχαία μεταβλητή λέμε ότι είναι *χωρίς μνήμη* (memoryless) αν

$$P\left\{X > s+t | X > t\right\} = P\left\{X > s\right\} \tag{4.1}$$

για όλα τα $s, t \geq 0$. Αν για παράδειγμα υποθέσουμε ότι η X περιγράφει το χρόνο λειτουργίας μιας μηχανής, τότε η εξίσωση (4.1) δηλώνει ότι η πιθανότητα η μηχανή να λειτουργήσει για $s+t$ ώρες δεδομένου ότι έχει ήδη λειτουργήσει για t ώρες είναι η ίδια με την πιθανότητα να λειτουργήσει τις πρώτες s ώρες. Δηλαδή, η μηχανή δεν ''θυμάται'' ότι είναι ήδη σε λειτουργία t ώρες.

Η εξίσωση (4.1) είναι ισοδύναμη με:

$$\frac{P\left\{X > s+t, X > t\right\}}{P\left\{X > t\right\}} = P\left\{X > s\right\}$$

ή

$$P\left\{X > s+t\right\} = P\left\{X > s\right\} P\left\{X > t\right\}. \tag{4.2}$$

Η εξίσωση (4.2) ικανοποιείται στην περίπτωση που η X ακολουθεί εκθετική κατανομή (αφού $e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t}$), επομένως οι εκθετικές τυχαίες μεταβλητές έχουν την ιδιότητα της έλλειψης μνήμης. Αποδεικνύεται μάλιστα ότι η εκθετική κατανομή είναι η μοναδική κατανομή που έχει αυτή την ιδιότητα.

Παράδειγμα 4.2.1 *Υποθέστε ότι ο χρόνος που κάποιος ξοδεύει σε μια τράπεζα ακολουθεί εκθετική κατανομή με μέση τιμή 10 λεπτά. Ποια η πιθανότητα ένας πελάτης να μείνει πάνω από 15 λεπτά στην τράπεζα και ποια η ίδια πιθανότητα αν μας πουν ότι ήδη περίμενε 10 λεπτά;*

Λύση. Η μέση τιμή είναι 10 λεπτά, άρα $\lambda = \frac{1}{10}$. Η πρώτη πιθανότητα είναι ίση με

$$P\left\{X > 15\right\} = 1 - P\left\{X \leq 15\right\} = 1 - (1 - e^{-15\lambda}) = e^{-\frac{3}{2}} \approx 0.22.$$

Η δεύτερη πιθανότητα είναι η

$$P\{X > 15|X > 10\} = P\{X > 5\} = e^{-5\lambda} \approx 0.604.$$

∎

Παράδειγμα 4.2.2 *Υποθέστε ότι X_1 και X_2 είναι δύο ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την εκθετική κατανομή με παραμέτρους λ_1 και λ_2 αντίστοιχα. Ποια είναι η πιθανότητα $P\{X_1 < X_2\}$;*

Λύση.

$$P\{X_1 < X_2\} = \int_0^\infty P\{X_1 < X_2 | X_1 = x\} f_{X_1}(x) dx$$

$$= \int_0^\infty P\{x < X_2\} \lambda_1 e^{-\lambda_1 x} dx = \int_0^\infty e^{-\lambda_2 x} \lambda_1 e^{-\lambda_1 x} dx$$

$$= \int_0^\infty \lambda_1 e^{-(\lambda_1+\lambda_2)x} dx = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Προφανώς αν χρησιμοποιήσουμε συνθήκη ως προς την τυχαία μεταβλητή X_2 θα πάρουμε το ίδιο αποτέλεσμα. ∎

Το παραπάνω παράδειγμα μπορεί να φανεί πολύ χρήσιμο σε πρακτικά προβλήματα. Αν θεωρήσουμε ότι οι τυχαίες μεταβλητές X_1, X_2 αποτελούν τις διάρκειες δύο διεργασιών που εκτελούνται ταυτόχρονα, τότε μας ενδιαφέρει η πιθανότητα η πρώτη ή η δεύτερη διεργασία να τελειώσει πρώτη. Είδαμε ότι $P\{X_1 < X_2\} = \frac{\lambda_1}{\lambda_1+\lambda_2}$. Όμοια ισχύει ότι $P\{X_2 < X_1\} = \frac{\lambda_2}{\lambda_1+\lambda_2}$. Πολλές φορές επίσης μας ενδιαφέρει η κατανομή του διαστήματος X μέχρι να τελειώσει κάποια από τις δύο διεργασίες ή ισοδύναμα η κατανομή του $X = min(X_1, X_2)$. Ισχύει ότι:

$$F_X(x) = P\{X \le x\} = 1 - P\{X > x\} = 1 - P\{X_1 > x, X_2 > x\} =$$

$$1 - P\{X_1 > x\} P\{X_2 > x\} = 1 - e^{-\lambda_1 x} e^{-\lambda_2 x} = 1 - e^{-(\lambda_1+\lambda_2)x}$$

Άρα το διάστημα X είναι εκθετικά κατανεμημένο με παράμετρο $\lambda_1 + \lambda_2$ και δεν εξαρτάται από το ποια διεργασία τελειώνει πρώτη. Το αποτέλεσμα αυτό μπορεί να γενικευτεί για περισσότερες από δύο τυχαίες μεταβλητές.

4.2.3 Η Συνάρτηση Ρυθμού Βλάβης

Έστω X μια συνεχής, θετική τυχαία μεταβλητή με συνάρτηση κατανομής F και συνάρτηση πυκνότητας πιθανότητας f. Τότε η *συνάρτηση ρυθμού βλάβης* (failure or hazard rate function) $r(t)$ ορίζεται από τη σχέση:

$$r(t) = \frac{f(t)}{1 - F(t)}. \tag{4.3}$$

Για να ερμηνεύσουμε την ακριβή σημασία της συνάρτησης $r(t)$, θεωρείστε το εξής παράδειγμα: έστω μια μηχανή με διάρκεια λειτουργίας X, για την οποία γνωρίζουμε ότι λειτουργεί ήδη t ώρες. Θέλουμε να βρούμε την πιθανότητα να πάψει να λειτουργεί μετά από χρονικό διάστημα dt. Επομένως, ζητάμε την πιθανότητα $P\{X \in (t, t+dt)|X > t\}$. Ισχύει ότι:

$$P\{X \in (t, t+dt)|X > t\} = \frac{P\{X \in (t, t+dt), X > t\}}{P\{X > t\}}$$

$$= \frac{P\{X \in (t, t+dt)\}}{P\{X > t\}} \approx \frac{f(t)dt}{1-F(t)} = r(t)dt.$$

Δηλαδή η συνάρτηση $r(t)$ αναπαριστά την δεσμευμένη πιθανότητα μια μηχανή που ήδη λειτουργεί για χρόνο t, να χαλάσει.

Ας υποθέσουμε τώρα ότι η κατανομή της διάρκειας λειτουργίας της μηχανής είναι εκθετική. Από την ιδιότητα έλλειψης μνήμης της εκθετικής κατανομής, προκύπτει ότι η κατανομή του υπόλοιπου λειτουργίας μιας μηχανής που δουλεύει ήδη t ώρες είναι ίδια με την αντίστοιχη κατανομή μιας μηχανής που δεν έχει λειτουργήσει ακόμα. Περιμένουμε λοιπόν η συνάρτηση ρυθμού βλάβης να είναι σταθερή για την εκθετική κατανομή, αφού δεν εξαρτάται από το t. Πράγματι, $r(t) = \frac{f(t)}{1-F(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$.

Η συνάρτηση $r(t)$ ορίζει επίσης με μονοσήμαντο τρόπο τη συνάρτηση κατανομής F. Από την εξίσωση (4.3) έχουμε ότι:

$$r(t) = \frac{d/dtF(t)}{1-F(t)} \Rightarrow ln\left(1-F(t)\right) = -\int_0^t r(t)dt + k$$

$$\Rightarrow 1-F(t) = e^k exp\left\{-\int_0^t r(t)dt\right\}.$$

Θέτοντας $t = 0$ προκύπτει ότι $k = 0$, επομένως

$$F(t) = 1 - exp\left\{-\int_0^t r(t)dt\right\}.$$

Η παραπάνω σχέση μας αποδεικνύει επίσης ότι οι εκθετικές τυχαίες μεταβλητές είναι οι μόνες χωρίς μνήμη. Αν υποθέσουμε ότι X είναι μια τυχαία μεταβλητή χωρίς μνήμη, τότε η συνάρτηση ρυθμού βλάβης της είναι σταθερή, δηλαδή $r(t) = c$. Άρα, $F(t) = 1 - e^{-ct}$. Επομένως η κατανομή της X είναι εκθετική.

4.3 Η Διαδικασία Poisson

4.3.1 Εισαγωγή στις Στοχαστικές Διαδικασίες

Μια *στοχαστική διαδικασία* $\{X(t), t \in T\}$ είναι μια συλλογή από τυχαίες μεταβλητές. Δηλαδή, για κάθε $t \in T$, $X(t)$ είναι μια τυχαία μεταβλητή. Το t συχνά καλείται *χρόνος* και σαν αποτέλεσμα η $X(t)$ αναφέρεται ως η *κατάσταση* της διαδικασίας στον χρόνο t. Για παράδειγμα, η $X(t)$ μπορεί να είναι ίση με το συνολικό αριθμό των πελατών που έχουν μπει σε ένα κατάστημα μέχρι τη χρονική στιγμή t ή ο αριθμός των πελατών στο κατάστημα τη χρονική στιγμή t ή ο αριθμός των προϊόντων που έχουν πωληθεί μέχρι τη στιγμή t κ.τ.λ.

Το σύνολο T καλείται *παραμετρικός χώρος*. Όταν το σύνολο T είναι μετρήσιμο, τότε η στοχαστική διαδικασία καλείται *διαδικασία διακριτού χρόνου*. Διαφορετικά ονομάζεται *διαδικασία συνεχούς χρόνου*. Για παράδειγμα, η $\{X_n, n = 0, 1, \ldots\}$ είναι μια στοχαστική διαδικασία διακριτού χρόνου, ενώ η $\{X(t), t \geq 0\}$ είναι μία συνεχούς χρόνου διαδικασία. Ο *χώρος καταστάσεων* μιας στοχαστικής διαδικασίας είναι το σύνολο όλων των πιθανών τιμών που μπορούν να πάρουν οι τυχαίες μεταβλητές $X(t)$. Συμπερασματικά, μια στοχαστική διαδικασία είναι μια οικογένεια τυχαίων μεταβλητών που περιγράφει την εξέλιξη μέσα στον χρόνο μιας (φυσικής) διαδικασίας.

4.3.2 Διαδικασίες Καταμέτρησης

Μια στοχαστική διαδικασία $\{N(t), t \geq 0\}$ καλείται *διαδικασία καταμέτρησης* αν η τυχαία μεταβλητή $N(t)$ αντιπροσωπεύει το συνολικό αριθμό ‘‘γεγονότων’’ που έχουν συμβεί μέχρι τη χρονική στιγμή t. Παρακάτω δίνονται μερικά παραδείγματα διαδικασιών καταμέτρησης:

- Αν $N(t)$ είναι ο αριθμός των ανθρώπων που έχουν μπει σε ένα κατάστημα πριν τη χρονική στιγμή t, τότε η $\{N(t), t \geq 0\}$ είναι μια διαδικασία καταμέτρησης όπου ένα ‘‘γεγονός ’’ αντιστοιχεί σε κάθε άνθρωπο που μπαίνει στο κατάστημα.

- Αν $N(t)$ είναι ο αριθμός των γκολ που έχει βάλει ένας ποδοσφαιριστής μέχρι το χρόνο t, τότε η $\{N(t), t \geq 0\}$ είναι επίσης μια διαδικασία καταμέτρησης.

Μια διαδικασία καταμέτρησης πρέπει να ικανοποιεί τις εξής συνθήκες:

1. $N(t) \geq 0$.

2. Η $N(t)$ παίρνει ακέραιες τιμές.

3. Αν $s < t$ τότε $N(s) \leq N(t)$.

4. Για $s < t$, η ποσότητα $N(t) - N(s)$ παριστάνει τον αριθμό των γεγονότων που έχουν συμβεί στο διάστημα $(s, t]$.

Ορισμός 4.3.1 *Μια διαδικασία καταμέτρησης λέμε ότι παρουσιάζει* ανεξάρτητες προσαυξήσεις*(independent increments) αν οι αριθμοί των γεγονότων που λαμβάνουν χώρα σε μη επικαλυπτόμενα χρονικά διαστήματα είναι μεταξύ τους ανεξάρτητοι.*

Για παράδειγμα, αυτό σημαίνει ότι ο αριθμός των γεγονότων που έχουν συμβεί μέχρι τη χρονική στιγμή t (δηλαδή $N(t)$) πρέπει να είναι ανεξάρτητος του αριθμού των γεγονότων που συνέβησαν μεταξύ των χρόνων t και $t+s$ (δηλαδή $N(t+s) - N(t)$).

Ορισμός 4.3.2 *Μια διαδικασία καταμέτρησης λέμε ότι παρουσιάζει* στάσιμες προσαυξήσεις*(stationary increments) αν η κατανομή του αριθμού των γεγονότων που συμβαίνουν σε οποιοδήποτε χρονικό διάστημα εξαρτάται μόνο από το μήκος του χρονικού διαστήματος.*

Με άλλα λόγια, μια διαδικασία καταμέτρησης παρουσιάζει στάσιμες προσαυξήσεις αν ο αριθμός των γεγονότων στο διάστημα $(t_1 + s, t_2 + s]$ (δηλαδή $N(t_2 + s) - N(t_1 + s)$) έχει την ίδια κατανομή με τον αριθμό των γεγονότων στο διάστημα $(t_1, t_2]$ (δηλαδή $N(t_2) - N(t_1)$) για όλα τα $t_1 < t_2$ και $s > 0$.

4.3.3 Ορισμός Διαδικασίας Poisson

Ορισμός 4.3.3 *Μια διαδικασία καταμέτρησης* $\{N(t), t \geq 0\}$ *ονομάζεται* διαδικασία Poisson *με ρυθμό* $\lambda > 0$*, αν:*

1. $N(0) = 0$.
2. *Η διαδικασία έχει ανεξάρτητες προσαυξήσεις.*
3. *Ο αριθμός των γεγονότων σε κάθε διάστημα μήκους t ακολουθεί κατανομή Poisson με παράμετρο λt. Δηλαδή για όλα τα $s, t \geq 0$ ισχύει*

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

για $n = 0, 1, \ldots$

Σύμφωνα με την τρίτη συνθήκη η διαδικασία Poisson έχει στάσιμες προσαυξήσεις και ισχύει ότι:

$$E[N(t)] = \lambda t.$$

Για να αποδείξει κανείς ότι μια διαδικασία καταμέτρησης είναι διαδικασία Poisson θα πρέπει να ελέγξει τις παραπάνω τρεις συνθήκες. Οι δύο πρώτες συνθήκες μπορούν συνήθως εύκολα να επαληθευτούν μέσω της γνώσης που έχουμε για τη διαδικασία. Ωστόσο, δεν είναι ξεκάθαρο πώς μπορούμε να επαληθεύσουμε την τρίτη συνθήκη. Για το λόγο αυτό δίνεται παρακάτω ένας εναλλακτικός ορισμός της διαδικασίας Poisson.

Πριν δώσουμε τον ορισμό, πρέπει να ορίσουμε την έννοια των $o(h)$ συναρτήσεων:

Ορισμός 4.3.4 *Μια συνάρτηση $f()$ λέμε ότι είναι $o(h)$ αν*

$$lim_{h\to 0}\frac{f(h)}{h} = 0.$$

Για παράδειγμα, η συνάρτηση $f(x) = x^2$ είναι $o(h)$ ενώ η $f(x) = x$ δεν είναι. Ισχύει ακόμα ότι οποιοσδήποτε γραμμικός συνδυασμός $o(h)$ συναρτήσεων είναι $o(h)$ επίσης. Προφανώς αν μία συνάρτηση $f()$ είναι $o(h)$ τότε $lim_{h\to 0}f(h) = 0$. Μπορούμε τώρα να δώσουμε τον παρακάτω ορισμό για τη διαδικασία Poisson:

Ορισμός 4.3.5 *Μια διαδικασία καταμέτρησης* $\{N(t), t \geq 0\}$ *ονομάζεται* διαδικασία Poisson *με ρυθμό* $\lambda > 0$, *αν:*

1. $N(0) = 0$.
2. *Η διαδικασία έχει ανεξάρτητες και στάσιμες προσαυξήσεις.*
3. $P\{N(h) = 1\} = \lambda h + o(h)$.
4. $P\{N(h) \geq 2\} = o(h)$.

Οι τρεις τελευταίες συνθήκες μπορούν να αντικατασταθούν και από τις:

1. Η διαδικασία έχει ανεξάρτητες προσαυξήσεις.
2. $P\{N(t+h) - N(t) = 1\} = \lambda h + o(h)$.
3. $P\{N(t+h) - N(t) \geq 2\} = o(h)$.

Ουσιαστικά η τρίτη συνθήκη του Ορισμού 4.3.5 ορίζει πως σε ένα ασυμπτωτικά μικρό χρονικό διάστημα h είτε θα συμβεί ένα γεγονός με πιθανότητα λh ή κανένα με πιθανότητα $(1-\lambda h)$. Η πιθανότητα πολλαπλών γεγονότων σύμφωνα με την τέταρτη συνθήκη είναι ασυμπτωτικά μηδενική.

Παράδειγμα 4.3.1 *Αυτοκίνητα διασχίζουν ένα δρόμο σύμφωνα με μια διαδικασία Poisson με ρυθμό* $\lambda = 3$ *αυτοκίνητα ανά λεπτό. Αν ένας λαγός χρειάζεται 5 δευτερόλεπτα για να περάσει από τη μια πλευρά του δρόμου στην απέναντι, ποια είναι η πιθανότητα να μην τραυματιστεί από κάποιο αυτοκίνητο;*

Λύση. Η πιθανότητα που ζητείται από την άσκηση είναι η πιθανότητα να μην έχουμε κανένα γεγονός στο διάστημα $(0, s)$ όπου $s = \frac{5}{60} = \frac{1}{12}$ αφού ο χρόνος μετριέται σε λεπτά για τη διαδικασία Poisson. Επομένως η ζητούμενη πιθανότητα είναι:

$$P\{N(t+s) - N(t) = 0\} = P\left\{N(t + \frac{1}{12}) - N(t) = 0\right\} = e^{-3\frac{1}{12}} \approx 0.78.$$

∎

Παράδειγμα 4.3.2 *Πελάτες καταφθάνουν σε μια τράπεζα σύμφωνα με μία διαδικασία Poisson με ρυθμό* λ *πελάτες ανά ώρα. Έστω ότι καταφθάνουν δύο πελάτες κατά την πρώτη ώρα. Ποια η πιθανότητα ότι:*

1. *και οι δύο κατέφθασαν τα πρώτα 20 λεπτά;*
2. *τουλάχιστον ένας κατέφθασε τα πρώτα 20 λεπτά;*

Λύση. Για το πρώτο ερώτημα η ζητούμενη πιθανότητα είναι $P\{$ 2 αφίξεις στο$(0,s)|$ 2 αφίξεις στο$(0,1)\}$ όπου $s=\frac{1}{3}$ γιατί τα 20 λεπτά είναι το ένα τρίτο της ώρας. Έχουμε ότι:

$$P\left\{ \text{ 2 αφίξεις στο}(0,\frac{1}{3})|\text{ 2 αφίξεις στο}(0,1)\right\} = \frac{P\{\text{ 2 αφίξεις στο}(0,\frac{1}{3}),\text{ 2 αφίξεις στο}(0,1)\}}{P\{\text{ 2 αφίξεις στο}(0,1)\}}$$

$$= \frac{P\{\text{ 2 αφίξεις στο}(0,\frac{1}{3}),\text{ 0 αφίξεις στο}(\frac{1}{3},1)\}}{e^{-\lambda}\frac{\lambda^2}{2}} = \frac{e^{-\lambda\frac{1}{3}}\frac{(\lambda\frac{1}{3})^2}{2}e^{-\lambda\frac{2}{3}}}{e^{-\lambda}\frac{\lambda^2}{2}} = \frac{1}{9}.$$

Η πιθανότητα που ζητείται στο δεύτερο ερώτημα είναι:

$$1-P\left\{ \text{ 0 αφίξεις στο}(0,\frac{1}{3})|\text{ 2 αφίξεις στο}(0,1)\right\}$$

$$=1-\frac{P\{\text{ 0 αφίξεις στο}(0,\frac{1}{3}),\text{ 2 αφίξεις στο}(\frac{1}{3},1)\}}{e^{-\lambda}\frac{\lambda^2}{2}} = 1-\frac{e^{-\lambda\frac{2}{3}}\frac{(\lambda\frac{2}{3})^2}{2}e^{-\lambda\frac{1}{3}}}{e^{-\lambda}\frac{\lambda^2}{2}} = \frac{5}{9}.$$

∎

Παράδειγμα 4.3.3 *Ένα νήπιο κλαίει με συχνότητα 10 φορές την ώρα, σύμφωνα με μία διαδικασία Poisson. Αν η μητέρα του το σηκώνει μόνο αφού κλάψει για τρίτη φορά, ποια η πιθανότητα να μην το σηκώσει καθόλου σε διάστημα 18 λεπτών;*

Λύση. Η πιθανότητα είναι ίση με

$$P\left\{ \text{ 0 ή 1 ή 2 γεγονότα στο}(s,s+\frac{18}{60})\right\} = e^{-10\frac{18}{60}}+e^{-10\frac{18}{60}}10\frac{18}{60}+e^{-10\frac{18}{60}}\frac{(10\frac{18}{60})^2}{2} = 5e^{-2}.$$

∎

4.3.4 Ιδιότητες Διαδικασίας Poisson

Θεωρείστε μια διαδικασία Poisson $\{N(t),t\geq 0\}$ και έστω ότι ο χρόνος άφιξης του πρώτου γεγονότος είναι ίσος με T_1. Γενικότερα, για $n>1$ έστω T_n ο χρόνος που πέρασε από την $(n-1)$-οστή άφιξη μέχρι την n-οστή άφιξη. Η ακολουθία

$$\{T_n, n=1,2,\ldots\}$$

καλείται *ακολουθία των ενδιάμεσων χρόνων*. Για παράδειγμα, αν $T_1=5$ και $T_2=10$, τότε το πρώτο γεγονός της διαδικασίας Poisson έχει συμβεί τη χρονική στιγμή 5 και το δεύτερο τη χρονική στιγμή 15.

Θέλουμε τώρα να προσδιορίσουμε την κατανομή των T_n. Παρατηρούμε αρχικά ότι το γεγονός $\{T_1>t\}$ ισχύει αν και μόνο αν κανένα γεγονός της διαδικασίας Poisson δεν έχει συμβεί στο διάστημα $[0,t]$. Επομένως,

$$P\{T_1>t\}=P\{N(t)=0\}=e^{-\lambda t}.$$

Άρα $P\{T_1 \leq t\} = 1 - e^{-\lambda t}$ και η τυχαία μεταβλητή T_1 ακολουθεί εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$. Για την τυχαία μεταβλητή T_2 ισχύει:

$$P\{T_2 > t\} = \int_{-\infty}^{\infty} P\{T_2 > t | T_1 = s\} f_{T_1}(s) ds.$$

Η πιθανότητα $P\{T_2 > t | T_1 = s\}$ είναι ίση με:

$$P\{T_2 > t | T_1 = s\} = P\{0 \text{ γεγονότα στο } (s, s+t] \,| T_1 = s\}$$

$$= P\{0 \text{ γεγονότα στο } (s, s+t]\} = P\{N(s+t) - N(s) = 0\} = e^{-\lambda t}.$$

Η ισότητα $P\{0 \text{ γεγονότα στο } (s, s+t] \,| T_1 = s\} = P\{0 \text{ γεγονότα στο } (s, s+t]\}$ προκύπτει από το γεγονός ότι η διαδικασία Poisson έχει ανεξάρτητες προσαυξήσεις. Οπότε έχουμε ότι:

$$P\{T_2 > t\} = \int_{-\infty}^{\infty} e^{-\lambda t} f_{T_1}(s) ds = e^{-\lambda t}.$$

Επομένως και η τυχαία μεταβλητή T_2 ακολουθεί εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$ και από τα παραπάνω φάνηκε ότι είναι ανεξάρτητη της T_1. Γενικότερα ισχύει ότι:

Πρόταση 4.3.1 *Οι τυχαίες μεταβλητές T_n, με $n = 1, 2, \ldots$, είναι ανεξάρτητες, εκθετικές τυχαίες μεταβλητές με μέση τιμή $\frac{1}{\lambda}$.*

Κάτι που μας ενδιαφέρει επίσης στην περίπτωση των διαδικασιών Poisson είναι ο χρόνος άφιξης του n-οστού γεγονότος. Προφανώς ισχύει ότι:

$$S_n = \sum_{i=1}^{n} T_i$$

για κάθε $n \geq 1$. Παρατηρούμε ότι το n-οστό γεγονός θα συμβεί πριν (ή κατά) τον χρόνο t αν και μόνο αν ο αριθμός των γεγονότων που θα συμβούν μέχρι τη χρονική στιγμή t είναι τουλάχιστον n. Δηλαδή,

$$N(t) \geq n \Leftrightarrow S_n \leq t.$$

Άρα,

$$F_{S_n}(t) = P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} P\{N(t) = j\}$$

$$= \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}.$$

Παραγωγίζοντας την παραπάνω σχέση έχουμε ότι:

$$f_{S_n}(t) = -\sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!}$$

$$= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{j=n+1}^{\infty} \lambda e^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty} \lambda e^{-\lambda t}\frac{(\lambda t)^{j}}{j!}$$

$$= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}.$$

Άρα οι τυχαίες μεταβλητές S_n ακολουθούν κατανομή Γάμμα με παραμέτρους n και λ.

Παράδειγμα 4.3.4 *Έστω ότι άνθρωποι μεταναστεύουν σε μια περιοχή σύμφωνα με μια διαδικασία Poisson με ρυθμό $\lambda = 1$ ανά μέρα.*

1. *Ποιος είναι ο αναμενόμενος χρόνος μέχρι την άφιξη του δέκατου μετανάστη;*
2. *Ποια η πιθανότητα ο χρόνος μεταξύ της άφιξης του δέκατου και του ενδέκατου μετανάστη να ξεπερνά τις δύο μέρες;*

Λύση. Στο πρώτο ερώτημα ζητείται η τιμή $E[S_{10}]$. Έχουμε ότι:

$$E[S_{10}] = E[T_1 + T_2 + \ldots + T_{10}] = 10E[T_i] = \frac{10}{\lambda} = 10 \text{ μέρες}.$$

Για το δεύτερο ερώτημα πρέπει να υπολογίσουμε την πιθανότητα

$$P\{T_{11} > 2\} = 1 - P\{T_{11} \le 2\} = 1 - (1 - e^{-2\lambda}) = e^{-2} \approx 0.133.$$

∎

Η πρόταση 4.3.1 μπορεί να μας δώσει ένα διαφορετικό τρόπο ορισμού των διαδικασιών Poisson. Αν υποθέσουμε ότι αρχίζουμε με μια ακολουθία $\{T_n, n \ge 1\}$ από ανεξάρτητες, εκθετικές τυχαίες μεταβλητές με μέση τιμή $\frac{1}{\lambda}$ και ορίσουμε μια διαδικασία καταμέτρησης τέτοια ώστε το n-οστό γεγονός συμβαίνει τη χρονική στιγμή $S_n = T_1 + T_2 + \ldots + T_n$, τότε η διαδικασία είναι στοχαστική διαδικασία Poisson με ρυθμό λ.

Επίσης, αν ξέρουμε ότι ένα γεγονός έχει συμβεί σε κάποιο διάστημα $(s, s+t]$ μήκους t, μας ενδιαφέρει σε πολλές περιπτώσεις η κατανομή του χρόνου που συνέβη το γεγονός. Αν θεωρήσουμε ότι $(n-1)$ γεγονότα έχουν συμβεί μέχρι τη χρονική στιγμή s, τότε μέχρι το χρόνο t θα έχουν συμβεί n γεγονότα. Έστω $(s, s+t_1]$ ένα υποδιάστημα του $(s, s+t]$ (δηλαδή $t_1 < t$). Ισχύει ότι:

$$P\{T_n < t_1 | N(s+t) - N(s) = 1\} = \frac{P\{T_n < t_1, N(s+t) - N(s) = 1\}}{P\{N(s+t) - N(s) = 1\}}$$

$$= \frac{P\{\ 1 \text{ γεγονός στο}(s, s+t_1),\ 0 \text{ γεγονότα στο}(s+t_1, s+t)\}}{\lambda t e^{-\lambda t}}$$

$$= \frac{P\{\ 1 \text{ γεγονός στο}(s, s+t_1)\}\, P\{\ 0 \text{ γεγονότα στο}(s+t_1, s+t)\}}{\lambda t e^{-\lambda t}}$$

$$= \frac{\lambda t_1 e^{-\lambda t_1} e^{-\lambda(t-t_1)}}{\lambda t e^{-\lambda t}} = \frac{t_1}{t}.$$

Η κατανομή δηλαδή είναι ομοιόμορφη. Για παράδειγμα, αν γνωρίζουμε ότι ένα γεγονός έχει συμβεί σε διάστημα μίας ώρας, η πιθανότητα να συνέβη το πρώτο μισάωρο είναι 0.5 ή η πιθανότητα να συνέβη το πρώτο τέταρτο είναι 0.25 κ.τ.λ.

Παράδειγμα 4.3.5 *Γεγονότα συμβαίνουν σύμφωνα με μία στοχαστική διαδικασία Poisson με ρυθμό $\lambda = 3$ γεγονότα ανά ώρα.*

1. *Ποια η πιθανότητα ότι δε συμβαίνει τίποτα μεταξύ 8 και 10 το πρωί;*

2. *Ποια η μέση τιμή του αριθμού των γεγονότων που συμβαίνουν μεταξύ 8 και 10 το πρωί;*

3. *Ποια η αναμενόμενη ώρα εμφάνισης του πέμπτου γεγονότος μετά από τις 2 το απόγευμα;*

Λύση.

1. Η πιθανότητα που ζητείται είναι ίση με:

$$P\{N(10) - N(8) = 0\} = e^{-6}.$$

2. Η μέση τιμή του αριθμού των γεγονότων που συμβαίνουν μεταξύ 8 και 10 είναι η $E[N(10) - N(8)]$. Γνωρίζουμε όμως ότι ο αριθμός των γεγονότων που συμβαίνουν σε ένα διάστημα μήκους t ακολουθεί κατανομή Poisson με παράμετρο λt. Επομένως, $E[N(10) - N(8)] = 2\lambda = 6$.

3. Ο αναμενόμενος χρόνος μέχρι την άφιξη του πέμπτου γεγονότος είναι $E[S_5] = E[T_1 + T_2 + T_3 + T_4 + T_5] = 5\frac{1}{3}$, δηλαδή μία ώρα και 40 λεπτά. Άρα η αναμενόμενη ώρα εμφάνισης του πέμπτου γεγονότος μετά τις 2 είναι η 3:40.

■

Παράδειγμα 4.3.6 *Έστω δύο ανεξάρτητες στοχαστικές διαδικασίες Poisson $\{N(t), t \geq 0\}$ και $\{M(t), t \geq 0\}$ με ρυθμούς αφίξεων a και b αντίστοιχα. Ποια η πιθανότητα η πρώτη άφιξη να προέρχεται από τη στοχαστική διαδικασία $\{N(t), t \geq 0\}$ και ποια η πιθανότητα ότι οι επόμενες τρεις αφίξεις προέρχονται από την ίδια στοχαστική διαδικασία (δηλαδή είτε από την $\{N(t), t \geq 0\}$ είτε από την $\{M(t), t \geq 0\}$);*

Λύση. Ας συμβολίσουμε με T_n και $T_n^{'}$ τους χρόνους μεταξύ των αφίξεων για τις στοχαστικές διαδικασίες $\{N(t), t \geq 0\}$ και $\{M(t), t \geq 0\}$ αντίστοιχα. Οι τυχαίες μεταβλητές T_n και $T_n^{'}$ είναι εκθετικά κατανεμημένες με παραμέτρους a και b αντίστοιχα. Επομένως, η πρώτη πιθανότητα που ζητείται από την άσκηση είναι ίση με $P\left\{T_1 < T_1^{'}\right\}$. Από προηγούμενη άσκηση γνωρίζουμε ότι η πιθανότητα $P\{X < Y\}$ για δύο εκθετικά κατανεμημένες τυχαίες μεταβλητές είναι ίση με $\frac{a}{a+b}$ αν a είναι η παράμετρος της X και b της Y. Άρα, $P\left\{T_1 < T_1^{'}\right\} = \frac{a}{a+b}$.

Όταν συμβαίνει μία άφιξη, αυτή προέρχεται από τη διαδικασία $\{N(t), t \geq 0\}$ με πιθανότητα $\frac{a}{a+b}$ και με πιθανότητα $\frac{b}{a+b}$ από τη διαδικασία $\{M(t), t \geq 0\}$. Όμως το γεγονός μίας άφιξης είναι ανεξάρτητο από την μία άφιξη στην επόμενη. Επομένως, η πιθανότητα οι τρεις πρώτες αφίξεις να προέρχονται από την στοχαστική διαδικασία $\{N(t), t \geq 0\}$ είναι ίση με $\left(\frac{a}{a+b}\right)^3$ και ίση με $\left(\frac{b}{a+b}\right)^3$ αν προέρχονται από την $\{M(t), t \geq 0\}$. Δηλαδή η ζητούμενη πιθανότητα είναι ίση με $\left(\frac{a}{a+b}\right)^3 + \left(\frac{b}{a+b}\right)^3$. ∎

4.3.5 Διάσπαση Διαδικασιών Poisson

Έστω μια διαδικασία Poisson $\{N(t), t \geq 0\}$ ρυθμού λ και υποθέστε ότι κάθε γεγονός της διαδικασίας ανήκει στην κατηγορία 1 με πιθανότητα p και στην κατηγορία 2 με πιθανότητα $1-p$. Για παράδειγμα, αν θεωρήσουμε ότι πελάτες μπαίνουν σε ένα κατάστημα σύμφωνα με μια διαδικασία Poisson ρυθμού λ, τότε κάθε πελάτης (δηλαδή κάθε γεγονός της διαδικασίας) είναι με πιθανότητα $\frac{1}{2}$ άνδρας (κατηγορία 1) και με πιθανότητα $\frac{1}{2}$ γυναίκα (κατηγορία 2). Αν συμβολίσουμε με $N_1(t)$ το πλήθος των γεγονότων που ανήκουν στην κατηγορία 1 και με $N_2(t)$ το πλήθος των γεγονότων που ανήκουν στην κατηγορία 2 μέσα στο χρονικό διάστημα $[0, t]$ τότε ισχύει το εξής:

Πρόταση 4.3.2 *Οι $\{N_1(t), t \geq 0\}$ και $\{N_2(t), t \geq 0\}$ είναι διαδικασίες Poisson με ρυθμούς λp και $\lambda(1-p)$ αντίστοιχα. Επιπλέον, οι διαδικασίες είναι ανεξάρτητες μεταξύ τους.*

Απόδειξη. Θα αποδείξουμε την πρόταση για την περίπτωση της κατηγορίας 1. Η ίδια απόδειξη ισχύει και για την δεύτερη κατηγορία. Σύμφωνα με τον ορισμό 4.3.5 πρέπει να αποδείξουμε τα εξής:

1. $N_1(0) = 0$. Ισχύει λόγω του ότι $N(0) = 0$.

2. Αφού οι αριθμοί των γεγονότων που συμβαίνουν σε μη επικαλυπτόμενα διαστήματα είναι ανεξάρτητοι μεταξύ τους και εξαρτώνται μόνο από το μήκος του διαστήματος, το ίδιο θα συμβαίνει και για τα γεγονότα που ανήκουν στην κατηγορία 1. Άρα η διαδικασία $\{N_1(t), t \geq 0\}$ έχει ανεξάρτητες και στάσιμες προσαυξήσεις.

3. $P\{N_1(h) = 1\} = \sum_{n=0}^{\infty} P\{N_1(h) = 1 | N(h) = n\} P\{N(h) = n\}$
$= P\{N_1(h) = 1 | N(h) = 1\} P\{N(h) = 1\} + P\{N_1(h) = 1 | N(h) \geq 2\} P\{N(h) \geq 2\}$
$= p(\lambda h + o(h)) + o(h) = \lambda p h + o(h)$.

4. $P\{N_1(h) \geq 2\} \leq P\{N(h) \geq 2\} = o(h)$.

■

Το παραπάνω συμπέρασμα γενικεύεται για περισσότερες από δύο κατηγορίες. Δηλαδή, αν κάθε γεγονός της διαδικασίας ανήκει στην κατηγορία 1 με πιθανότητα p_1, στην κατηγορία 2 με πιθανότητα p_2, στην κατηγορία 3 με πιθανότητα p_3 κ.ο.κ. μέχρι την κατηγορία n με πιθανότητα p_n, τότε οι $\{N_k(t), t \geq 0\}$ για $k = 1, 2, \ldots, n$ είναι διαδικασίες Poisson με ρυθμούς λp_k αντίστοιχα. Αντίστροφα, αν θεωρήσουμε την υπέρθεση n διαδικασιών Poisson με ρυθμούς $\lambda_1, \lambda_2, ..., \lambda_n$, τότε η διαδικασία που θα προκύψει είναι επίσης διαδικασία Poisson με ρυθμό $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.

Παράδειγμα 4.3.7 *Έστω ότι μετανάστες φτάνουν σε μια περιοχή σύμφωνα με μια διαδικασία Poisson με ρυθμό $\lambda = 10$ την βδομάδα. Αν γνωρίζουμε ότι ένας στους δώδεκα μετανάστες είναι Κινέζος, ποια η πιθανότητα να μην φτάσει στην περιοχή κανένας Κινέζος στη διάρκεια του Δεκέμβρη;*

Λύση. Σύμφωνα με την προηγούμενη πρόταση, το πλήθος των Κινέζων ακολουθεί διαδικασία Poisson με ρυθμό $10\frac{1}{12}$ τη βδομάδα, δηλαδή $40\frac{1}{12} = \frac{10}{3}$ το μήνα. Επομένως η πιθανότητα που ζητείται είναι ίση με:

$$P\{\ 0 \text{ γεγονότα σε ένα μήνα}\} = e^{-\frac{10}{3}}.$$

■

Κεφάλαιο 5

Προσομοίωση

5.1 Εισαγωγή

Οι τυχαίοι αριθμοί χρησιμοποιούνται σε πολλές εφαρμογές στην πληροφορική και στην στατιστική. Στο κεφάλαιο αυτό θα δούμε πως μπορούν να δημιουργηθούν τυχαίοι αριθμοί που ακολουθούν κάποια συγκεκριμένη κατανομή. Οι μέθοδοι που χρησιμοποιούνται για το σκοπό αυτό, μπορούν να χρησιμοποιηθούν επίσης για την αριθμητική επίλυση διαφόρων προβλημάτων. Έστω για παράδειγμα, ότι θέλουμε να υπολογίσουμε το ολοκλήρωμα

$$I = \int_0^1 g(x)dx.$$

Για το σκοπό αυτό, ορίζουμε μία τυχαία μεταβλητή X που ακολουθεί την ομοιόμορφη κατανομή στο $(0,1)$ και σχηματίζουμε την τυχαία μεταβλητή $Y = g(X)$. Τότε, ισχύει ότι

$$E\left[g(X)\right] = \int_0^1 g(x)f_X(x)dx = \int_0^1 g(x)dx = I.$$

Έτσι η άγνωστη τιμή I που θέλουμε να υπολογίσουμε μπορεί να εκφραστεί ως η αναμενόμενη τιμή της τυχαίας μεταβλητής Y.

Αν η συνάρτηση $g()$ έχει n μεταβλητές, τότε ο υπολογισμός του ολοκληρώματος

$$E\left[g(X)\right] = \int\int \ldots \int g(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$$

γίνεται ακόμα πιο πολύπλοκος. Επομένως, σε πολλές περιπτώσεις δεν είναι δυνατό να υπολογιστεί το πολλαπλό ολοκλήρωμα με μεγάλη ακρίβεια. Στην περίπτωση αυτή, ο μόνος αποδοτικός τρόπος να υπολογίσουμε την τιμή $E\left[g(X)\right]$ είναι με κάποια μέθοδο προσομοίωσης.

Συγκεκριμένα, για τον υπολογισμό της τιμής $E\left[g(X)\right] = \int_0^1 g(x)dx$ μπορούμε να ακολουθήσουμε την εξής διαδικασία: να δημιουργήσουμε n τυχαίους αριθμούς $x_i \in (0,1)$, τις αντίστοιχες τιμές $y_i = g(x_i)$ και να προσεγγίσουμε την αναμενόμενη τιμή ως το άθροισμα $E[g(X)] \approx \frac{1}{n}\sum_{i=1}^{n} g(x_i)$. Η μέθοδος αυτή καλείται *προσομοίωση Monte Carlo*.

Το πρώτο βήμα για την προσομοίωση τυχαίων μεταβλητών είναι η δημιουργία τυχαίων μεταβλητών που ακολουθούν την ομοιόμορφη κατανομή στο $(0,1)$. Βασιζόμενοι σε τυχαίους αριθμούς, ομοιόμορφα κατανεμημένους στο $(0,1)$, μπορούμε στη συνέχεια να δημιουργήσουμε τυχαίους αριθμούς που ακολουθούν οποιαδήποτε κατανομή. Ένας απλός τρόπος για τη δημιουργία πραγματικά τυχαίων αριθμών στο $(0,1)$ θα ήταν να βασιστούμε σε ένα τυχαίο πείραμα όπως το παρακάτω: να βάλουμε σε ένα καπέλο 10 μπάλες, αριθμημένες από το 0 μέχρι το 9, να διαλέγουμε κάθε φορά μία μπάλα και να την ξαναβάζουμε πίσω. Η ακολουθία των αριθμών που θα προκύψουν, μπορούμε να θεωρήσουμε ότι είναι τα δεκαδικά ψηφία ενός αριθμού στο $(0,1)$. Για παράδειγμα, αν έχουμε επιλέξει τις μπάλες 3,8,7,1,3 ο τυχαίος αριθμός 0.38713 που θα προκύψει, μπορεί να θεωρηθεί ως η τιμή μιας τυχαίας μεταβλητής ομοιόμορφα κατανεμημένης στο $(0,1)$.

Ωστόσο, οι υπολογιστές δεν μπορούν να προσομοιώσουν τυχαίες μεταβλητές στο $(0,1)$ με κάποιο παρόμοιο, πραγματικά τυχαίο τρόπο. Στην πράξη, παράγουν ψευδοτυχαίους αριθμούς, δηλαδή αριθμούς που προκύπτουν μέσω μιας ντετερμινιστικής διαδικασίας, ενός προγράμματος ουσιαστικά, που έχει μικρές απαιτήσεις μνήμης και χρησιμοποιεί απλές αριθμητικές μεθόδους. Ένα τέτοιο πρόγραμμα καλείται γεννήτρια ψευδοτυχαίων αριθμών και θα πρέπει να είναι έτσι σχεδιασμένο ώστε κανένα στατιστικό τεστ να μην μπορεί να ξεχωρίσει τους ψευδοτυχαίους αριθμούς που παράγει από πραγματικά τυχαίους αριθμούς.

Οι περισσότερες γεννήτριες ψευδοτυχαίων αριθμών αρχίζουν από μια τιμή x_0 που καλείται seed και στην συνέχεια υπολογίζουν επαναληπτικά τιμές

$$x_{n+1} = (ax_n + c) \pmod{m}, n \geq 0$$

για θετικούς ακεραίους a, c και m. Η πράξη $mod\ m$ μας δίνει το υπόλοιπο της διαίρεσης του $ax_n + c$ με το m. Κάθε x_n είναι ένας ακέραιος από $0, 1, \ldots, m-1$ και η τιμή $\frac{x_n}{m}$ αποτελεί προσέγγιση μιας τιμής μιας ομοιόμορφης τυχαίας μεταβλητής στο $(0,1)$.

Για να προσομοιώσουμε τυχαίες μεταβλητές από μια συγκεκριμένη κατανομή, υποθέτουμε ότι μπορούμε να προσομοιώσουμε τυχαίες μεταβλητές που είναι ομοιόμορφες στο $(0,1)$ (για παράδειγμα με τον τρόπο που περιγράφηκε προηγουμένως). Στη συνέχεια του κεφαλαίου, όταν θα μιλάμε για τυχαίους αριθμούς θα εννοούμε τιμές ομοιόμορφων μεταβλητών στο $(0,1)$.

Παράδειγμα Χρήσης Τυχαίων Αριθμών: Έστω ότι θέλουμε να δημιουργήσουμε μία μετάθεση των αριθμών $1, 2, \ldots, n$, τέτοια ώστε όλοι οι $n!$ συνδυασμοί να είναι εξίσου πιθανοί. Για να λύσουμε αυτό το πρόβλημα, πρώτα επιλέγουμε έναν αριθμό από 1 ως n τυχαία και τον τοποθετούμε στη θέση n. Στη συνέχεια επιλέγεται τυχαία κάποιος από τους υπόλοιπους $n-1$ αριθμούς, τοποθετείται στη θέση $n-1$ κ.ο.κ. Ωστόσο, για να μην αντιμετωπίζουμε το πρόβλημα ποιος αριθμός έχει μείνει για να τοποθετηθεί, είναι βολικό να κρατάμε τους αριθμούς σε έναν πίνακα και στη συνέχεια να επιλέγουμε τη θέση που βρίσκεται ο αριθμός και όχι τον ίδιο τον αριθμό. Δηλαδή, αρχίζοντας από μια διάταξη $p_1, p_2, \ldots, p_n$ των n αριθμών, διαλέγουμε μία από τις θέσεις $1, \ldots, n$ τυχαία και αλλάζουμε τον αριθμό αυτής της θέσης με τον αριθμό στη θέση n. Στη συνέχεια, διαλέγουμε μία από τις θέσεις $1, \ldots, n-1$ και αλλάζουμε τον αριθμό που βρίσκεται σε αυτή τη θέση με τον αριθμό στη θέση $n-1$ κ.ο.κ.

Για να υλοποιήσουμε τον αλγόριθμο που μόλις περιγράφηκε, θα πρέπει να μπορούμε να δημιουργήσουμε μία τυχαία μεταβλητή που είναι εξίσου πιθανό να πάρει μία από τις τιμές $1, 2, \ldots, k$. Παρατηρήστε ότι αν U είναι ένας τυχαίος αριθμός στο $(0, 1)$, τότε ο αριθμός kU είναι ένας τυχαίος αριθμός στο $(0, k)$. Δηλαδή, αποτελεί μια τιμή μιας ομοιόμορφης τυχαίας μεταβλητής στο $(0, k)$. Επομένως, $P\{i-1 < kU < i\} = \frac{1}{k}$ με $i = 1, \ldots, k$. Αν ορίσουμε μία τυχαία μεταβλητή I ως $I = \lfloor kU \rfloor + 1$ τότε:

$$P\{I = i\} = P\{\lfloor kU \rfloor = i-1\} = P\{i-1 < kU < i\} = \frac{1}{k}.$$

Καταλήγουμε τελικά στον ακόλουθο αλγόριθμο:

ΑΛΓΟΡΙΘΜΟΣ 5.1.1 ***Αλγόριθμος δημιουργίας μιας τυχαίας μετάθεσης n αριθμών***
Είσοδος: *Μια διάταξη $p_1, p_2, \ldots, p_n$ των αριθμών $1, 2, \ldots, n$*
Έξοδος: *Μια τυχαία μετάθεση των n αριθμών*

Βήμα 1: Θέσε $k = n$
Βήμα 2: Δημιούργησε έναν τυχαίο αριθμό U στο $(0, 1)$ και θέσε $I = \lfloor kU \rfloor + 1$
Βήμα 3: Άλλαξε τις τιμές στις θέσεις p_I και p_k
Βήμα 4: Θέσε $k = k - 1$ και αν $k > 1$ πήγαινε στο βήμα 2
Βήμα 5: Επέστρεψε την νέα διάταξη $p_1, p_2, \ldots, p_n$ των n αριθμών

Για παράδειγμα, έστω ότι $n = 4$ και η αρχική διάταξη των αριθμών είναι η 1,2,3,4. Αν η πρώτη τιμή του I είναι ίση με 3, τότε στο βήμα 3 η νέα διάταξη που θα προέκυπτε θα ήταν η 1,2,4,3. Αν στη συνέχεια είχαμε ότι $I = 2$, η νέα μετάθεση που θα προέκυπτε θα ήταν η 1,4,2,3 και αν στην τελευταία επανάληψη υπολογίζαμε $I = 2$ πάλι, η τελική μετάθεση των αριθμών θα ήταν η 1,4,2,3 (για $I = 2$ δεν αλλάζει η σειρά των δύο πρώτων στοιχείων).

Ο προηγούμενος αλγόριθμος μπορεί να χρησιμοποιηθεί για τη δημιουργία ενός τυχαίου υποσυνόλου μεγέθους r, των ακεραίων $1, \ldots, n$. Συγκεκριμένα, εφαρμόζεται ο αλγόριθμος μέχρι να καλυφθούν οι θέσεις $n, n-1, \ldots, n-r+1$ και τα στοιχεία αυτών των θέσεων αποτελούν το τυχαίο υποσύνολο που θέλουμε να δημιουργήσουμε.

5.2 Προσομοίωση Συνεχών Τυχαίων Μεταβλητών

5.2.1 Η Μέθοδος του Αντίστροφου Μετασχηματισμού

Μια γενική μέθοδος για την προσομοίωση συνεχών τυχαίων μεταβλητών είναι η *μέθοδος του αντίστροφου μετασχηματισμού*. Η μέθοδος αυτή βασίζεται στην παρακάτω πρόταση:

Πρόταση 5.2.1 *Έστω U μια ομοιόμορφη τυχαία μεταβλητή στο $(0,1)$. Για κάθε συνεχή συνάρτηση κατανομής F, αν ορίσουμε την τυχαία μεταβλητή X ως:*

$$X = F^{-1}(U)$$

τότε η τυχαία μεταβλητή X έχει συνάρτηση κατανομής την F.

Απόδειξη. Για τη συνάρτηση κατανομής της τυχαίας μεταβλητής X ισχύει ότι:

$$F_X(a) = P\left\{X \le a\right\} = P\left\{F^{-1}(U) \le a\right\}.$$

Λόγω του ότι οι συναρτήσεις κατανομής είναι αύξουσες συναρτήσεις, προκύπτει ότι $F^{-1}(U) \le a$ αν και μόνο αν $U \le F(a)$. Έτσι έχουμε ότι:

$$F_X(a) = P\left\{F^{-1}(U) \le a\right\} = P\left\{U \le F(a)\right\} = F(a).$$

Επομένως, αποδείχθηκε ότι η συνάρτηση κατανομής της τυχαίας μεταβλητής X είναι ίση με την F. ∎

Συνοψίζοντας, μπορούμε να προσομοιώσουμε μία τυχαία μεταβλητή X που έχει συνάρτηση κατανομής F, όταν η συνάρτηση F^{-1} μπορεί να υπολογιστεί εύκολα, προσομοιώνοντας μία ομοιόμορφη τυχαία μεταβλητή U και θέτοντας $X = F^{-1}(U)$.

Παράδειγμα 5.2.1 *Προσομοιώστε μία εκθετική τυχαία μεταβλητή με παράμετρο 1 με τη μέθοδο του αντίστροφου μετασχηματισμού.*

Λύση. Η συνάρτηση κατανομής μιας εκθετικής τυχαίας μεταβλητής X με παράμετρο 1 είναι η $F(x) = 1 - e^{-x}$. Για να βρούμε την αντίστροφη συνάρτηση $F^{-1}(u)$ λύνουμε την εξίσωση $1 - e^{-x} = u$ ως προς x και έχουμε ότι $x = -\log(1-u)$. Επομένως, η $X = F^{-1}(U) = -\log(1-U)$ είναι εκθετική με μέση τιμή 1. Επειδή η τυχαία μεταβλητή $1-U$ είναι (όπως και η U) ομοιόμορφη στο $(0,1)$ έχουμε ότι η $X = -\log U$ είναι επίσης εκθετική με παράμετρο 1. Αν η παράμετρος της εκθετικής τυχαίας μεταβλητής ήταν λ, τότε θα είχαμε ότι $X = -\frac{\log U}{\lambda}$. ∎

5.2.2 Η Μέθοδος Απόρριψης

Ο υπολογισμός της αντίστροφης συνάρτησης F^{-1} για κάθε συνάρτηση κατανομής F δεν είναι πάντα εύκολος. Στις περιπτώσεις αυτές χρησιμοποιείται μια διαφορετική μέθοδος για την προσομοίωση των τυχαίων μεταβλητών που καλείται *μέθοδος απόρριψης*. Η μέθοδος απόρριψης βασίζεται στην εξής ιδέα: αν υποθέσουμε ότι έχουμε μία μέθοδο για την προσομοίωση μιας τυχαίας μεταβλητής με συνάρτηση πυκνότητας πιθανότητας $g(x)$, τότε μπορούμε να την χρησιμοποιήσουμε ως βάση για την προσομοίωση μιας τυχαίας μεταβλητής που έχει συνάρτηση πυκνότητας πιθανότητας $f(x)$. Μπορούμε να προσομοιώσουμε μια τυχαία μεταβλητή Y με τη συνάρτηση $g(x)$ και στη συνέχεια να δεχτούμε την τιμή που προέκυψε με πιθανότητα ανάλογη της τιμής $f(Y)/g(Y)$.

Συγκεκριμένα, έστω c μια σταθερά για την οποία

$$\frac{f(y)}{g(y)} \leq c$$

για όλες τις πιθανές τιμές της y. Τότε προκύπτει η εξής μέθοδος για την προσομοίωση μιας τυχαίας μεταβλητής X με συνάρτηση πυκνότητας πιθανότητας $f(x)$:

Βήμα 1: Προσομοίωσε μια τυχαία μεταβλητή Y με συνάρτηση πυκνότητας πιθανότητας $g(x)$ και μια ομοιόμορφη τυχαία μεταβλητή U στο $(0,1)$.
Βήμα 2: Αν $U \leq \frac{f(Y)}{cg(Y)}$ θέσε $X = Y$.Διαφορετικά γύρνα στο βήμα 1.

Πρόταση 5.2.2 *Η τυχαία μεταβλητή X που δημιουργείται με τη μέθοδο απόρριψης έχει συνάρτηση πυκνότητας πιθανότητας την $f(x)$.*

Απόδειξη. Θέλουμε να αποδείξουμε ότι $P\{X \leq x\} = \int_{-\infty}^{x} f(y)dy$. Έστω N ο αριθμός των επαναλήψεων που χρειάστηκαν για τον υπολογισμό μιας τιμής X. Τότε:

$$P\{X \leq x\} = P\{Y_N \leq x\} = P\left\{Y \leq x | U \leq \frac{f(Y)}{cg(Y)}\right\} =$$

$$\frac{P\left\{Y \leq x, U \leq \frac{f(Y)}{cg(Y)}\right\}}{K}$$

όπου $K = P\left\{U \leq \frac{f(Y)}{cg(Y)}\right\}$. Για να υπολογίσουμε την πιθανότητα $P\left\{Y \leq x, U \leq \frac{f(Y)}{cg(Y)}\right\}$ χρησιμοποιούμε δεσμευμένη πιθανότητα ως προς την τυχαία μεταβλητή Y. Δηλαδή,

$$P\left\{Y \leq x, U \leq \frac{f(Y)}{cg(Y)}\right\} = \int_{-\infty}^{\infty} P\left\{Y \leq x, U \leq \frac{f(Y)}{cg(Y)} | Y = y\right\} g(y)dy =$$

$$\int_{-\infty}^{x} P\left\{U \leq \frac{f(y)}{cg(y)}\right\} g(y)dy = \int_{-\infty}^{x} \frac{f(y)}{c} dy.$$

Άρα, $P\{X \leq x\} = \frac{\int_{-\infty}^{x} f(y)dy}{Kc}$. Αν $x \to \infty$ τότε $P\{X \leq x\} = 1$ και $\int_{-\infty}^{x} f(y)dy = 1$. Επομένως, προκύπτει ότι $Kc = 1$ και $P\{X \leq x\} = \int_{-\infty}^{x} f(y)dy$. ∎

Με την προηγούμενη απόδειξη, δείξαμε ότι η πιθανότητα $K = P\left\{U \leq \frac{f(Y)}{cg(Y)}\right\} = \frac{1}{c}$. Η σχέση αυτή καταδεικνύει ότι κατά μέσο όρο η μέθοδος απόρριψης θα χρειαστεί c επαναλήψεις για να βρει μια τιμή X. Σημειώνεται επίσης ότι για να εφαρμόσουμε τη μέθοδο απόρριψης για δύο συναρτήσεις $f(x)$ και $g(x)$, θα πρέπει οι συναρτήσεις αυτές να έχουν το ίδιο πεδίο ορισμού έτσι ώστε για κάθε τιμή y που ορίζεται η $g(y)$ να ορίζεται και η $f(y)$ και το αντίστροφο.

Παράδειγμα 5.2.2 *Περιγράψτε τη μέθοδο απόρριψης για τη δημιουργία μίας τυχαίας μεταβλητής με συνάρτηση πυκνότητας πιθανότητας* $f(x) = 20x(1-x)^3$, *με* $0 < x < 1$.

Λύση. Επειδή η $f(x)$ ορίζεται στο $(0,1)$, θεωρούμε ως $g(x) = 1$ τη συνάρτηση πυκνότητας πιθανότητας μιας ομοιόμορφης τυχαίας μεταβλητής στο $(0,1)$. Αρχικά, πρέπει να βρούμε τη σταθερά c για την οποία $\frac{f(x)}{g(x)} \leq c$ για κάθε $x \in (0,1)$. Έχουμε ότι $\frac{f(x)}{g(x)} = 20x(1-x)^3$. Για να υπολογίσουμε το μέγιστο αυτής της συνάρτησης, βρίσκουμε καταρχήν την παράγωγο $\left(\frac{f(x)}{g(x)}\right)' = 20\left[(1-x)^3 - 3x(1-x)^2\right] = 20(1-x)^2(1-4x)$. Το μέγιστο της συνάρτησης βρίσκεται στο σημείο x_0 για το οποίο $\left(\frac{f(x_0)}{g(x_0)}\right)' = 0$, δηλαδή στο σημείο $x_0 = \frac{1}{4}$. Άρα, $\frac{f(x)}{g(x)} \leq 20\frac{1}{4}\left(\frac{3}{4}\right)^3 = \frac{135}{64} = c$ και

$$\frac{f(x)}{cg(x)} = \frac{256}{27}x(1-x)^3.$$

Επομένως, η μέθοδος απόρριψης είναι η:

Βήμα 1: Προσομοίωσε δύο ομοιόμορφες τυχαίες μεταβλητές Y και U στο $(0,1)$.
Βήμα 2: Αν $U \leq \frac{256}{27}Y(1-Y)^3$ θέσε $X = Y$.Διαφορετικά γύρνα στο βήμα 1.
∎

Παράδειγμα 5.2.3 *Χρησιμοποιήστε τη μέθοδο της απόρριψης για να προσομοιώσετε μια τυπική κανονική κατανομή.*

Λύση. Η συνάρτηση πυκνότητας πιθανότητας μιας συνεχούς τυχαίας μεταβλητής X που ακολουθεί την τυπική κανονική κατανομή είναι η $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ και ορίζεται για $-\infty < x < \infty$. Από προηγούμενη άσκηση γνωρίζουμε μία μέθοδο προσομοίωσης για την εκθετική τυχαία μεταβλητή με παράμετρο 1, η οποία όμως ορίζεται στο διάστημα $0 < x < \infty$. Άρα, μπορούμε να προσομοιώσουμε με τη μέθοδο της απόρριψης την τυχαία μεταβλητή $Z = |X|$ που ορίζεται στο διάστημα $0 < x < \infty$ και στη συνέχεια να βρούμε την τιμή της X που θα είναι είτε ίση με Z ή με $-Z$.

Η τυχαία μεταβλητή Z θα έχει συνάρτηση πυκνότητας πιθανότητας

$$f_Z(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{5.1}$$

Αυτό προκύπτει άμεσα από το ότι $P\{Z \leq x\} = P\{|X| \leq x\} = \int_{-x}^{x} f(y)dy = \int_0^x 2f(y)dy = \int_0^x f_Z(y)dy$. Ως $g(x)$ χρησιμοποιούμε τη συνάρτηση πυκνότητας πιθανότητας μιας εκθστικής τυχαίας μεταβλητής με παράμετρο 1, δηλαδή $g(x) = e^{-x}$. Προκύπτει ότι $\frac{f_Z(x)}{g(x)} = \sqrt{2e/\pi} e^{-\frac{(x-1)^2}{2}} \leq \sqrt{2e/\pi} = c$. Άρα η μέθοδος απόρριψης για την προσομοίωση της τυχαίας μεταβλητής Z είναι:

Βήμα 1: Προσομοίωσε μία εκθετική τυχαία μεταβλητή Y με παράμετρο 1 και μια ομοιόμορφη U στο $(0, 1)$.
Βήμα 2: Αν $U \leq e^{-\frac{(U-1)^2}{2}}$ θέσε $Z = Y$.Διαφορετικά γύρνα στο βήμα 1.

Αφού προσομοιώσαμε μια τυχαία μεταβλητή Z με συνάρτηση πυκνότητας πιθανότητας που δίνεται από τη σχέση 5.1, μπορούμε τώρα να προσομοιώσουμε την X επιλέγοντας με τυχαίο τρόπο η X να είναι ίση είτε με Z ή με $-Z$. Για παράδειγμα, μπορούμε να δημιουργήσουμε μία τυχαία τιμή u στο $(0, 1)$ και αν $u < 0.5$ τότε να θέσουμε $X = Z$, διαφορετικά $X = -Z$. Αν θέλαμε να προσομοιώσουμε μία κανονική τυχαία μεταβλητή με παραμέτρους μ και σ^2, τότε θα υπολογίζαμε την τιμή $\mu + \sigma X$. ∎

5.2.3 Η Μέθοδος Ρυθμού Βλάβης

Έστω F μια συνάρτηση κατανομής. Υπενθυμίζουμε ότι η συνάρτηση ρυθμού βλάβης $\lambda(t)$ της F, ορίζεται από τη σχέση

$$\lambda(t) = \frac{f(t)}{1 - F(t)}$$

με $t \geq 0$. Υποθέστε ότι μας δίνεται μία συνάρτηση $\lambda(t)$, τέτοια ώστε $\int_0^\infty \lambda(t)dt = \infty$, και ότι θέλουμε να προσομοιώσουμε μία τυχαία μεταβλητή S που έχει ως συνάρτηση ρυθμού βλάβης την $\lambda(t)$.

Για να το πετύχουμε αυτό, θεωρούμε αρχικά ότι λ_1 είναι ένας αριθμός τέτοιος ώστε

$$\lambda(t) \leq \lambda_1$$

για όλες τις τιμές $t \geq 0$. Για να προσομοιώσουμε μία τυχαία μεταβλητή μέσω της $\lambda(t)$ θα ακολουθήσουμε τα εξής βήματα:

Βήμα 1: Προσομοιώνουμε μία διαδικασία Poisson με ρυθμό λ_1.
Βήμα 2: Θα μετράμε μόνο ορισμένα από τα γεγονότα της διαδικασίας.
Συγκεκριμένα, θα μετράμε ένα γεγονός που συμβαίνει στο χρόνο t με πιθανότητα $\frac{\lambda(t)}{\lambda_1}$.

Ένα ερώτημα που προκύπτει άμεσα είναι πώς μπορούμε να προσομοιώσουμε μία στοχαστική διαδικασία Poisson με ρυθμό λ. Ο απλούστερος τρόπος είναι να προσομοιώσουμε τους χρόνους μεταξύ των αφίξεων T_i για $i \geq 1$. Οι μεταβλητές T_i γνωρίζουμε από το προηγούμενο κεφάλαιο ότι είναι εκθετικές τυχαίες μεταβλητές με παράμετρο λ. Επομένως, μπορούν να κατασκευαστούν με τη μέθοδο προσομοίωσης που περιγράφηκε στο παράδειγμα 5.2.1. Από τη στιγμή που θα έχουμε προσομοιώσει τους χρόνους μεταξύ των αφίξεων, θα ξέρουμε ότι το πρώτο γεγονός της διαδικασίας Poisson συνέβη στο χρόνο T_1, το δεύτερο στο χρόνο $T_1 + T_2$, το τρίτο στο χρόνο $T_1 + T_2 + T_3$ κ.ο.κ.

Η επόμενη πρόταση μας δίνει ουσιαστικά τον τρόπο για την προσομοίωση μιας τυχαίας μεταβλητής της οποίας η κατανομή έχει συνάρτηση ρυθμού βλάβης $\lambda(t)$.

Πρόταση 5.2.3 *Ο χρόνος S στον οποίο θα μετρήσουμε το πρώτο γεγονός είναι μια τυχαία μεταβλητή της οποίας η κατανομή έχει συνάρτηση ρυθμού βλάβης $\lambda(t)$.*

Η μέθοδος ρυθμού βλάβης για την προσομοίωση μιας τυχαίας μεταβλητής S με συνάρτηση ρυθμού βλάβης $\lambda(t)$ συνοψίζεται παρακάτω. Έστω λ_1 ένας αριθμός για τον οποίο $\lambda(t) \leq \lambda_1$ για όλες τις τιμές $t \geq 0$. Δημιουργούμε ζευγάρια U_i, X_i, όπου U_i είναι ομοιόμορφες τυχαίες μεταβλητές στο $(0,1)$ και X_i εκθετικές με παράμετρο λ_1, σταματώντας στο σημείο

$$N = min\left\{n : U_n \leq \frac{\lambda\left(\sum_{i=1}^{n} X_i\right)}{\lambda_1}\right\}.$$

Τότε θα έχουμε ότι $S = \sum_{i=1}^{N} X_i$.

5.2.4 Ο Αλγόριθμος του Von Neumann

Εκτός των τριών μεθόδων που περιγράφηκαν παραπάνω για την προσομοίωση συνεχών τυχαίων μεταβλητών, υπάρχουν κάποιες ειδικές μέθοδοι που χρησιμοποιούνται αποκλειστικά για την προσομοίωση συγκεκριμένων τυχαίων μεταβλητών. Ο αλγόριθμος Von Neumann αποτελεί μια τέτοια ειδική μέθοδο για την προσομοίωση εκθετικών τυχαίων μεταβλητών με παράμετρο 1.

Στο παράδειγμα 5.2.1 είδαμε ότι μια εκθετική τυχαία μεταβλητή με παράμετρο 1 μπορεί να προσομοιωθεί με τον υπολογισμό του λογαρίθμου ενός τυχαίου αριθμού. Τα περισσότερα προγράμματα που υπάρχουν για τον υπολογισμό του λογαρίθμου ενός αριθμού χρησιμοποιούν δυναμοσειρές (π.χ. σειρές Taylor) και επομένως ο υπολογισμός δεν είναι ιδιαίτερα αποδοτικός. Ο αλγόριθμος του Von Neumann δεν χρησιμοποιεί λογαρίθμους, αλλά μόνο τυχαίους αριθμούς στο $(0,1)$ και συνεπώς είναι πολύ πιο αποδοτικός από τη μέθοδο προσομοίωσης που περιγράφηκε στο παράδειγμα 5.2.1.

Τα βήματα του αλγορίθμου για την προσομοίωση μιας εκθετικής τυχαίας μεταβλητής X με παράμετρο 1 είναι τα εξής:

Βήμα 1: Δημιούργησε τυχαίους αριθμούς $U_1, U_2, \ldots$ στο $(0,1)$ μέχρι το σημείο $N = min\{n : U_1 \geq U_2 \ldots \geq U_{n-1} < U_n\}$.
Βήμα 2: Αν ο αριθμός N είναι ζυγός, προχώρησε στο βήμα 3. Διαφορετικά, γύρνα στο βήμα 1.
Βήμα 3: Η τιμή της τυχαίας μεταβλητής X είναι ίση με τον αριθμό των αποτυχημένων επαναλήψεων του αλγορίθμου (δηλαδή πόσες φορές ξαναγυρίσαμε από το βήμα 2 στο βήμα 1) συν τον αριθμό U_1.

Για να αποδείξουμε την ορθότητα του παραπάνω αλγορίθμου, ας θεωρήσουμε μια τυχαία μεταβλητή N τέτοια ώστε

$$N = min\{n : U_1 \geq U_2 \ldots \geq U_{n-1} < U_n\}$$

όπου $U_1, U_2, \ldots$ είναι ομοιόμορφες τυχαίες μεταβλητές στο $(0,1)$. Η από κοινού κατανομή των τυχαίων μεταβλητών N και U_1 είναι η:

$$P\{N > n, U_1 \leq y\} = \int_0^1 P\{N > n, U_1 \leq y | U_1 = x\} f_{U_1}(x)dx$$

$$= \int_0^y P\{N > n, U_1 \leq y | U_1 = x\} dx = \int_0^y P\{N > n | U_1 = x\} dx.$$

Τώρα, δοθέντος ότι $U_1 = x$, η τυχαία μεταβλητή N θα είναι μεγαλύτερη από n αν $x \geq U_2 \ldots \geq U_n$ ή ισοδύναμα αν $U_i \leq x$ για κάθε $i = 2, \ldots, n$ και $U_2 \ldots \geq U_n$. Το γεγονός $U_i \leq x$ έχει πιθανότητα x^{n-1} να συμβεί, ενώ το γεγονός $U_2 \ldots \geq U_n$ πιθανότητα $\frac{1}{(n-1)!}$ αφού το σύνολο όλων των πιθανών διατάξεων $n-1$ αριθμών είναι $(n-1)!$. Επομένως,

$$P\{N > n | U_1 = x\} = \frac{x^{n-1}}{(n-1)!}$$

και

$$P\{N > n, U_1 \leq y\} = \int_0^y \frac{x^{n-1}}{(n-1)!} dx = \frac{y^n}{n!}.$$

Από την παραπάνω εξίσωση προκύπτει ότι:

$$P\{N = n, U_1 \leq y\} = P\{N > n-1, U_1 \leq y\} - P\{N > n, U_1 \leq y\}$$

$$= \frac{y^{n-1}}{(n-1)!} - \frac{y^n}{n!}.$$

Αν θέσουμε $n = 2, 4, \ldots$ έχουμε ότι

$$P\{N \text{ ζυγός}, U_1 \leq y\} = P\{N = 2, U_1 \leq y\} + P\{N = 4, U_1 \leq y\} + \ldots$$

$$= y - \frac{y^2}{2!} + \frac{y^3}{3!} - \frac{y^4}{4!} + \ldots = 1 - e^{-y}.$$

Στην εξίσωση αυτή αν θέσουμε $y = 1$ έχουμε ότι:

$$P\{N \text{ ζυγός}, U_1 \leq 1\} = P\{N \text{ ζυγός}\} = 1 - e^{-1}$$

και

$$P\{N \text{ περιττός}\} = e^{-1}.$$

Για να αποδείξουμε τώρα ότι η τυχαία μεταβλητή X που δημιουργείται με τον αλγόριθμο Von Neumann ακολουθεί την εκθετική κατανομή με παράμετρο 1, πρέπει να δείξουμε ότι $P\{X > x\} = e^{-x}$. Για να είναι προκύψει μία τιμή της τυχαίας μεταβλητής X μεγαλύτερη ενός αριθμού x, θα πρέπει ο αλγόριθμος να αποτύχει (δηλαδή από το βήμα 2 να ξαναγυρίσει στο βήμα 1) $\lfloor x \rfloor$ φορές και την επόμενη είτε να αποτύχει πάλι ή να επιτύχει και η τιμή U_1 να είναι μεγαλύτερη από $x - \lfloor x \rfloor$. Ισχύει ότι:

$$P\{X > x\} = P\{\lfloor x \rfloor \text{ αποτυχίες}\} P\{\text{επόμενη επανάληψη αποτυχία ή επιτυχία με} U_1 > x - \lfloor x \rfloor\}.$$

Η πιθανότητα να έχουμε μία αποτυχία είναι ίση με την πιθανότητα $P\{N \text{ περιττός}\} = e^{-1}$. Άρα,

$$P\{\lfloor x \rfloor \text{ αποτυχίες}\} = e^{-\lfloor x \rfloor}.$$

Επίσης, η πιθανότητα $P\{\text{επόμενη επανάληψη επιτυχία και} U_1 > x - \lfloor x \rfloor\}$ είναι ίση με

$$\begin{aligned} P\{N \text{ ζυγός}, U_1 > x - \lfloor x \rfloor\} &= P\{N \text{ ζυγός}\} - P\{N \text{ ζυγός}, U_1 \leq x - \lfloor x \rfloor\} \\ &= 1 - e^{-1} - (1 - e^{-(x - \lfloor x \rfloor)}) = e^{-(x - \lfloor x \rfloor)} - e^{-1}. \end{aligned}$$

Επομένως, έχουμε ότι $P\{\text{επόμενη επανάληψη αποτυχία ή επιτυχία με} U_1 > x - \lfloor x \rfloor\}$ είναι ίση με:

$$P\{\text{επόμενη επανάληψη αποτυχία}\} + P\{\text{επιτυχία με} U_1 > x - \lfloor x \rfloor\} = e^{-1} + e^{-(x - \lfloor x \rfloor)} - e^{-1}.$$

Με βάση τα παραπάνω, $P\{X > x\} = e^{-\lfloor x \rfloor}(e^{-1} + e^{-(x - \lfloor x \rfloor)} - e^{-1}) = e^{-x}$.

5.3 Προσομοίωση Διακριτών Τυχαίων Μεταβλητών

Η προσομοίωση διακριτών τυχαίων μεταβλητών γίνεται με έναν αρκετά απλό τρόπο. Για παράδειγμα, αν θέλουμε να προσομοιώσουμε μια διακριτή τυχαία μεταβλητή X με συνάρτηση πιθανότητας

$$P\{X = x_j\} = P_j$$

για $j = 1, 2, \ldots$ μπορούμε να χρησιμοποιήσουμε την εξής μέθοδο:

Βήμα 1: Δημιούργησε έναν τυχαίο αριθμό U στο $(0, 1)$.

Βήμα 2: Θέσε

$$X = \begin{cases} x_1 & \text{αν } U < P_1 \\ x_2 & \text{αν } P_1 < U < P_1 + P_2 \\ \cdot & \\ \cdot & \\ \cdot & \\ x_j & \text{αν } \sum_{i=1}^{j-1} P_i < U < \sum_{i=1}^{j} P_i \\ \cdot & \\ \cdot & \end{cases}$$

Με βάση την παραπάνω μέθοδο έχουμε ότι

$$P\{X=x_j\}=P\left\{\sum_{i=1}^{j-1}P_i<U<\sum_{i=1}^{j}P_i\right\}.$$

Αλλά για την ομοιόμορφη τυχαία μεταβλητή U ισχύει ότι

$$P\left\{\sum_{i=1}^{j-1}P_i<U<\sum_{i=1}^{j}P_i\right\}=P_j.$$

Επομένως, διαπιστώνουμε ότι η παραπάνω μέθοδος προσομοιώνει με ορθό τρόπο τη διακριτή τυχαία μεταβλητή X.

Παράδειγμα 5.3.1 *Προσομοιώστε μία διωνυμική τυχαία μεταβλητή X με παραμέτρους (n,p).*

Λύση. Γνωρίζουμε ότι μία διωνυμική τυχαία μεταβλητή με παραμέτρους (n,p) μπορεί να εκφραστεί ως το άθροισμα n ανεξάρτητων τυχαίων μεταβλητών Bernoulli με παράμετρο p. Μια τυχαία μεταβλητή Bernoulli X_i με παράμετρο p μπορεί να προσομοιωθεί με τον εξής τρόπο:

$$X_i=\begin{cases}1 & \text{αν } U_i<p\\ 0 & \text{διαφορετικά}\end{cases}$$

όπου U_i είναι μια ομοιόμορφη τυχαία μεταβλητή στο $(0,1)$. Επομένως, μπορούμε να κατασκευάσουμε n τυχαίες μεταβλητές Bernoulli και να θέσουμε την διωνυμική τυχαία μεταβλητή X ίση με $X=\sum_{i=1}^{n}X_i$. ∎

Παράδειγμα 5.3.2 *Προσομοιώστε μία γεωμετρική τυχαία μεταβλητή X με παράμετρο p.*

Λύση. Η συνάρτηση πιθανότητας μιας γεωμετρικής τυχαίας μεταβλητής X είναι ίση με:

$$P\{X=i\}=p(1-p)^{i-1}$$

με $i\geq 1$. Καθώς,

$$\sum_{i=1}^{j-1}P\{X=i\}=1-P\{X>j-1\}=1-\sum_{i=j}^{\infty}p(1-p)^{i-1}$$

$$=1-\sum_{i=j}^{\infty}\left((1-p)^{i-1}-(1-p)^{i}\right)=1-(1-p)^{j-1}$$

μπορούμε να προσομοιώσουμε την τυχαία μεταβλητή X δημιουργώντας έναν τυχαίο αριθμό U στο $(0,1)$ και να θέσουμε στην μεταβλητή X την τιμή j για την οποία

$$1-(1-p)^{j-1}<U<1-(1-p)^{j}$$

ή ισοδύναμα
$$(1-p)^j < 1-U < (1-p)^{j-1}.$$

Λόγω του ότι και ο αριθμός $1-U$ είναι ένας τυχαίος αριθμός στο $(0,1)$ μπορούμε να ορίσουμε την τυχαία μεταβλητή X ως:

$$X = min\left\{j : (1-p)^j < U\right\} = min\left\{j : j > \frac{\log U}{\log(1-p)}\right\}$$
$$= 1 + \left\lfloor \frac{\log U}{\log(1-p)} \right\rfloor.$$

■

Παράδειγμα 5.3.3 *Προσομοιώστε μία τυχαία μεταβλητή Poisson με παράμετρο λ.*

Λύση. Μια τυχαία μεταβλητή Poisson N με παράμετρο λ μπορεί να προσομοιωθεί μέσω μιας στοχαστικής διαδικασίας Poisson $\{N(t), t \geq 0\}$ με ρυθμό 1. Για κάθε χρονικό διάστημα t, η τυχαία μεταβλητή $N(t)$ της στοχαστικής διαδικασίας θα είναι μια τυχαία μεταβλητή Poisson με παράμετρο t. Ουσιαστικά δηλαδή, η τυχαία μεταβλητή N θα είναι ίση με την τυχαία μεταβλητή $N(\lambda)$ που μας δίνει το πλήθος των γεγονότων της στοχαστικής διαδικασίας που έχουν συμβεί μέχρι τον χρόνο λ. Πως μπορούμε όμως να μετρήσουμε πόσα γεγονότα έχουν συμβεί μέχρι τη χρονική στιγμή λ; Πολύ απλά, μπορούμε να δημιουργήσουμε εκθετικές τυχαίες μεταβλητές X_i με παράμετρο 1 που θα συμβολίζουν τους χρόνους μεταξύ των αφίξεων και να σταματήσουμε στο ελάχιστο σημείο n για το οποίο $\sum_{i=1}^{n} X_i > \lambda$. Τότε η τιμή της τυχαίας μεταβλητής N θα είναι $n-1$ (δηλαδή όσα και τα γεγονότα που συνέβησαν στο χρονικό διάστημα $(0,\lambda)$). Επομένως,

$$N+1 = min\left\{n : \sum_{i=1}^{n} X_i > \lambda\right\}.$$

Αν αντικαταστήσουμε τις τιμές X_i με τις $-\log U_i$ όπου U_i είναι ομοιόμορφες τυχαίες μεταβλητές στο $(0,1)$, τότε έχουμε ότι:

$$N+1 = min\left\{n : \sum_{i=1}^{n} (-\log U_i) > \lambda\right\}$$

ή ισοδύναμα

$$N+1 = min\left\{n : \prod_{i=1}^{n} U_i < e^{-\lambda}\right\}.$$

■

5.4 Προσομοίωση Στοχαστικών Διαδικασιών

Μια στοχαστική διαδικασία $\{X(t), t \in T\}$ μπορεί να προσομοιωθεί με την προσομοίωση των τυχαίων μεταβλητών $X(t)$ που την αποτελούν. Για παράδειγμα, μια διαδικασία Poisson $\{N(t), t \geq 0\}$ μπορεί να προσομοιωθεί με την δημιουργία των τυχαίων μεταβλητών $N(t)$. Σε προηγούμενη ενότητα είδαμε ότι μια διαδικασία Poisson μπορεί να προκύψει πολύ εύκολα με την προσομοίωση εκθετικών τυχαίων μεταβλητών που μας δίνουν τους χρόνους μεταξύ των αφίξεων των γεγονότων. Υπάρχει ωστόσο μια εναλλακτική προσέγγιση για την προσομοίωση στοχαστικών διαδικασιών Poisson αρκετά αποδοτική.

Έστω ότι θέλουμε να προσομοιώσουμε τα γεγονότα μιας στοχαστικής διαδικασίας Poisson με ρυθμό λ που συνέβησαν μέχρι τη χρονική στιγμή t. Δηλαδή να βρούμε πόσα γεγονότα συνέβησαν στο διάστημα $(0, t)$ και σε ποιες χρονικές στιγμές ακριβώς. Για να το κάνουμε αυτό, μπορούμε πρώτα να προσομοιώσουμε την τυχαία μεταβλητή $N(t)$ που μας δίνει το πλήθος των γεγονότων μέχρι τη χρονική στιγμή t και μετά να χρησιμοποιήσουμε την ιδιότητα των διαδικασιών Poisson ότι οι χρόνοι που συνέβησαν τα $N(t)$ γεγονότα κατανέμονται ομοιόμορφα μέσα στο διάστημα $(0, t)$. Άρα, προσομοιώνουμε την τυχαία μεταβλητή Poisson $N(t)$ που έχει παράμετρο λt με τη μέθοδο που περιγράφηκε στο παράδειγμα 5.3.3 και στη συνέχεια, αν βρήκαμε ότι $N(t) = n$, δημιουργούμε ένα σύνολο n τυχαίων μεταβλητών U_i στο $(0, 1)$. Τότε, το σύνολο $(tU_1, tU_2, \ldots, tU_n)$ θα αντιπροσωπεύει τους χρόνους που συνέβησαν τα $N(t)$ γεγονότα. Το πρόβλημα σε αυτή τη μέθοδο είναι ότι συχνά θέλουμε να έχουμε τους χρόνους των γεγονότων σε αύξουσα σειρά για να μπορούμε να υπολογίσουμε κάθε τιμή $N(s)$ με $s \leq t$, αφού

$$N(s) = \text{πλήθος των } U_i \text{ για τα οποία } tU_i \leq s.$$

Άρα, θα πρέπει να ταξινομήσουμε πρώτα τις τιμές U_i και μετά να δημιουργήσουμε το σύνολο $(tU_1, tU_2, \ldots, tU_n)$.

Κεφάλαιο 6

Μαρκοβιανές Αλυσίδες

6.1 Εισαγωγή

Μαρκοβιανή αλυσίδα είναι μια διακριτή στοχαστική διαδικασία της μορφής $\{X_n, n = 0, 1, 2, \ldots\}$. Κάθε τυχαία μεταβλητή X_n αυτής της στοχαστικής διαδικασίας μπορεί να πάρει τις τιμές $\{0, 1, 2, \ldots\}$. Αν $X_n = i$ τότε λέμε ότι η μαρκοβιανή αλυσίδα βρίσκεται στην κατάσταση i τη χρονική στιγμή n. Υποθέτουμε ότι κάθε φορά που η μαρκοβιανή αλυσίδα είναι στην κατάσταση i, υπάρχει μια συγκεκριμένη πιθανότητα P_{ij} να βρεθεί την επόμενη ακριβώς χρονική στιγμή στην κατάσταση j. Δηλαδή, υποθέτουμε ότι:

$$P\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = P_{ij} \tag{6.1}$$

για όλες τις καταστάσεις $i_0, i_1, \ldots, i_{n-1}, i, j$ και όλες τις χρονικές στιγμές $n \geq 0$. Η εξίσωση 6.1 ουσιαστικά δηλώνει ότι για μια μαρκοβιανή αλυσίδα $\{X_n, n = 0, 1, 2, \ldots\}$, η κατάσταση j της τυχαίας μεταβλητής X_{n+1} εξαρτάται μόνο από την κατάσταση i στην οποία βρίσκεται η τυχαία μεταβλητή X_n και όχι από τις καταστάσεις των προηγούμενων $X_{n-1}, X_{n-2}, \ldots$ τυχαίων μεταβλητών. Δηλαδή,

$$P_{ij} = P\{X_{n+1} = j | X_n = i\}.$$

Η τιμή P_{ij} αντιπροσωπεύει όπως είδαμε την πιθανότητα η μαρκοβιανή αλυσίδα να μετακινηθεί από την κατάσταση i στην κατάσταση j και καλείται πιθανότητα μετάβασης. Για όλες τις τιμές P_{ij} ισχύει:

$$P_{ij} \geq 0 \text{ για } i, j \geq 0$$

και

$$\sum_{j=0}^{\infty} P_{ij} = 1 \text{ για } i = 0, 1, \ldots$$

Ο πίνακας μετάβασης πρώτης τάξης (one step transition matrix) $\mathbb{P}$ μιας μαρκοβιανής αλυσίδας είναι ίσος με:

$$\mathbb{P} = \left\| \begin{array}{cccc} P_{00} & P_{01} & P_{02} & \ldots \\ P_{10} & P_{11} & P_{12} & \ldots \\ . & & & \\ . & & & \\ . & & & \\ P_{i0} & P_{i1} & P_{i2} & \ldots \\ . & & & \\ . & & & \end{array} \right\|$$

Σύμφωνα με τη δεύτερη ιδιότητα των πιθανοτήτων μετάβασης (δηλαδή ότι $\sum_{j=0}^{\infty} P_{ij} = 1$) προκύπτει ότι το άθροισμα των στοιχείων κάθε γραμμής του πίνακα $\mathbb{P}$ θα είναι ίσο με 1.

Παράδειγμα 6.1.1 *Υποθέστε ότι η πιθανότητα να βρέξει αύριο εξαρτάται μόνο από τις συνθήκες που επικρατούν σήμερα. Υποθέστε επίσης ότι αν βρέχει σήμερα, τότε θα βρέξει και αύριο με πιθανότητα a και αν δεν βρέχει σήμερα, τότε θα βρέξει αύριο με πιθανότητα b. Πώς μοντελοποιείται το παράδειγμα αυτό με μια μαρκοβιανή αλυσίδα;*

Λύση. Μια μαρκοβιανή αλυσίδα $\{X_n, n = 0, 1, 2, \ldots\}$ όπου κάθε τυχαία μεταβλητή X_n συμβολίζει τον καιρό την n-οστή μέρα και παίρνει την τιμή $X_n = 0$ αν βρέχει και την τιμή $X_n = 1$ αν δεν βρέχει, μπορεί να χρησιμοποιηθεί για την μοντελοποίηση του παραπάνω παραδείγματος. Ο πίνακας μετάβασης πρώτης τάξης της αλυσίδας είναι ίσος με:

$$\mathbb{P} = \left\| \begin{array}{cc} a & 1-a \\ b & 1-b \end{array} \right\|$$

∎

Παράδειγμα 6.1.2 *Έστω ότι μέσω ενός καναλιού επικοινωνίας θέλουμε να στείλουμε ένα bit. Κάθε bit που μεταδίδεται πρέπει να περάσει από διάφορα στάδια και σε κάθε στάδιο υπάρχει πιθανότητα p να αλλάξει η τιμή του. Μοντελοποιείστε το παράδειγμα αυτό με μια μαρκοβιανή αλυσίδα.*

Λύση. Αν θεωρήσουμε ότι X_n είναι η τιμή του bit πριν περάσει από το στάδιο n, τότε η στοχαστική διαδικασία $\{X_n, n = 0, 1, \ldots\}$ είναι μια μαρκοβιανή αλυσίδα δύο καταστάσεων που έχει πίνακα μετάβασης πρώτης τάξης τον:

$$\mathbb{P} = \left\| \begin{array}{cc} 1-p & p \\ p & 1-p \end{array} \right\|$$

∎

Παράδειγμα 6.1.3 *Κάθε μέρα ένας άνθρωπος μπορεί να είναι είτε χαρούμενος (Χ) ή έτσι και έτσι (Ε) ή λυπημένος (Λ). Αν είναι χαρούμενος σήμερα, τότε θα είναι Χ, Ε ή Λ αύριο με πιθανότητες 0.5, 0.4 και 0.1 αντίστοιχα. Αν σήμερα είναι έτσι και έτσι, τότε θα είναι Χ, Ε ή Λ αύριο με πιθανότητες 0.3, 0.4 και 0.3. Τέλος, αν είναι λυπημένος σήμερα, θα είναι Χ, Ε ή Λ αύριο με πιθανότητες 0.2, 0.3 και 0.5. Περιγράψτε το παράδειγμα με μια μαρκοβιανή αλυσίδα.*

Λύση. Έστω X_n η διάθεση του ανθρώπου τη n-οστή μέρα. Τότε, η στοχαστική διαδικασία $\{X_n, n = 0, 1, \ldots\}$ θα είναι μαρκοβιανή διαδικασία με πίνακα μετάβασης πρώτης τάξης τον

$$\mathbb{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix}$$

αν θεωρήσουμε ότι η κατάσταση 0 της αλυσίδας συμβολίζει το γεγονός ο άνθρωπος να είναι χαρούμενος, η κατάσταση 1 να είναι έτσι και έτσι και η κατάσταση 2 να είναι λυπημένος. ∎

Παράδειγμα 6.1.4 *Υποθέστε ότι το γεγονός να βρέξει μια οποιαδήποτε μέρα εξαρτάται μόνο από τις καιρικές συνθήκες των προηγούμενων δύο ημερών. Συγκεκριμένα, αν έβρεχε χτες και σήμερα, τότε θα βρέξει και αύριο με πιθανότητα 0.7, αν έβρεχε σήμερα αλλά όχι χτες, τότε θα βρέξει αύριο με πιθανότητα 0.5, αν δεν έβρεχε σήμερα και έβρεξε χτες, τότε θα βρέξει αύριο με πιθανότητα 0.4 και τέλος, αν δεν έβρεχε ούτε σήμερα ούτε χτες, τότε θα βρέξει αύριο με πιθανότητα 0.2. Μοντελοποιήστε το παράδειγμα αυτό με μια μαρκοβιανή αλυσίδα.*

Λύση. Αν θεωρήσουμε μια στοχαστική διαδικασία $\{X_n, n = 0, 1, \ldots\}$ όπου η τυχαία μεταβλητή X_n να προσδιορίζει τις καιρικές συνθήκες την n-οστή μέρα, τότε προφανώς η διαδικασία αυτή δεν είναι μαρκοβιανή αλυσίδα. Για να μοντελοποιήσουμε το παράδειγμα που δίνεται, θα πρέπει να θεωρήσουμε μια στοχαστική διαδικασία της οποίας οι καταστάσεις θα εξαρτώνται από τις καιρικές συνθήκες σήμερα και χτες. Άρα θα πρέπει κάθε τυχαία μεταβλητή X_n να προσδιορίζει τις καιρικές συνθήκες δύο ημερών: της n-οστής και της $n-1$-οστής μέρας. Συγκεκριμένα, αν ορίσουμε ως καταστάσεις τις:

- κατάσταση 0: αν έβρεχε σήμερα και χτες
- κατάσταση 1: αν έβρεχε σήμερα και όχι χτες
- κατάσταση 2: αν δεν έβρεχε σήμερα αλλά έβρεξε χτες
- κατάσταση 3: αν δεν έβρεχε ούτε σήμερα ούτε χτες

τότε η στοχαστική διαδικασία $\{X_n, n = 0, 1, \ldots\}$ θα είναι μαρκοβιανή αλυσίδα με πίνακα μετάβασης πρώτης τάξης τον:

$$\mathbb{P} = \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix}$$

∎

Μια κατηγορία μαρκοβιανών αλυσίδων που χρησιμοποιείται συχνά για την επίλυση διαφόρων μαθηματικών προβλημάτων (π.χ. στην κρυπτογραφία) είναι οι *τυχαίοι περίπατοι*. Τυχαίος περίπατος είναι μια μαρκοβιανή αλυσίδα της οποίας οι πιθανές καταστάσεις δίνονται από τους ακέραιους $i = 0, \pm 1, \pm 2, \ldots$ και για κάποιον αριθμό $0 < p < 1$ ισχύει:

$$P_{i,i+1} = p = 1 - P_{i,i-1} \text{ για } i = 0, \pm 1, \pm 2, \ldots$$

Αν φανταστούμε τη συγκεκριμένη μαρκοβιανή αλυσίδα ως μοντέλο για έναν άνθρωπο που περπατάει πάνω σε μια ευθεία γραμμή και σε κάθε χρονική στιγμή διαλέγει να κάνει ένα βήμα προς τα δεξιά με πιθανότητα p ή ένα βήμα αριστερά με πιθανότητα $1 - p$, γίνεται κατανοητό γιατί καλείται τυχαίος περίπατος.

Ας δούμε ένα παράδειγμα στο οποίο έχουν εφαρμογή οι τυχαίοι περίπατοι. Θεωρείστε έναν τζογαδόρο ο οποίος σε κάθε γύρο του παιχνιδιού που παίζει είτε κερδίζει 1 ευρώ με πιθανότητα p ή χάνει 1 ευρώ με πιθανότητα $1 - p$. Αν υποθέσουμε ότι ο τζογαδόρος συνεχίζει να παίζει μέχρι είτε να χάσει όλα του τα χρήματα ή όταν αποκτήσει N ευρώ, τότε το ποσό που έχει ο τζογαδόρος ανά πάσα χρονική στιγμή αποτελεί μια μαρκοβιανή αλυσίδα. Κάθε τιμή X_n της μαρκοβιανής αλυσίδας $\{X_n, n = 0, 1, \ldots\}$ δηλώνει το ποσό που έχει ο τζογαδόρος μετά τον n-οστό γύρο. Δηλαδή, κάθε $X_n \in \{0, 1, 2, \ldots, N\}$ και οι πιθανότητες μετάβασης δίνονται από τις σχέσεις

$$P_{i,i+1} = p = 1 - P_{i,i-1} \text{ για } i = 1, 2, \ldots, N - 1$$

και

$$P_{00} = P_{NN} = 1.$$

Οι καταστάσεις 0 και N καλούνται *καταστάσεις απορρόφησης* λόγω του ότι αν φτάσει κανείς σε αυτές δεν μπορεί να φύγει. Προφανώς η μαρκοβιανή αλυσίδα του παραπάνω παραδείγματος αποτελεί έναν τυχαίο περίπατο.

Παράδειγμα 6.1.5 *Έστω $\{X_n, n = 0, 1, \ldots\}$ μια ακολουθία ανεξάρτητων, διακριτών, τυχαίων μεταβλητών για τις οποίες $P\{X_n = -1\} = P\{X_n = 1\} = 0.5$. Δείξτε ότι η ακολουθία $\{Y_n, n = 0, 1, \ldots\}$ που ορίζεται ως:*

$$Y_n = \frac{1}{2}\left(X_n + X_{n+1}\right)$$

δεν είναι μαρκοβιανή αλυσίδα.

Λύση. Για να αποδείξουμε ότι η στοχαστική διαδικασία $\{Y_n, n = 0, 1, \ldots\}$ δεν είναι μαρκοβιανή αλυσίδα αρκεί να βρούμε τιμές y_1, y_2, y_3 για τις οποίες

$$P\{Y_{n+1} = y_1 | Y_n = y_2\} \neq P\{Y_{n+1} = y_1 | Y_n = y_2, Y_{n-1} = y_3\}.$$

Οι τιμές που μπορούν να πάρουν οι τυχαίες μεταβλητές Y_n είναι οι -1, 0 και 1. Ας δούμε τί συμβαίνει στην περίπτωση που επιλέγουμε $y_1 = 1$, $y_2 = 0$ και $y_3 = 1$:

$$P\{Y_{n+1} = 1 | Y_n = 0\} = \frac{P\{Y_{n+1} = 1, Y_n = 0\}}{P\{Y_n = 0\}}$$

$$= \frac{P\{X_{n+1} + X_{n+2} = 2, X_n + X_{n+1} = 0\}}{P\{X_n + X_{n+1} = 0\}} = \frac{P\{X_n = -1, X_{n+1} = 1, X_{n+2} = 1\}}{P\{X_n = -1, X_{n+1} = 1\} + P\{X_n = 1, X_{n+1} = -1\}}$$

$$= \frac{P\{X_n = -1\} P\{X_{n+1} = 1\} P\{X_{n+2} = 1\}}{P\{X_n = -1\} P\{X_{n+1} = 1\} + P\{X_n = 1\} P\{X_{n+1} = -1\}} = \frac{(0.5)^3}{2(0.5)^2} = 0.25.$$

Όμως, η πιθανότητα,

$$P\{Y_{n+1} = 1 | Y_n = 0, Y_{n-1} = 1\} = \frac{P\{X_{n+1} + X_{n+2} = 2, X_n + X_{n+1} = 0, X_{n-1} + X_n = 2\}}{P\{Y_n = 0, Y_{n-1} = 1\}} = 0.$$

Επομένως, η στοχαστική διαδικασία $\{Y_n, n = 0, 1, \ldots\}$ δεν είναι μαρκοβιανή αλυσίδα. ∎

Παράδειγμα 6.1.6 *Σε ένα ταχυδρομείο που διαθέτει ένα μόνο ταμείο οι πελάτες που φτάνουν σχηματίζουν ουρά και περιμένουν τη σειρά τους για να εξυπηρετηθούν. Στο χρονικό διάστημα που το ταμείο εξυπηρετεί τον n-οστό πελάτη, έστω X_n ο αριθμός των πελατών που φτάνουν στο ταχυδρομείο και μπαίνουν στην ουρά. Υποθέτουμε ότι οι $X_0, X_1, \ldots$ είναι ανεξάρτητες τυχαίες μεταβλητές με $P\{X_n = i\} = p_i$ για $i, n = 0, 1, \ldots$. Έστω Y_n ο αριθμός των πελατών που περιμένουν στην ουρά τη στιγμή που το ταμείο έχει τελειώσει με την εξυπηρέτηση του n-οστού πελάτη. Δείξτε ότι η ακολουθία των τυχαίων μεταβλητών Y_n είναι μια μαρκοβιανή αλυσίδα και υπολογίστε τις πιθανότητες μετάβασης πρώτης τάξης.*

Λύση. Αν $Y_n = 0$ τότε $Y_{n+1} = X_{n+1}$. Διαφορετικά, αν $Y_n \neq 0$ τότε $Y_{n+1} = X_{n+1} + Y_n - 1$. Επειδή η τυχαία μεταβλητή X_{n+1} είναι ανεξάρτητη των μεταβλητών $Y_1, Y_2, \ldots, Y_n$, προκύπτει ότι δεν χρειάζεται να γνωρίζουμε τις τιμές των τυχαίων μεταβλητών $Y_1, Y_2, \ldots, Y_{n-1}$ για τον υπολογισμό της δεσμευμένης συνάρτησης πιθανότητας της τυχαίας μεταβλητής Y_{n+1} (βλέπε εξίσωση 6.1). Άρα η ακολουθία $\{Y_n, n = 0, 1, \ldots\}$ είναι μαρκοβιανή αλυσίδα με πιθανότητες μετάβασης πρώτης τάξης τις:

$$P_{0j} = P\{Y_{n+1} = j | Y_n = 0\} = P\{X_{n+1} = j\} = p_j$$

για $j = 0, 1, 2 \ldots$ και για $i = 1, 2 \ldots$

$$P_{ij} = P\{Y_{n+1} = j | Y_n = i\} = P\{X_{n+1} = j - i + 1\} = p_{j-i+1}$$

για $j \geq i - 1$ και

$$P_{ij} = 0$$

για $j < i - 1$. ∎

6.2 Εξισώσεις Chapman-Kolmogorov

Εκτός από τις πιθανότητες μετάβασης πρώτης τάξης P_{ij}, μπορούν να οριστούν και οι πιθανότητες μετάβασης n-οστής τάξης $P_{ij}^{(n)}$. Κάθε πιθανότητα $P_{ij}^{(n)}$ δηλώνει την πιθανότητα η μαρκοβιανή αλυσίδα να μεταβεί από την κατάσταση i στην κατάσταση j μετά από n βήματα. Δηλαδή,

$$P_{ij}^{(n)} = P\left\{X_{n+k} = j | X_k = i\right\}$$

για $n, i, j \geq 0$. Προφανώς ισχύει ότι $P_{ij}^{(1)} = P_{ij}$. Παρατηρήστε επίσης ότι οι πιθανότητες $P_{ij}^{(n)}$ ορίζονται και για τιμές $n = 0$. Στην περίπτωση αυτή ισχύει ότι

$$P_{ij}^{(0)} = \begin{cases} 1 & \text{αν } i = j \\ 0 & \text{αν } i \neq j \end{cases}$$

για όλα τα $i, j \geq 0$.

Οι εξισώσεις Chapman-Kolmogorov παρέχουν μία μέθοδο για τον υπολογισμό των πιθανοτήτων μετάβασης n-οστής τάξης. Συγκεκριμένα, οι εξισώσεις αυτές είναι οι:

$$P_{ij}^{(n+m)} = \sum_{k=0}^{\infty} P_{ik}^{(n)} P_{kj}^{(m)} \tag{6.2}$$

για όλες τις τιμές $n, m, i, j \geq 0$. Η πιθανότητα $P_{ik}^{(n)} P_{kj}^{(m)}$ αντιπροσωπεύει την πιθανότητα να πάμε από την κατάσταση i στην κατάσταση j μετά από $n + m$ βήματα μέσω ενός μονοπατιού που περνάει από την κατάσταση k μετά από n βήματα. Επομένως, αθροίζοντας τις πιθανότητες $P_{ik}^{(n)} P_{kj}^{(m)}$ για όλες τις πιθανές καταστάσεις k παίρνουμε την τελική πιθανότητα $P_{ij}^{(n+m)}$.

Αν με $\mathbb{P}^{(n)}$ συμβολίσουμε τον πίνακα μετάβασης n-οστής τάξης, δηλαδή

$$\mathbb{P}^{(n)} = \left\| \begin{array}{cccc} P_{00}^{(n)} & P_{01}^{(n)} & P_{02}^{(n)} & \cdots \\ P_{10}^{(n)} & P_{11}^{(n)} & P_{12}^{(n)} & \cdots \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ P_{i0}^{(n)} & P_{i1}^{(n)} & P_{i2}^{(n)} & \cdots \\ \cdot & & & \\ \cdot & & & \end{array} \right\|$$

τότε από την εξίσωση 6.2 προκύπτει εύκολα ότι

$$\mathbb{P}^{(n+m)} = \mathbb{P}^{(n)} \bullet \mathbb{P}^{(m)}.$$

Η πράξη $\bullet$ αντιπροσωπεύει πολλαπλασιασμό πινάκων. Συγκεκριμένα, ισχύει ότι

$$\mathbb{P}^{(2)} = \mathbb{P}^{(1)} \bullet \mathbb{P}^{(1)} = \mathbb{P} \bullet \mathbb{P} = \mathbb{P}^2$$

και με επαγωγή έχουμε τελικά ότι

$$\mathbb{P}^{(n)} = \mathbb{P}^n.$$

Δηλαδή ο πίνακας μετάβασης n-οστής τάξης προκύπτει από τον πολλαπλασιασμό του πίνακα μετάβασης πρώτης τάξης $\mathbb{P}$ με τον εαυτό του n φορές.

Παράδειγμα 6.2.1 *Θεωρείστε ξανά το παράδειγμα 6.1.1 όπου ο καιρός μιας μέρας μοντελοποιείται με μια μαρκοβιανή αλυσίδα δύο καταστάσεων. Αν* $a = 0.7$ *και* $b = 0.4$, *υπολογίστε την πιθανότητα να βρέξει τέσσερις μέρες από σήμερα δοθέντος ότι σήμερα βρέχει.*

Λύση. Η πιθανότητα που ζητείται να υπολογίσουμε είναι η $P_{00}^{(4)}$. Για να τη βρούμε πρέπει να υπολογίσουμε τον πίνακα μετάβασης 4ης τάξης. Έχουμε ότι

$$\mathbb{P} = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

άρα

$$\mathbb{P}^2 = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} \bullet \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} = \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix}$$

και

$$\mathbb{P}^4 = \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix} \bullet \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix} = \begin{Vmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{Vmatrix}.$$

Επομένως η ζητούμενη πιθανότητα είναι ίση με 0.5749. ∎

Παράδειγμα 6.2.2 *Θεωρείστε ξανά το παράδειγμα 6.1.4. Δοθέντος ότι έβρεξε τη Δευτέρα και την Τρίτη, ποια η πιθανότητα να βρέξει την Πέμπτη;*

Λύση. Επειδή η κατάσταση 0 δηλώνει ότι σήμερα και χτες έβρεξε και η κατάσταση 1 ότι σήμερα έβρεξε αλλά όχι χτες, τότε η πιθανότητα να βρέξει την Πέμπτη θα είναι ίση με την πιθανότητα η μαρκοβιανή αλυσίδα να είναι είτε στην κατάσταση 0 ή στην κατάσταση 1 την Πέμπτη. Επομένως, αφού μας δίνεται η κατάσταση της αλυσίδας την Τρίτη (είναι στην κατάσταση 0) τότε η ζητούμενη πιθανότητα είναι η $P_{00}^{(2)} + P_{01}^{(2)}$. Ο πίνακας μετάβασης πρώτης τάξης είναι ο

$$\mathbb{P} = \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix}$$

και ο $\mathbb{P}^2 = \mathbb{P} \bullet \mathbb{P}$ είναι ίσος με:

$$\mathbb{P}^2 = \begin{Vmatrix} 0.49 & 0.12 & 0.21 & 0.18 \\ 0.35 & 0.20 & 0.15 & 0.30 \\ 0.20 & 0.12 & 0.20 & 0.48 \\ 0.10 & 0.16 & 0.10 & 0.64 \end{Vmatrix}.$$

Άρα $P_{00}^{(2)} + P_{01}^{(2)} = 0.49 + 0.12 = 0.61.$ ∎

Μέχρι στιγμής, όλες οι πιθανότητες που υπολογίσαμε είναι δεσμευμένες πιθανότητες. Για παράδειγμα, κάθε $P_{ij}^{(n)}$ είναι η πιθανότητα η κατάσταση μιας μαρκοβιανής αλυσίδας τη χρονική στιγμή n να είναι η j δοθέντος ότι τη χρονική στιγμή 0 η κατάσταση ήταν η i. Αν μας ζητείται η μη δεσμευμένη πιθανότητα μια χρονική στιγμή n η μαρκοβιανή αλυσίδα να είναι στην κατάσταση j, δηλαδή η $P\{X_n = j\}$, μπορούμε να την υπολογίσουμε μέσω της κατανομής της τυχαίας μεταβλητής X_0. Δηλαδή, έστω ότι η συνάρτηση πιθανότητας της τυχαίας μεταβλητής X_0 δίνεται από τη σχέση

$$a_i = P\{X_0 = i\}$$

με $i \geq 0$ και $\sum_{i=0}^{\infty} a_i = 1$. Τότε

$$P\{X_n = j\} = \sum_{i=0}^{\infty} P\{X_n = j | X_0 = i\} P\{X_0 = i\} = \sum_{i=0}^{\infty} P_{ij}^{(n)} a_i.$$

Για παράδειγμα, αν υποθέσουμε ότι $a_0 = 0.4$ και $a_1 = 0.6$ για την άσκηση 6.2.1 και θεωρήσουμε ότι ο χρόνος μετράει από σήμερα (δηλαδή η τυχαία μεταβλητή X_0 προσδιορίζει τον καιρό σήμερα), τότε η πιθανότητα ότι θα βρέξει σε 4 μέρες από σήμερα θα είναι ίση με

$$P\{X_4 = 0\} = 0.4P_{00}^{(4)} + 0.6P_{10}^{(4)} = 0.4 \cdot 0.5749 + 0.6 \cdot 0.5668 = 0.57.$$

Παράδειγμα 6.2.3 *Ένας συνταξιούχος λαμβάνει 200 ευρώ στην αρχή κάθε μήνα. Το ποσό των χρημάτων που χρειάζεται να ξοδέψει στη διάρκεια κάθε μήνα είναι ανεξάρτητο του ποσού που έχει και είναι ίσο με* $100 \cdot i$ *ευρώ με πιθανότητα* P_i*, με* $i = 1, 2, 3, 4$ *και* $\sum_{i=1}^{4} P_i = 1$*. Αν ο συνταξιούχος έχει παραπάνω από 300 ευρώ στο τέλος του μήνα δίνει το ποσό που είναι πάνω από τα 300 ευρώ στο γιο του. Αν στο τέλος ενός μήνα ο συνταξιούχος έχει 300 ευρώ, ποια η πιθανότητα το ποσό που θα έχει μετά από 4 μήνες να είναι 100 ευρώ ή λιγότερα;*

Λύση. Αρχικά θα πρέπει να μοντελοποιήσουμε το πρόβλημα αυτό με μία μαρκοβιανή αλυσίδα. Αν θεωρήσουμε ως κατάσταση της αλυσίδας το ποσό που έχει ο συνταξιούχος στο τέλος του μήνα, θα έχουμε τις εξής καταστάσεις:

- κατάσταση 0: έχει 100 ευρώ ή λιγότερα
- κατάσταση 1: έχει 200 ευρώ
- κατάσταση 2: έχει 300 ευρώ

Ο πίνακας μετάβασης πρώτης τάξης αυτής της μαρκοβιανής αλυσίδας θα είναι ίσος με:

$$\mathbb{P} = \left\| \begin{array}{ccc} 1 & 0 & 0 \\ P_3 + P_4 & P_2 & P_1 \\ P_4 & P_3 & P_1 + P_2 \end{array} \right\|.$$

Για να γίνει κατανοητό το πώς προέκυψε ο πίνακας αυτός, θα υπολογίσουμε αναλυτικά ένα στοιχείο του, το P_{10}. Με όμοιο τρόπο γίνεται και ο υπολογισμός των υπολοίπων στοιχείων του πίνακα. Η πιθανότητα P_{10} είναι η πιθανότητα ένας μήνας που τελειώνει με τον συνταξιούχο να έχει 200 ευρώ, να καταλήγει στο τέλος του επόμενου μήνα να έχει 100 ευρώ ή λιγότερα. Επειδή στην αρχή του μήνα θα έχει $200 + 200 = 400$ ευρώ, θα πρέπει να ξοδέψει 300 ή 400 ευρώ. Άρα $P_{10} = P_3 + P_4$. Η πιθανότητα που μας ζητάει η άσκηση είναι η $P_{20}^{(4)}$. Αν υποθέσουμε ότι $P_1 = P_2 = P_3 = P_4 = 0.25$ τότε ο πίνακας $\mathbb{P}^4$ είναι ίσος με:

$$\mathbb{P}^4 = \begin{Vmatrix} 1 & 0 & 0 \\ \frac{222}{256} & \frac{13}{256} & \frac{21}{256} \\ \frac{201}{256} & \frac{21}{256} & \frac{34}{256} \end{Vmatrix}$$

και επομένως η πιθανότητα $P_{20}^{(4)} = \frac{201}{256}$.

∎

6.3 Κατηγοριοποίηση Καταστάσεων

Μια κατάσταση j λέμε ότι είναι *προσβάσιμη* από την κατάσταση i και θα συμβολίζουμε $i \to j$, αν $P_{ij}^{(n)} > 0$ για κάποιο $n \geq 0$. Δηλαδή υπάρχει μη μηδενική πιθανότητα να φτάσουμε στην κατάσταση j μετά από κάποια βήματα, ξεκινώντας από την κατάσταση i. Δύο καταστάσεις i και j που είναι προσβάσιμες η μία από την άλλη, δηλαδή $i \to j$ και $j \to i$, λέμε ότι *επικοινωνούν* και συμβολίζουμε το γεγονός αυτό με $i \leftrightarrow j$. Ισχύουν οι εξής ιδιότητες:

- Κάθε κατάσταση i επικοινωνεί με τον εαυτό της i, λόγω του ότι $P_{ii}^{(0)} = P\{X_0 = i | X_0 = i\} = 1$.
- Αν η κατάσταση i επικοινωνεί με την κατάσταση j, τότε και η j επικοινωνεί με την i.
- Αν η κατάσταση i επικοινωνεί με την κατάσταση j και η j με την k, τότε και η i επικοινωνεί με την k.

Δύο καταστάσεις που επικοινωνούν λέμε ότι ανήκουν στην ίδια ομάδα ή *κλάση ισοδυναμίας*. Από τις παραπάνω ιδιότητες προκύπτει ότι οποιεσδήποτε δύο καταστάσεις είτε θα ανήκουν στην ίδια κλάση ισοδυναμίας αν επικοινωνούν μεταξύ τους ή σε διαφορετικές αν δεν επικοινωνούν. Αυτό οδηγεί στο συμπέρασμα ότι οι καταστάσεις μιας μαρκοβιανής αλυσίδας χωρίζονται σε κλάσεις ισοδυναμίας που δεν έχουν μεταξύ τους κοινά στοιχεία. Μια μαρκοβιανή αλυσίδα καλείται *ανάγωγη* αν αποτελείται από μία μόνο κλάση επικοινωνίας, δηλαδή όλες οι καταστάσεις επικοινωνούν μεταξύ τους. Για παράδειγμα, η μαρκοβιανή αλυσίδα τριών καταστάσεων με πίνακα πιθανοτήτων μετάβασης

$$\mathbb{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{Vmatrix}$$

είναι ανάγωγη.

Παράδειγμα 6.3.1 *Έστω μια μαρκοβιανή αλυσίδα τεσσάρων καταστάσεων $\{0,1,2,3\}$ με πίνακα πιθανοτήτων μετάβασης τον*

$$\mathbb{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

Ποιες είναι οι κλάσεις ισοδυναμίας της αλυσίδας;

Λύση. Σχεδιάζοντας το γράφημα της μαρκοβιανής αλυσίδας, προκύπτει εύκολα ότι οι κλάσεις ισοδυναμίας είναι οι $\{0,1\}$, $\{2\}$ και $\{3\}$. ∎

Μια κλάση ισοδυναμίας K καλείται *κλειστή ή κλειστό σύνολο καταστάσεων* αν $P_{ij}^{(n)} = 0$ για κάθε $i \in K$, $j \notin K$ και $n \geq 1$. Αυτό σημαίνει ότι δεν υπάρχει μονοπάτι που να οδηγεί από μια κατάσταση του κλειστού συνόλου σε κάποια κατάσταση εκτός συνόλου. Προφανώς, αν το σύνολο K αποτελείται από μια κατάσταση μόνο, τότε αυτή είναι κατάσταση απορρόφησης.

Για κάθε κατάσταση i, έστω η πιθανότητα f_i να ξαναβρεθούμε στην κατάσταση i, αν υποθέσουμε ότι ξεκινήσαμε από αυτή, μετά από κάποια βήματα. Αν $f_i = 1$ τότε η κατάσταση καλείται *έμμονη* και αν $f_i < 1$ *μεταβατική*. Ο παραπάνω ορισμός δηλώνει ότι αν μια κατάσταση i είναι έμμονη, τότε η μαρκοβιανή αλυσίδα θα βρεθεί στην κατάσταση αυτή άπειρες φορές. Στην περίπτωση που η κατάσταση i είναι μεταβατική, τότε η πιθανότητα να βρεθούμε για $n \geq 1$ ακριβώς φορές στην κατάσταση i είναι $f_i^{n-1}(1-f_i)$. Παρατηρήστε ότι καθώς η τιμή του n αυξάνει, μειώνεται η πιθανότητα. Δηλαδή, η μαρκοβιανή αλυσίδα επισκέπτεται πεπερασμένο πλήθος φορών μια μεταβατική κατάσταση και από ένα σημείο και μετά δεν ξαναπερνάει από την κατάσταση αυτή. Αυτό οδηγεί στο συμπέρασμα ότι δεν μπορεί όλες οι καταστάσεις μιας μαρκοβιανής αλυσίδας να είναι μεταβατικές.

Αν ορίσουμε τυχαίες μεταβλητές I_n με $n \geq 0$ ως:

$$I_n = \begin{cases} 1 & \text{αν } X_n = i \\ 0 & \text{αν } X_n \neq i \end{cases}$$

τότε το άθροισμα $\sum_{n=0}^{\infty} I_n$ αντιπροσωπεύει το πλήθος των φορών που η μαρκοβιανή αλυσίδα είναι στην κατάσταση i. Επίσης, ισχύει ότι:

$$E\left[\sum_{n=0}^{\infty} I_n | X_0 = i\right] = \sum_{n=0}^{\infty} E\left[I_n | X_0 = i\right] = \sum_{n=0}^{\infty} P\left\{X_n = i | X_0 = i\right\} = \sum_{n=0}^{\infty} P_{ii}^n.$$

Πρόταση 6.3.1 *Μια κατάσταση i είναι έμμονη αν $\sum_{n=0}^{\infty} P_{ii}^n = \infty$ και μεταβατική αν $\sum_{n=0}^{\infty} P_{ii}^n < \infty$.*

Πόρισμα 6.3.1 *Αν η κατάσταση i είναι έμμονη και επικοινωνεί με την κατάσταση j, τότε και η j είναι έμμονη.*

Απόδειξη. Λόγω του ότι οι καταστάσεις i και j επικοινωνούν, θα υπάρχουν ακέραιοι k και m, τέτοιοι ώστε $P_{ij}^{(k)} > 0$ και $P_{ji}^{(m)} > 0$. Τώρα, για κάθε ακέραιο n θα ισχύει ότι:

$$P_{jj}^{(m+n+k)} \geq P_{ji}^{(m)} P_{ii}^{(n)} P_{ij}^{(k)}.$$

Ο λόγος που ισχύει η παραπάνω σχέση είναι ότι η πρώτη πιθανότητα $P_{jj}^{(m+n+k)}$ μας δίνει την πιθανότητα να πάμε από την κατάσταση j πίσω πάλι στην j από οποιοδήποτε μονοπάτι μήκους $m+n+k$, ενώ η πιθανότητα $P_{ji}^{(m)} P_{ii}^{(n)} P_{ij}^{(k)}$ δίνει την πιθανότητα να ξεκινήσουμε από την j και να καταλήξουμε πάλι σε αυτή μέσα από ένα πιο συγκεκριμένο μονοπάτι το οποίο οδηγεί στην κατάσταση i μετά από m βήματα, από την i ξαναγυρνά στην i μετά από n βήματα και καταλήγει στην j μετά από k βήματα. Επομένως,

$$\sum_{n=0}^{\infty} P_{jj}^{(m+n+k)} \geq P_{ji}^{(m)} P_{ij}^{(k)} \sum_{n=0}^{\infty} P_{ii}^{(n)} = \infty.$$

Άρα και η κατάσταση j είναι έμμονη. ∎

Από το παραπάνω πόρισμα προκύπτει άμεσα ότι αν μια κατάσταση i είναι μεταβατική και επικοινωνεί με μια κατάσταση j, τότε και η j είναι μεταβατική. Επίσης, αν μια πεπερασμένη μαρκοβιανή αλυσίδα είναι ανάγωγη, τότε όλες οι καταστάσεις της είναι έμμονες.

Πόρισμα 6.3.2 *Αν η κατάσταση i είναι έμμονη και η j μεταβατική, τότε η j δεν μπορεί να είναι προσβάσιμη από την i.*

Δηλαδή μια ομάδα έμμονων καταστάσεων αποτελεί κλειστό σύνολο γιατί αν η διαδικασία βρεθεί σε κάποια από τις καταστάσεις της δεν μπορεί να ξαναφύγει από την ομάδα. Αντίθετα, μία κλάση ισοδυναμίας που αποτελείται από μεταβατικές καταστάσεις δεν μπορεί να είναι κλειστό σύνολο αφού από κάθε μεταβατική κατάσταση υπάρχει κάποιο μονοπάτι που οδηγεί σε μία τουλάχιστον έμμονη κατάσταση.

Παράδειγμα 6.3.2 *Έστω μια μαρκοβιανή αλυσίδα με καταστάσεις τις $\{0,1,2,3,4\}$ και πίνακα πιθανοτήτων μετάβασης πρώτης τάξης τον:*

$$\mathbb{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \end{Vmatrix}$$

Ποιες είναι οι κλάσεις ισοδυναμίας της αλυσίδας και ποιες οι έμμονες και μεταβατικές καταστάσεις;

Λύση. Από το γράφημα της αλυσίδας προκύπτει ότι οι κλάσεις ισοδυναμίας είναι οι $\{0,1\}$, $\{2,3\}$ και $\{4\}$. Οι δύο πρώτες είναι έμμονες και η τρίτη μεταβατική. ∎

Με βάση τα όσα ορίσαμε παραπάνω, αν γνωρίζουμε ποιες καταστάσεις είναι μεταβατικές και ποιες έμμονες, μπορούμε να ορίσουμε τον πίνακα πιθανοτήτων μετάβασης πρώτης τάξης για κάθε μαρκοβιανή αλυσίδα ως εξής:

$$\mathbb{P} = \begin{Vmatrix} \mathbb{P}_T & \mathbb{M} \\ 0 & \mathbb{P}_R \end{Vmatrix}.$$

Οι πίνακες $\mathbb{P}_T$ και $\mathbb{P}_R$ είναι τετραγωνικοί πίνακες που αντιστοιχούν στις πιθανότητες μετάβασης των μεταβατικών και των έμμονων καταστάσεων αντίστοιχα, 0 είναι ο μηδενικός πίνακας και $\mathbb{M}$ είναι ο πίνακας που περιέχει τις πιθανότητες μετάβασης από τις μεταβατικές καταστάσεις στις έμμονες. Για να γίνει πιο κατανοητός ο μετασχηματισμός του πίνακα μετάβασης σε αυτή τη μορφή, θεωρείστε μια μαρκοβιανή αλυσίδα με καταστάσεις τις $\{0,1,2,3,4\}$ και πίνακα μετάβασης τον

$$\mathbb{P} = \begin{Vmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Οι έμμονες καταστάσεις είναι οι $\{0\}$ και $\{4\}$, ενώ οι μεταβατικές οι $\{1,2,3\}$. Ο πίνακας $\mathbb{P}_T$ θα είναι ίσος με

$$\mathbb{P}_T = \begin{Vmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{Vmatrix}$$

ο $\mathbb{P}_R$ θα είναι ο

$$\mathbb{P}_R = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$$

και ο $\mathbb{M}$ θα είναι ίσος με

$$\mathbb{M} = \begin{Vmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{Vmatrix}.$$

Άρα ο πίνακας μετάβασης στη μορφή

$$\mathbb{P} = \begin{Vmatrix} \mathbb{P}_T & \mathbb{M} \\ 0 & \mathbb{P}_R \end{Vmatrix}$$

είναι ο

$$\mathbb{P} = \begin{Vmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

6.4 Οριακές Πιθανότητες

Στο παράδειγμα 6.2.1 είχαμε υπολογίσει τον πίνακα πιθανοτήτων μετάβασης τέταρτης τάξης για μια μαρκοβιανή αλυσίδα δύο καταστάσεων ο οποίος ήταν ίσος με:

$$\mathbb{P}^{(4)} = \left\| \begin{array}{cc} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{array} \right\|$$

Αν υπολογίσουμε τον πίνακα $\mathbb{P}^{(8)} = \mathbb{P}^{(4)}\mathbb{P}^{(4)}$ βρίσκουμε ότι είναι ίσος με:

$$\mathbb{P}^{(8)} = \left\| \begin{array}{cc} 0.572 & 0.428 \\ 0.570 & 0.430 \end{array} \right\|.$$

Παρατηρείστε ότι οι πίνακες $\mathbb{P}^{(8)}$ και $\mathbb{P}^{(4)}$ είναι περίπου ίδιοι και επίσης ότι οι γραμμές του $\mathbb{P}^{(8)}$ έχουν περίπου τις ίδιες τιμές. Συγκεκριμένα, φαίνεται ότι κάθε τιμή $P_{ij}^{(n)}$ συγκλίνει σε μια τιμή όταν $n \to \infty$ η οποία είναι ίδια για όλες τις πιθανές καταστάσεις i. Με άλλα λόγια φαίνεται ότι υπάρχει μία οριακή πιθανότητα η μαρκοβιανή αλυσίδα να βρεθεί σε μια κατάσταση j μετά από μεγάλο αριθμό βημάτων, και ότι η πιθανότητα αυτή δεν εξαρτάται από την αρχική κατάσταση της αλυσίδας.

Για να ορίσουμε με μαθηματικές σχέσεις την παραπάνω παρατήρηση, χρειάζεται να προσδιορίσουμε δύο ιδιότητες των καταστάσεων μιας μαρκοβιανής αλυσίδας. Μια κατάσταση i λέμε ότι έχει περίοδο d αν $P_{ii}^{(n)} = 0$ όποτε το n δεν είναι πολλαπλάσιο του d, και d είναι ο μικρότερος ακέραιος με αυτή την ιδιότητα. Για παράδειγμα, αρχίζοντας από την κατάσταση i, μπορεί να είναι πιθανό να μπορούμε να ξαναπάμε στην i μετά από $2, 4, 6, 8, \ldots$ βήματα. Τότε η κατάσταση i θα έχει περίοδο 2. Μια κατάσταση με περίοδο 1 λέμε ότι είναι *απεριοδική*. Μπορεί να δειχθεί ότι η περιοδικότητα είναι μια ιδιότητα που χαρακτηρίζει μια ολόκληρη κλάση ισοδυναμίας. Δηλαδή, αν μια κατάσταση i έχει περίοδο d και επικοινωνεί με την j, τότε και η j θα έχει περίοδο d.

Αν μια κατάσταση i είναι έμμονη, τότε λέμε ότι είναι *θετικά έμμονη* αν ξεκινώντας από την i, ο αναμενόμενος χρόνος μέχρι η διαδικασία να επιστρέψει στην κατάσταση i είναι πεπερασμένος. Επιπλέον, αν μια θετικά έμμονη κατάσταση ανήκει σε μια κλάση ισοδυναμίας, τότε όλες οι καταστάσεις που ανήκουν σε αυτή την κλάση είναι θετικά έμμονες. Σε μια μαρκοβιανή αλυσίδα με πεπερασμένο αριθμό καταστάσεων όλες οι έμμονες καταστάσεις είναι θετικά έμμονες. Οι θετικά έμμονες, απεριοδικές καταστάσεις καλούνται *εργοδικές*.

Μπορούμε τώρα να δώσουμε το παρακάτω θεώρημα:

Θεώρημα 6.4.1 *Για μια ανάγωγη, εργοδική μαρκοβιανή αλυσίδα υπάρχει το όριο* $\lim_{n\to\infty} P_{ij}^{(n)}$ *και είναι ανεξάρτητο της κατάστασης* i*. Επιπλέον, αν* $\pi_j = \lim_{n\to\infty} P_{ij}^{(n)}$ *για* $j \geq 0$ *τότε τα* π_j *αποτελούν την μοναδική λύση των εξισώσεων:*

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}$$

και

$$\sum_{j=0}^{\infty} \pi_j = 1.$$

Απόδειξη. Οι τιμές π_j είναι ίσες με $\pi_j = \lim_{n\to\infty} P_{ij}^{(n)} = \lim_{n\to\infty} P\{X_n = j\}$. Υπολογίζοντας ότι

$$P\{X_{n+1} = j\} = \sum_{i=0}^{\infty} P\{X_{n+1} = j | X_n = i\} P\{X_n = i\} = \sum_{i=0}^{\infty} P_{ij} P\{X_n = i\}$$

προκύπτει ότι για $n \to \infty$ έχουμε $\pi_j = \sum_{i=0}^{\infty} P_{ij}\pi_i$. Επίσης, αφού π_j είναι η πιθανότητα η αλυσίδα να βρεθεί στην κατάσταση j μετά από άπειρο αριθμό βημάτων, προφανώς ισχύει ότι $\sum_{j=0}^{\infty} \pi_j = 1$. ∎

Αν μια μαρκοβιανή αλυσίδα δεν είναι ανάγωγη και εργοδική, δεν ορίζονται οι οριακές πιθανότητες. Ωστόσο, αν μια κλάση ισοδυναμίας μιας μαρκοβιανής αλυσίδας είναι εργοδική και αποτελεί επιπλέον κλειστό σύνολο, μπορούμε να την θεωρήσουμε ως μια επιμέρους μαρκοβιανή αλυσίδα μέσα στην αρχική και να ορίσουμε με τον τρόπο που περιγράψαμε νωρίτερα τις οριακές πιθανότητες για τις καταστάσεις αυτής της κλάσης.

Παράδειγμα 6.4.1 *Βρείτε για το παράδειγμα 6.1.1 τις οριακές πιθανότητες.*

Λύση. Προφανώς η μαρκοβιανή αλυσίδα που ορίζεται από το παράδειγμα 6.1.1 είναι ανάγωγη και εργοδική, επομένως υπάρχουν οι οριακές πιθανότητες. Ο πίνακας πιθανοτήτων μετάβασης πρώτης τάξης της αλυσίδας είναι ίσος με:

$$\mathbb{P} = \left\| \begin{array}{cc} a & 1-a \\ b & 1-b \end{array} \right\|.$$

Άρα οι εξισώσεις που προκύπτουν για τις οριακές πιθανότητες είναι οι:

$$\pi_0 = a\pi_0 + b\pi_1$$

$$\pi_1 = (1-a)\pi_0 + (1-b)\pi_1$$

και

$$\pi_0 + \pi_1 = 1.$$

Λύνοντας προκύπτει ότι $\pi_0 = \frac{b}{1+b-a}$ και $\pi_1 = \frac{1-a}{1+b-a}$.

∎

Παράδειγμα 6.4.2 *Βρείτε τις οριακές πιθανότητες των καταστάσεων της μαρκοβιανής αλυσίδας του παραδείγματος 6.1.3.*

Λύση. Ο πίνακας μετάβασης πρώτης τάξης για την μαρκοβιανή αλυσίδα του παραδείγματος 6.1.3 είναι ίσος με:

$$\mathbb{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix}.$$

Οι εξισώσεις που προκύπτουν είναι:

$$\pi_0 = 0.5\pi_0 + 0.3\pi_1 + 0.2\pi_2$$

$$\pi_1 = 0.4\pi_0 + 0.4\pi_1 + 0.3\pi_2$$

$$\pi_2 = 0.1\pi_0 + 0.3\pi_1 + 0.5\pi_2$$

και

$$\pi_0 + \pi_1 + \pi_2 = 1.$$

Λύνοντας προκύπτει ότι $\pi_0 = \frac{21}{62}$, $\pi_1 = \frac{23}{62}$ και $\pi_2 = \frac{18}{62}$.

∎

6.5 Χρόνος Παραμονής στις Μεταβατικές Καταστάσεις

Θεωρείστε ότι οι καταστάσεις μιας μαρκοβιανής αλυσίδας είναι αριθμημένες με τέτοιο τρόπο ώστε το σύνολο $T = \{1, 2, \ldots, t\}$ αποτελείται από τις μεταβατικές καταστάσεις της αλυσίδας. Έστω

$$\mathbb{P}_T = \begin{Vmatrix} P_{11} & P_{12} & \ldots & P_{1t} \\ \vdots & \vdots & \vdots & \vdots \\ P_{t1} & P_{t2} & \ldots & P_{tt} \end{Vmatrix}$$

όπου $\mathbb{P}_T$ είναι ο πίνακας πιθανοτήτων μετάβασης για τις μεταβατικές καταστάσεις μόνο.

Για δύο μεταβατικές καταστάσεις i και j, έστω s_{ij} το αναμενόμενο πλήθος των χρονικών στιγμών που η μαρκοβιανή αλυσίδα είναι στην κατάσταση j, δοθέντος ότι η διαδικασία άρχισε από την κατάσταση i. Ισχύει η εξής σχέση:

$$s_{ij} = \delta_{ij} + \sum_{k=1}^{t} P_{ik} s_{kj}$$

όπου $\delta_{ij} = 1$ αν $i = j$ και $\delta_{ij} = 0$ διαφορετικά. Αν $\mathbb{S}$ είναι ο πίνακας

$$\mathbb{S} = \begin{Vmatrix} s_{11} & s_{12} & \ldots & s_{1t} \\ \vdots & \vdots & \vdots & \vdots \\ s_{t1} & s_{t2} & \ldots & s_{tt} \end{Vmatrix}$$

τότε η παραπάνω εξίσωση θα σημαίνει για τους πίνακες $\mathbb{S}$ και $\mathbb{P}_T$ ότι:

$$\mathbb{S} = I + \mathbb{P}_T \mathbb{S} \tag{6.3}$$

όπου I είναι ο μοναδιαίος πίνακας. Η εξίσωση 6.3 είναι ισοδύναμη με την

$$\mathbb{S} = (I - \mathbb{P}_T)^{-1}.$$

Δηλαδή, οι ποσότητες s_{ij} προκύπτουν με το να αντιστρέψουμε τον πίνακα $I - \mathbb{P}_T$.

Παράδειγμα 6.5.1 *Θεωρείστε ότι ένας παίκτης σε ένα καζίνο ξεκινά με 3 μάρκες και ότι έχει πιθανότητα* 0.4 *σε κάθε γύρο του παιχνιδιού να κερδίσει μία μάρκα και* 0.6 *να χάσει μία μάρκα. Αν υποθέσουμε ότι ο παίκτης σταμάταει να παίζει όταν χάσει όλες τις μάρκες ή κερδίσει 7, καθορίστε:*

1. *τον αναμενόμενο χρόνο που ο παίκτης έχει 5 μάρκες*
2. *τον αναμενόμενο χρόνο που ο παίκτης έχει 2 μάρκες*
3. *τον αναμενόμενο χρόνο μέχρι ο παίκτης να σταματήσει να παίζει*

Λύση. Η μαρκοβιανή αλυσίδα που ορίζεται από το παράδειγμα της άσκησης έχει οκτώ καταστάσεις $\{0, 1, 2, 3, 4, 5, 6, 7\}$ όπου οι $\{0\}$ και $\{7\}$ είναι έμμονες, ενώ οι $\{1, 2, 3, 4, 5, 6\}$ είναι μεταβατικές. Συγκεκριμένα, ο πίνακας πιθανοτήτων μετάβασης $\mathbb{P}_T$ είναι ίσος με:

$$\mathbb{P}_T = \begin{Vmatrix} 0 & 0.4 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 0.6 & 0 \end{Vmatrix}.$$

Αντιστρέφοντας τον πίνακα $I - \mathbb{P}_T$ έχουμε ότι

$$\mathbb{S} = \begin{Vmatrix} 1.6149 & 1.0248 & 0.6314 & 0.3691 & 0.1943 & 0.0777 \\ 1.5372 & 2.5619 & 1.5784 & 0.9228 & 0.4857 & 0.1943 \\ 1.4206 & 2.3677 & 2.9990 & 1.7533 & 0.9228 & 0.3691 \\ 1.2458 & 2.0763 & 2.6299 & 2.9990 & 1.5784 & 0.6314 \\ 0.9835 & 1.6391 & 2.0763 & 2.3677 & 2.5619 & 1.0248 \\ 0.5901 & 0.9835 & 1.2458 & 1.4206 & 1.5372 & 1.6149 \end{Vmatrix}.$$

Επομένως, η απάντηση στα δύο πρώτα ερωτήματα είναι ίση με $s_{35} = 0.9228$ και $s_{32} = 2.3677$ αντίστοιχα. Για να βρούμε τον αναμενόμενο χρόνο μέχρι ο παίκτης να σταματήσει να παίζει, θα πρέπει να αθροίσουμε όλες τις τιμές $s_{31} + s_{32} + s_{33} + s_{34} + s_{35} = 9.8325$. ∎