



# Εθνικό Μετσόβιο Πολυτεχνείο

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ  
ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

## Προβλήματα Βελτιστοποίησης και Εκμάθησης σε Στοχαστικά Πρότυπα

*Γιώργος Βερναρδάκης*

Επιβλέπων: Μιχαήλ Λουλάκης

*Η εργασία είναι αφιερωμένη σε εκείνους που μου απέδειξαν την ύπαρξη αστείρευτων πηγών υπομονής και μου συμπαραστάθηκαν σε μια περίοδο άρνησης για δημιουργία.*

*Αθήνα, Ιούλιος 2020*

# Περιεχόμενα

<b>1</b>	<b>Γενικές Εισαγωγικές Έννοιες Πιθανοτήτων</b>	<b>5</b>
1.1	Έννοιες & Ορισμοί . . . . .	5
1.1.1	Χώροι Πιθανότητας . . . . .	5
1.1.2	Τυχαία Μεταβλητή . . . . .	8
1.1.3	Ανεξαρτησία . . . . .	10
1.2	Ολοκλήρωμα Lebesgue και Μέση Τιμή . . . . .	11
1.2.1	Ολοκλήρωμα Lebesgue . . . . .	11
1.2.2	Μέση Τιμή & Ροπές . . . . .	13
1.2.3	Κατανομή . . . . .	15
1.2.4	Δεσμευμένη Μέση Τιμή . . . . .	19
1.3	Χώροι $L^p$ και Τρόποι Σύγκλισης Τυχαίων Μεταβλητών . . . . .	20
1.3.1	Χώροι $L^p$ . . . . .	20
1.3.2	Τρόποι Σύγκλισης Τυχαίων Μεταβλητών . . . . .	21
1.4	Χρήσιμα Αποτελέσματα & Εργαλεία . . . . .	23
1.4.1	Παραδείγματα διακριτών Κατανομών . . . . .	23
1.4.2	Παραδείγματα Συνεχών Κατανομών . . . . .	26
1.4.3	Διάσημες Ανισότητες . . . . .	30
1.4.4	Θεωρήματα Σύγκλισης . . . . .	32
1.4.5	Θεωρήματα Σύγκλισης Μέσης Τιμής . . . . .	33
1.4.6	Οριακά Θεωρήματα . . . . .	35
<b>2</b>	<b>Στοχαστικές Διαδικασίες</b>	<b>37</b>
2.1	Γενικές Βασικές Έννοιες . . . . .	37
2.1.1	Έννοιες & Ορισμοί . . . . .	37
2.1.2	Συνέχεια & Ολοκληρωσιμότητα . . . . .	39
2.1.3	Διαδικασίες Gauss, Wiener και Lévy . . . . .	42
2.2	Αλυσίδες Markov . . . . .	43
2.2.1	Μαρκοβιανή Ιδιότητα . . . . .	43

2.2.2	Χρόνοι Διακοπής . . . . .	46
2.2.3	Χαρακτηρισμοί Αλυσίδας . . . . .	47
2.2.4	Ασυμπτωτική Κατανομή & Εργοδικό Θεώρημα . . . . .	51
2.3	Διαδικασίες Martingale . . . . .	52
2.3.1	Ορισμοί & Ειδικές Περιπτώσεις . . . . .	52
2.3.2	Στοιχειώδη Θεωρήματα . . . . .	54
2.3.3	Μεγιστικές Ανισότητες . . . . .	57
<b>3</b>	<b>Βέλτιστη Διακοπή</b>	<b>61</b>
3.1	Το Πρόβλημα σε Διακριτό Χρόνο . . . . .	62
3.1.1	Η περίπτωση των Martingales . . . . .	62
3.1.2	Η περίπτωση των Μαρχοβιανών Αλυσίδων . . . . .	73
3.2	Το Πρόβλημα σε Συνεχή Χρόνο . . . . .	87
3.2.1	Η περίπτωση των Martingales . . . . .	87
<b>4</b>	<b>Το Πρόβλημα του Ληστή με τα Πολλά Χέρια (Multi-Armed Bandit Problem)</b>	<b>91</b>
4.1	Εισαγωγή στο Πρόβλημα . . . . .	92
4.2	Bernoulli Bandit . . . . .	93
4.2.1	Ορισμός του Προβλήματος . . . . .	94
4.2.2	Ο Αλγόριθμος $\epsilon$ -Greedy . . . . .	96
4.2.3	Ο Αλγόριθμος Upper Confidence Bounds (UCB) . . . . .	97
4.3	Thompson Sampling . . . . .	98
4.4	Παράδειγμα Bernoulli Bandit . . . . .	101
<b>5</b>	<b>Εφαρμογές</b>	<b>111</b>
5.1	Εύρεση μιας στοχαστικά βέλτιστης διαδρομής . . . . .	111
5.2	Εύρεση λύσης σε ένα στοχαστικό Πρόβλημα του Πλανόδιου Πωλητή (Traveling Salesman Problem) . . . . .	133

# Εισαγωγή

Κάθε άνθρωπος με φιλοσοφικές ανησυχίες είχε ανέκαθεν την τάση να αναρωτιέται κατά πόσο ο κόσμος κυβερνάται από την τύχη ή τα πάντα έχουν προκαθοριστεί από κάποια ανώτερη δύναμη. Σε κάθε περίπτωση, στο στάδιο που βρίσκεται η εξέλιξη του ανθρώπινου είδους, εξακολουθούν και συμβαίνουν αμέτρητα γεγονότα τα οποία ο άνθρωπος δε δύναται να προβλέψει. Ωστόσο, έχουν γίνει υπέρογκες προσπάθειες προκειμένου να μελετηθούν αυτές οι αβεβαιότητες και να ληφθούν οι αντίστοιχες εκτιμήσεις, με τη βοήθεια μαθηματικών μοντέλων που βασίζονται στη Θεωρία Πιθανοτήτων. Πλέον, με σύμμαχο τους υπολογιστές και την τεχνητή νοημοσύνη, η μελέτη της αβεβαιότητας με προσομοιώσεις μοιάζει πιο επίκαιρη, αλλά και πιο εξελιγμένη από ποτέ. Εκεί, λοιπόν, παραπέμπει και ο τίτλος της παρούσας εργασίας, η οποία πραγματεύεται προβλήματα βελτιστοποίησης και εκμάθησης σε στοχαστικά πρότυπα.

Στις πρώτες δύο ενότητες παρατίθενται βασικές έννοιες από τη Θεωρία Μέτρου σε χώρους πιθανότητας, καθώς και του κλάδου της Στατιστικής, καταλήγοντας στις Στοχαστικές Διαδικασίες. Οι τελευταίες, θα μας επιτρέψουν να ορίσουμε τα Προβλήματα Βέλτιστης Διακοπής και να περιγράψουμε τη βασική ιδέα της μοντελοποίησής τους στην τρίτη ενότητα. Από εκεί, με τη βοήθεια του Θεωρήματος Gittins & Jones, θα μεταφερθούμε στην ενότητα τέσσερα, στα Προβλήματα του Ληστή με τα Πολλά Χέρια (Multi-Armed Bandit Problems) και στον αλγόριθμο Thompson Sampling που επιλύει αποδοτικά τέτοιου είδους προβλήματα. Τέλος, στην ενότητα πέντε θα παρουσιάσουμε δύο πραγματικές εφαρμογές που αποδεικνύουν το ενδιαφέρον των εν λόγω προβλημάτων καθώς και την αποτελεσματικότητα των αντίστοιχων αλγορίθμων επίλυσης. Όλοι οι κώδικες έχουν γραφεί στο στατιστικό πακέτο *R*. Επιβλέπων καθηγητής ήταν ο Μιχαήλ Λουλάκης.

# Ενότητα 1

## Γενικές Εισαγωγικές Έννοιες Πιθανοτήτων

Στην πρώτη ενότητα θα επιχειρήσουμε να προσφέρουμε στον αναγνώστη μία γρήγορη πλην στοχευμένη ανασκόπηση εννοιών από τον κλάδο των Πιθανοτήτων που πιστεύουμε ότι απαιτούνται, ούτως ώστε να συνδυαστούν η εύκολη και ομαλή ανάγνωση, με τη μέγιστη δυνατή κατανόηση και την αυστηρή μαθηματική συνέπεια. Για περαιτέρω ανάγνωση και πηγές, παραπέμπουμε τον αναγνώστη στα [1], [2], [3], [4], [5].

### 1.1 Έννοιες & Ορισμοί

#### 1.1.1 Χώροι Πιθανότητας

**Ορισμός 1.1.** Ορίζουμε ως  **$\sigma$ -Άλγεβρα** πάνω σε ένα σύνολο  $\Omega$  μία συλλογή  $\mathcal{F}$  από υποσύνολα του  $\Omega$  με τρόπο τέτοιοιον ώστε:

- Το  $\Omega$  ανήκει στη  $\mathcal{F}$ .
- Έστω  $A \subseteq \Omega$ . Αν  $A \in \mathcal{F}$ , τότε  $A^c \in \mathcal{F}$ .
- Έστω  $\{A_n\}_{n \in \mathbb{N}}$  μία ακολουθία υποσυνόλων του  $\Omega$ . Αν  $A_n \in \mathcal{F}$ ,  $\forall n \in \mathbb{N}$ , τότε  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$  (δηλαδή είναι κλειστή ως προς την ένωση).

Από τον ορισμό έπεται ότι η  $\mathcal{F}$  θα περιέχει και το κενό σύνολο  $\emptyset$ , ενώ είναι κλειστή και ως προς την τομή.

**Ορισμός 1.2. Δειγματικό Χώρο** ενός τυχαίου πειράματος θα ονομάζουμε το σύνολο με όλες τις πιθανές εκβάσεις του πειράματος αυτού. Από εδώ και στο εξής, με  $\Omega$  θα συμβολίζουμε τον δειγματικό χώρο και με  $\omega$  τις πιθανές εκβάσεις, έτσι ώστε  $\omega \in \Omega$ .

**Ορισμός 1.3.** Το ζευγάρι  $(\Omega, \mathcal{F})$  θα ονομάζεται **Μετρήσιμος Χώρος**.

**Ορισμός 1.4.** Έστω  $(\Omega, \mathcal{F})$  μετρήσιμος χώρος και  $\{A_n\}_{n \geq 1}$  ακολουθία στοιχείων της  $\mathcal{F}$ . Ορίζουμε:

- $\limsup_{n \rightarrow \infty} A_n = \bigcap_{n \geq 1} \bigcup_{j \geq n} A_j = \{\omega \in \Omega : \omega \text{ ανήκει σε άπειρα } A_n\}$
- $\liminf_{n \rightarrow \infty} A_n = \bigcup_{n \geq 1} \bigcap_{j \geq n} A_j = \{\omega \in \Omega : \omega \text{ ανήκει σε όλα τα } A_n \text{ από κάποιο } n_0 \text{ και μετά}\}$

Αν τα δύο παραπάνω όρια υπάρχουν και είναι ίσα, τότε ορίζουμε:

$$\lim_{n \rightarrow \infty} A_n = \liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n. \quad (1.1)$$

Γενικότερα, για μια οποιαδήποτε ακολουθία  $\{x_n\}_{n \in \mathbb{N}}$ , ορίζουμε:

- $\limsup_{n \rightarrow \infty} x_n := \inf_{n \geq 1} \sup_{j \geq n} x_j$
- $\liminf_{n \rightarrow \infty} x_n := \sup_{n \geq 1} \inf_{j \geq n} x_j$

Αντίστοιχα, στην περίπτωση που η  $\{x_n\}_{n \in \mathbb{N}}$  παίρνει τιμές στο  $[-\infty, +\infty]$  και τα δύο παραπάνω όρια υπάρχουν και είναι ίσα, θα έχουμε ότι:

$$\lim_{n \rightarrow \infty} x_n = \liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n. \quad (1.2)$$

**Ορισμός 1.5.** Μέτρο στον μετρήσιμο χώρο  $(\mathbb{X}, \mathcal{S})$  λέμε κάθε συνάρτηση  $\mu : \mathcal{S} \rightarrow [0, \infty]$  που ικανοποιεί τις εξής ιδιότητες:

1.  $\mu(\emptyset) = 0$ .
2.  **$\sigma$ -Αθροιστικότητα:** Για κάθε αριθμήσιμη οικογένεια ανά δύο ξένων συνόλων  $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}$ , έπεται ότι  $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ .

**Ορισμός 1.6.** Αν επιπλέον υπάρχει υποσύνολο  $S \in \mathcal{S}$ , το πολύ αριθμήσιμο, τέτοιο ώστε  $\mu(\mathbb{X} \setminus S) = 0$ , τότε θα λέμε ότι το  $\mu$  είναι **Διακριτό Μέτρο**.

Αν επίσης ισχύει ότι  $\mu(\mathbb{X}) < \infty$ , τότε το  $\mu$  θα λέγεται **Πεπερασμένο Μέτρο** και προφανώς, θα έπεται ότι  $\mu(A) < \infty$  για κάθε  $A \in \mathcal{S}$ .

**Ορισμός 1.7.** Έστω συνάρτηση  $P : \mathcal{F} \rightarrow [0, 1]$ . Θα λέμε ότι η  $P$  είναι **Μέτρο Πιθανότητας** στον  $(\Omega, \mathcal{F})$  αν και μόνον αν πληροί τις εξής προϋποθέσεις:

- Είναι μέτρο στον  $(\Omega, \mathcal{F})$ .
- **Κανονικότητα:** Το μέτρο όλου του δειγματικού χώρου  $\Omega$  είναι ίσο με 1. Δηλαδή,  $P(\Omega) = 1$ .

Παρατηρούμε ότι τα μέτρα πιθανότητας είναι πεπερασμένα μέτρα.

**Ορισμός 1.8.** Η τριάδα  $(\mathbb{X}, \mathcal{S}, \mu)$  ορίζεται ως **Χώρος Μέτρου**. Ειδικότερα, στην περίπτωση της τριάδας  $(\Omega, \mathcal{F}, P)$ , όπου  $\Omega$  ο δειγματικός χώρος,  $\mathcal{F}$  μια  $\sigma$ -άλγεβρα στον  $\Omega$  και  $P$  ένα μέτρο πιθανότητας, ο χώρος μέτρου ορίζεται ως **Χώρος Πιθανότητας**.

**Ορισμός 1.9.** Έστω  $(\mathbb{X}, \mathcal{S}, \mu)$ ,  $(\mathbb{Y}, \mathcal{T}, \nu)$  δύο χώροι μέτρου. Ορίζουμε ως **Χώρο Μέτρου-Γινόμενο** τον χώρο  $(\mathbb{X} \times \mathbb{Y}, \mathcal{S} \otimes \mathcal{T}, \mu \otimes \nu)$ , όπου  $\mathcal{S} \otimes \mathcal{T}$  είναι η  $\sigma$ -Άλγεβρα Γινόμενο στο καρτεσιανό γινόμενο  $\mathbb{X} \times \mathbb{Y}$  που παράγεται από τα ανοιχτά υποσύνολα της μορφής  $B_1 \times B_2$ , έτσι ώστε  $B_1 \in \mathcal{S}$  και  $B_2 \in \mathcal{T}$ .



Επίσης, με  $\mu \otimes \nu$  συμβολίζουμε το **Μέτρο Γινόμενο** στον  $(\mathbb{X} \times \mathbb{Y}, \mathcal{S} \otimes \mathcal{T})$ , το οποίο ικανοποιεί την εξής συνθήκη:

$$(\mu \otimes \nu)(B_1 \times B_2) = \mu(B_1)\nu(B_2), \quad (1.3)$$

για κάθε  $B_1 \in \mathcal{S}$ ,  $B_2 \in \mathcal{T}$ .

### 1.1.2 Τυχαία Μεταβλητή

**Ορισμός 1.10.** Έστω  $(\mathbb{X}, \mathcal{S})$  και  $(\mathbb{Y}, \mathcal{T})$  δύο μετρήσιμοι χώροι. Μια συνάρτηση  $f: X \rightarrow Y$  θα λέγεται  **$\mathcal{S}/\mathcal{T}$ -μετρήσιμη**, αν και μόνον αν, για κάθε  $E \in \mathcal{T}$ , η αντίστροφη εικόνα του  $E$  κάτω από την  $f$  θα ανήκει στη  $\mathcal{S}$ . Δηλαδή, αν και μόνον αν,

$$f^{-1}(E) := \{x \in X \mid f(x) \in E\} \in \mathcal{S}, \quad \forall E \in \mathcal{T}. \quad (1.4)$$

Από εδώ και στο εξής, αν οι  $\sigma$ -άλγεβρες είναι γνωστές, η  $f$  ενδέχεται να αναφέρεται απλώς ως «μετρήσιμη», εκτός και αν θέλουμε να δώσουμε έμφαση σε κάποια εκ των δύο  $\sigma$ -αλγεβρών (π.χ.  $\mathcal{S}$ -μετρήσιμη ή  $\mathcal{T}$ -μετρήσιμη).

**Ορισμός 1.11.** Έστω  $(\Omega, \mathcal{F}, P)$  χώρος πιθανότητας και  $(\mathbb{X}, \mathcal{S})$  ένας μετρήσιμος χώρος. Τότε, η συνάρτηση  $X: \Omega \rightarrow \mathbb{X}$  θα λέγεται **Τυχαία Μεταβλητή** με τιμές στον  $\mathbb{X}$ , αν και μόνον αν, είναι μετρήσιμη. Δηλαδή, αν και μόνον αν:

$$X^{-1}(B) \in \mathcal{F}, \quad \forall B \in \mathcal{S}. \quad (1.5)$$

Ειδικότερα, αν  $\mathbb{X} \subseteq \mathbb{R}$ , η  $X$  θα λέγεται **Πραγματική Τυχαία Μεταβλητή**.

**Ορισμός 1.12.** Borel σ-άλγεβρα ενός τοπολογικού χώρου  $\mathbb{X}$  θα λέμε την ελάχιστη σ-άλγεβρα που παράγεται από όλα τα ανοιχτά υποσύνολα του  $\mathbb{X}$  και θα τη συμβολίζουμε με  $\mathcal{B}(\mathbb{X})$ . Στην περίπτωση που  $\mathbb{X} = \mathbb{R}$  (χώρος με τοπολογία), η Borel σ-άλγεβρα θα παράγεται από τα υποσύνολα  $\{(-\infty, r] : r \in \mathbb{R}\}$ . Ως εκ τούτου, μπορούμε να ορίσουμε ως πραγματική τυχαία μεταβλητή τη συνάρτηση  $X : \Omega \rightarrow \mathbb{R}$ , αν και μόνον αν,

$$\{\omega : X(\omega) \leq r\} \in \mathcal{F} \quad \forall r \in \mathbb{R}, \quad (1.6)$$

καθώς ισχύει ότι  $\{\omega : X(\omega) \leq r\} = X^{-1}((-\infty, r])$ .

**Ορισμός 1.13.** Έστω μετρήσιμος χώρος  $(\mathbb{X}, \mathcal{B})$ , όπου  $\mathcal{B}$  η Borel σ-άλγεβρα. Τότε, ο χώρος θα λέγεται **Μετρήσιμος Χώρος Borel**.

**Ορισμός 1.14.** Έστω  $(\mathbb{X}, \mathcal{S})$  μετρήσιμος χώρος και  $(\mathbb{Y}, \mathcal{B})$  μετρήσιμος χώρος Borel. Ορίζουμε ως **Borel-μετρήσιμη** συνάρτηση, κάθε μετρήσιμη συνάρτηση  $f : \mathbb{X} \rightarrow \mathbb{Y}$ .

**Ορισμός 1.15.** Αν για τη  $X : \Omega \rightarrow \mathbb{X}$  το πεδίο τιμών  $\mathbb{X}$  είναι αριθμήσιμο, τότε αναγκαστικά και η εικόνα  $X(\Omega)$  της  $X$  μέσω του  $\Omega$  θα είναι αριθμήσιμη και άρα θα λέμε ότι η  $X$  είναι μια **διακριτή** τυχαία μεταβλητή. Διαφορετικά, αν το πεδίο τιμών  $\mathbb{X}$  είναι υπεραριθμήσιμο και η εικόνα  $X(\Omega)$  είναι υπεραριθμήσιμη, τότε η  $X$  θα λέγεται **συνεχής**.

**Ορισμός 1.16.** Έστω  $(\Omega, \mathcal{F}, \mathbb{P})$  χώρος πιθανότητας και  $X, Y : \Omega \rightarrow \mathbb{R}$  δύο πραγματικές τυχαίες μεταβλητές. Ορίζουμε **σχέση διάταξης**  $\preceq$  ανάμεσα στη  $X$  και στην  $Y$ , με τρόπο ώστε:

$$X \preceq Y \iff X \leq Y, \mathbb{P} - \sigma.β., \quad (1.7)$$

δηλαδή αν  $X(\omega) \leq Y(\omega)$ ,  $\forall \omega \in \Omega \setminus A$ , με  $P(A) = 0$ . Από εδώ και στο εξής, με τον συμβολισμό  $X \leq Y$  θα εννοούμε  $X \leq Y$ , **P-σχεδόν βεβαίως**.

Το μετροθεωρητικό ανάλογο σε έναν χώρο μέτρου  $(\mathbb{X}, \mathcal{S}, \mu)$  αφορά πραγματικές μετρήσιμες συναρτήσεις, έστω  $f, g$ , και λέμε ότι

$$f \preceq g \iff f \leq g, \mu - \sigma.π., \quad (1.8)$$

δηλαδή αν  $f(x) \leq g(x)$ ,  $\forall x \in \mathbb{X} \setminus A$ , με  $\mu(A) = 0$ . Αντίστοιχα, με τον συμβολισμό  $f \leq g$  θα εννοούμε  $f \leq g$ ,  **$\mu$ -σχεδόν παντού**.

### 1.1.3 Ανεξαρτησία

**Ορισμός 1.17.** Έστω χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ . Ορίζουμε την **Ανεξαρτησία Ενδεχομένων**, την **Ανεξαρτησία  $\sigma$ -Αλγεβρών** και την **Ανεξαρτησία Τυχαίων Μεταβλητών**:

- Δύο ενδεχόμενα  $A, B \subset \Omega$  θα λέγονται **ανεξάρτητα** ( $A \perp B$ ), αν και μόνο αν,

$$P(A \cap B) = P(A)P(B). \quad (1.9)$$

- Δύο υπο- $\sigma$ -άλγεβρες  $\mathcal{A}_1, \mathcal{A}_2$  θα λέγονται **ανεξάρτητες**, αν και μόνο αν, για κάθε  $A_1 \in \mathcal{A}_1$  και  $A_2 \in \mathcal{A}_2$  ισχύει ότι

$$P(A_1 \cap A_2) = P(A_1)P(A_2). \quad (1.10)$$

- Αντίστοιχα, δύο τυχαίες μεταβλητές  $X, Y : \Omega \rightarrow \mathbb{R}$  θα λέγονται **ανεξάρτητες**, αν και μόνο αν, για κάθε  $x \in X(\Omega)$  και  $y \in Y(\Omega)$  τα σύνολα  $\{X \leq x\}$  και  $\{Y \leq y\}$  είναι ανεξάρτητα. Δηλαδή, αν  $F_X(x)$  και  $F_Y(y)$  η συνάρτηση κατανομής πιθανότητας της  $X$  και της  $Y$ , αντίστοιχα, τότε θα

πρέπει για κάθε  $x, y$  η από κοινού συνάρτηση κατανομής πιθανότητας των  $(X, Y)$  να γράφεται ως

$$F_{X,Y}(x, y) = F_X(x)F_Y(y). \quad (1.11)$$

**Ορισμός 1.18.** Ορίζουμε τη **Δεσμευμένη Πιθανότητα** του ενδεχομένου  $A$  δεδομένου ενός ενδεχομένου  $B$  ως την ποσότητα

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1.12)$$

Άμεσα προκύπτει ότι:

1.  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$  (νόμος του Bayes).
2. Αν  $A \perp B$ , τότε  $P(A|B) = P(A)$ .

## 1.2 Ολοκλήρωμα Lebesgue και Μέση Τιμή

### 1.2.1 Ολοκλήρωμα Lebesgue

**Ορισμός 1.19.** Μια συνάρτηση  $f : \Omega \rightarrow [-\infty, +\infty]$  θα λέγεται **απλή** αν η εικόνα της είναι πεπερασμένο σύνολο, έστω  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . Αν επίσης θέσουμε  $A_i := f^{-1}(\{\alpha_i\})$ , τότε η  $f$  γράφεται

$$f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}. \quad (1.13)$$

Το  $\mathbb{A} = \{A_1, A_2, \dots, A_n\}$  θα αποτελεί μια διαμέριση του  $\Omega$  (δηλαδή τα στοιχεία του είναι μη κενά, ξένα ανά δύο, με ένωση το  $\Omega$ ) και η  $f$  θα είναι μετρήσιμη, αν και μόνο αν, για κάθε  $i = 1, 2, \dots, n$  τα  $A_i$  είναι μετρήσιμα. Σε αυτήν την περίπτωση, η παραπάνω γραφή της  $f$  είναι μοναδική και θα λέγεται **κανονική**. Η μοναδικότητα της  $f$  δεν ισχύει στις περιπτώσεις που το  $\mathbb{A}$  δεν διαμερίζει τον  $\Omega$ .

**Ορισμός 1.20.** Έστω χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ . Ορίζουμε το **Ολοκλήρωμα Lebesgue** μιας τυχαίας μεταβλητής  $X : \Omega \rightarrow [-\infty, +\infty]$  σε τρία στάδια:

1. Έστω  $X : \Omega \rightarrow [0, +\infty]$  απλή μετρήσιμη συνάρτηση (απλή τυχαία μεταβλητή) με κανονική μορφή  $X = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ . Ορίζουμε το ολοκλήρωμα Lebesgue της  $X$  ως προς το μέτρο πιθανότητας  $P$  ως:

$$\int X dP \equiv \int_{\Omega} X dP := \sum_{i=1}^n a_i P(A_i), \quad (1.14)$$

το οποίο παίρνει τιμές στο  $[0, +\infty]$ .

2. Έστω  $X : \Omega \rightarrow [0, +\infty]$  μετρήσιμη συνάρτηση (τυχαία μεταβλητή). Ορίζουμε το ολοκλήρωμα Lebesgue της μη αρνητικής  $X$  ως προς το μέτρο πιθανότητας  $P$  ως:

$$\int X dP := \sup_{Y \in \mathcal{S}} \left\{ \int Y dP \right\}, \quad (1.15)$$

όπου με  $\mathcal{S}$  συμβολίζουμε το σύνολο με τις απλές και μη αρνητικές τυχαίες μεταβλητές  $Y$  για τις οποίες ισχύει  $Y \leq X$ ,  $P$ -σ.β.

3. Έστω  $X : \Omega \rightarrow [-\infty, +\infty]$  μετρήσιμη συνάρτηση (τυχαία μεταβλητή). Ορίζουμε το ολοκλήρωμα Lebesgue της  $X$  ως προς το μέτρο πιθανότητας  $P$  ως:

$$\int X dP := \int X^+ dP - \int X^- dP, \quad (1.16)$$

εφόσον στο δεξί μέλος δεν παρουσιάζεται απροσδιοριστία της μορφής  $\infty - \infty$ .

Τέλος, στην περίπτωση που έχουμε ένα  $A \in \mathcal{F}$ , ορίζουμε το ολοκλήρωμα της  $X$  πάνω στο  $A$  ως προς το μέτρο  $P$  ως:

$$\int_A X dP := \int X 1_A dP. \quad (1.17)$$

## 1.2.2 Μέση Τιμή & Ροπές

**Ορισμός 1.21.** Έστω χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$  και τυχαία μεταβλητή  $X : \Omega \rightarrow [-\infty, +\infty]$ . Ορίζουμε τη **Μέση Τιμή** της  $X$  ως το ολοκλήρωμα Lebesgue της  $X$  και τη συμβολίζουμε με  $E[X]$ . Δηλαδή,

$$E[X] := \int X dP. \quad (1.18)$$

Με τον συμβολισμό  $E_P[X]$  θα καθιστούμε ξεκάθαρο το ότι η μέση τιμή υπολογίζεται ως προς το μέτρο  $P$ . Επίσης, αν η μέση τιμή της  $X$  είναι πεπερασμένη, θα λέμε ότι η  $X$  είναι **ολοκληρώσιμη**.

**Ιδιότητες** μέσης τιμής (ολοκληρώματος Lebesgue) για  $X, Y : \Omega \rightarrow [-\infty, +\infty]$  τυχαίες μεταβλητές με καλά ορισμένη μέση τιμή:

1. **Γραμμικότητα:**  $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$ , για κάθε  $\alpha, \beta \in \mathbb{R}$   
(Υπό την προϋπόθεση ότι δεν παρουσιάζεται απροσδιοριστία της μορφής  $\infty - \infty$ ).
2. **Τριγωνική Ανισότητα:**  $|E[X]| \leq E[|X|]$ .
3. **Μονοτονία:** Αν  $X \leq Y$   $P$ -σ.β., τότε  $E[X] \leq E[Y]$ .

Από την ιδιότητα 3 έπεται ότι αν η  $X$  είναι απολύτως ολοκληρώσιμη, θα είναι και ολοκληρώσιμη.

**Ορισμός 1.22.** Έστω χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$  και τυχαία μεταβλητή  $X : \Omega \rightarrow \mathbb{R}$  με  $E[X] < \infty$ . Ορίζουμε ως **Διασπορά** της  $X$  την ποσότητα

$$\text{Var}[X] := E\{[X - E[X]]^2\}, \quad (1.19)$$

που είναι στοιχείο του  $[0, +\infty]$  και για την οποία ισχύει ότι αν  $E[X^2] < \infty$ , τότε  $\text{Var}(X) < \infty$ .

**Ορισμός 1.23.** Έστω  $X, Y : \Omega \rightarrow \mathbb{R}$ , τυχαίες μεταβλητές, με  $E[X], E[Y] < \infty$ , ενώ η  $E[XY]$  ορίζεται. Ορίζουμε ως **Συνδιακύμανση** των  $X, Y$  την ποσότητα

$$\text{Cov}[X, Y] := E\{[X - E[X]][Y - E[Y]]\}, \quad (1.20)$$

η οποία είναι στοιχείο του  $[-\infty, +\infty]$ .

**Ορισμός 1.24.** Γενικότερα, **Κεντρική Ροπή Τάξης  $n$**  και **Ροπή Τάξης  $n$**  θα λέμε τις ποσότητες

$$\mu_n = E\{[X - E[X]]^n\}, n \in \mathbb{N} \quad (1.21)$$

και

$$m_n = E[X^n], n \in \mathbb{N}, \quad (1.22)$$

αντίστοιχα.

**Ορισμός 1.25.** Ακόμα πιο γενικά, ορίζουμε τη **Χαρακτηριστική Συνάρτηση** της πραγματικής τυχαίας μεταβλητής  $X$  ως εξής:

$$\phi_X(t) := E[e^{itX}], t \in \mathbb{R}, \quad (1.23)$$

όπου  $\phi_X : \mathbb{R} \rightarrow \mathbb{C}$  και  $i$  ο μοναδιαίος φανταστικός. Από την παραπάνω σχέση, μπορούμε να υπολογίσουμε όλες τις ροπές μιας τυχαίας μεταβλητής, όταν αυτές υπάρχουν, με τον εξής τρόπο:

$$E[X^n] = i^{-n} \phi_X^{(n)}(0), n \in \mathbb{N}, \quad (1.24)$$

όπου  $\phi_X^{(n)}$  είναι η  $n$ -οστή παράγωγος της  $\phi_X$ .

### 1.2.3 Κατανομή

**Ορισμός 1.26.** Έστω  $(\Omega, \mathcal{F}, P)$  χώρος πιθανότητας,  $(\mathbb{X}, \mathcal{S})$  μετρήσιμος χώρος και  $X : \Omega \rightarrow \mathbb{X}$  τυχαία μεταβλητή. Η συνολοσυνάρτηση  $P^X : \mathcal{S} \rightarrow [0, 1]$ , με

$$P^X(B) := P(X^{-1}(B)) = P(X \in B), \quad (1.25)$$

για κάθε  $B \in \mathcal{S}$  θα λέγεται **Κατανομή** της  $X$ . Εύκολα αποδεικνύεται ότι το  $P^X$  είναι μέτρο πιθανότητας στον  $(\mathbb{X}, \mathcal{S})$ . Επίσης, το  $P^X$  λέγεται και εικόνα του  $P$  μέσω της  $X$ .

**Ορισμός 1.27.** Έστω  $P^X$  η κατανομή μιας τυχαία μεταβλητής  $X$  με τιμές στον μετρήσιμο χώρο  $(\mathbb{X}, \mathcal{S})$ . Η **Πυκνότητα της  $X$** , έστω  $f$ , ως προς ένα μέτρο αναφοράς  $\mu$  θα είναι η παράγωγος Radon-Nikodym:

$$f = \frac{dP^X}{d\mu} \quad (1.26)$$

και θα ισχύει ότι



$$P(X \in A) = \int_{X^{-1}(A)} dP = \int_A dP^X = \int_A f d\mu. \quad (1.27)$$

**Ορισμός 1.28.** Έστω ο μετρήσιμος χώρος  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . Ως **Μέτρο Lebesgue** στον  $\mathbb{R}$  θα λέμε το μέτρο  $\lambda$  στον  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  για το οποίο ισχύει:

$$\lambda(I) = \text{length}(I), \quad (1.28)$$

για κάθε διάστημα  $I \subset \mathbb{R}$ , με τη συνάρτηση  $\text{length}$  να υπολογίζει το μήκος των διαστημάτων του  $\mathbb{R}$ , έτσι ώστε

$$\lambda([a, b]) = \text{length}([a, b]) = b - a. \quad (1.29)$$

**Ορισμός 1.29.** Έστω, τώρα, η ειδική περίπτωση (και πιο συνηθισμένη), όπου  $P^X$  μέτρο πιθανότητας στον  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ,  $\lambda$  το μέτρο Lebesgue και  $f : \mathbb{R} \rightarrow [0, \infty]$  μια Borel-μετρήσιμη συνάρτηση. Η  $f$  θα λέγεται **Πυκνότητα του  $P^X$** , αν και μόνο αν,

$$P^X(A) = \int_A f(x) d\lambda(x), \quad (1.30)$$

για κάθε  $A \in \mathcal{B}(\mathbb{R})$ .

**Ορισμός 1.30.** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ ,  $X : \Omega \rightarrow \mathbb{R}$  τυχαία μεταβλητή και  $f : \mathbb{R} \rightarrow [0, \infty]$  μια Borel-μετρήσιμη συνάρτηση. Θα λέμε ότι η  $f$  είναι **Πυκνότητα της  $X$** , αν και μόνο αν, η  $f$  είναι πυκνότητα της κατανομής  $P^X$  της  $X$ .

Από αυτό, έπεται ότι, αν  $f$  πυκνότητα της  $X$ , τότε για κάθε  $A \in \mathcal{B}(\mathbb{R})$ , θα ισχύει ότι:

$$P^X(A) = P(X \in A) = \int_A f(x) d\lambda(x), \quad (1.31)$$

**Ορισμός 1.31.** Στη λίγο ειδικότερη περίπτωση που το  $A = [a, b] \subset \mathbb{R}$  (ή της μορφής  $(a, b)$  ή  $(a, b]$  ή  $[a, b)$ ), τότε το μέτρο Lebesgue  $\lambda$  θα είναι το μήκος του διαστήματος και άρα θα ισχύει ότι:

$$P^X(A) = P(X \in A) = P(X \in [a, b]) = \int_a^b f(x) dx. \quad (1.32)$$

Το τελευταίο αποτέλεσμα υπονοεί ότι η  $P^X$  ορίζεται ως **Συνεχής Κατανομή** και η  $f$  ως **Συνάρτηση Πυκνότητας Πιθανότητας**.

Επίσης, για κάθε  $h : \mathbb{R} \rightarrow [-\infty, +\infty]$  θα ισχύει ότι:

$$E_P[h(X)] = \int h(x)f(x) dx, \quad (1.33)$$

όταν τουλάχιστον ένα από τα δύο μέλη ορίζεται.

**Ορισμός 1.32.** Αντίστοιχα, αν η  $X$  είναι διακριτή, έτσι ώστε η εικόνα  $\mathbb{X} = X(\Omega)$  να είναι αριθμήσιμη και  $P^X(\mathbb{X}) = 1$ , τότε για  $A \subset \mathbb{X}$  θα έχουμε ότι  $A = \bigcup_{x \in A} \{x\}$  και άρα:

$$P^X(A) = \sum_{x \in A} P^X(\{x\}) \quad (1.34)$$

ή

$$P(X \in A) = \sum_{x \in A} P(X = x). \quad (1.35)$$

Εδώ, η  $P^X$  ορίζεται να είναι μία **Διακριτή Κατανομή**, ενώ η  $f : \mathbb{X} \rightarrow [0, 1]$ , τέτοια ώστε  $f(x) := P(X = x)$  ορίζεται ως **Συνάρτηση Μάζας Πιθανότητας**.

Επίσης, το  $P$  θα είναι ένα διακριτό μέτρο και θα ισχύει ότι για κάθε  $h : \mathbb{R} \rightarrow [-\infty, +\infty]$ :

$$\int h(x) dP(x) = \sum_{x \in \mathbb{R}} h(x) P(\{x\}). \quad (1.36)$$

Τελικά, θα έχουμε και ότι

$$E_P[h(X)] = \sum_x h(x) f(x), \quad (1.37)$$

όταν τουλάχιστον ένα από τα δύο μέλη ορίζεται.

**Ορισμός 1.33.** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ ,  $X : \Omega \rightarrow \mathbb{R}$  τυχαία μεταβλητή και  $f : \mathbb{R} \rightarrow [0, \infty]$  η συνάρτηση πυκνότητας πιθανότητας της  $X$ . Ορίζουμε ως **Συνάρτηση Κατανομής Πιθανότητας** τη συνάρτηση  $F_x : \mathcal{B} \rightarrow [0, 1]$ , έτσι ώστε:

$$F_X(x) := P^X((-\infty, x]) = P(X \leq x) = \int_{-\infty}^x f(x) dx. \quad (1.38)$$

Το διακριτό ανάλογο (για διακριτό μέτρο πιθανότητας  $P$ ) μας δίνει ότι

$$F_X(x) := P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) \quad (1.39)$$

**Ορισμός 1.34.** Έστω γενικώς μία συνάρτηση  $f : E \rightarrow \mathbb{R}$ , όπου  $E$  ένα αυθαίρετο σύνολο. Ορίζουμε **Στήριγμα** της  $f$ , και θα το συμβολίζουμε ως  $\text{supp}(f)$ , το σύνολο με όλα τα στοιχεία του  $E$  για τα οποία η  $f$  είναι μη μηδενική:

$$\text{supp}(f) := \{x \in E \mid f(x) \neq 0\}. \quad (1.40)$$

## 1.2.4 Δεσμευμένη Μέση Τιμή

**Ορισμός 1.35.** Έστω χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ , ολοκληρώσιμη τυχαία μεταβλητή  $X : \Omega \rightarrow \mathbb{R}^n$  και  $\mathcal{H} \subseteq \mathcal{F}$  μία υπο-άλγεβρα της  $\mathcal{F}$ . Ορίζουμε **Δεσμευμένη Μέση Τιμή** της  $X$  δεδομένης της υπο-άλγεβρας  $\mathcal{H}$ , την οποία θα συμβολίζουμε με  $E[X | \mathcal{H}]$ , ως τη μοναδική  $\mathcal{H}$ -μετρήσιμη συνάρτηση  $\Omega \rightarrow \mathbb{R}^n$  που ικανοποιεί:

$$\int_{\mathcal{H}} E[X | \mathcal{H}] dP = \int_{\mathcal{H}} X dP, \quad (1.41)$$

για κάθε  $H \in \mathcal{H}$ .

Έστω επίσης ο μετρήσιμος χώρος  $(\mathbb{Y}, \mathcal{T})$  και η τυχαία μεταβλητή  $Y : \Omega \rightarrow \mathbb{Y}$ . Ορίζουμε **Δεσμευμένη Μέση Τιμή** της  $X$  δεδομένης της  $Y$ , την οποία θα συμβολίζουμε με  $E[X | Y]$ , ως τη μοναδική  $\mathcal{T}$ -μετρήσιμη συνάρτηση  $g : \mathbb{Y} \rightarrow \mathbb{R}^n$ , που ικανοποιεί:

$$E[g(Y)f(Y)] = E[Xf(Y)], \quad (1.42)$$

για κάθε  $\mathcal{T}$ -μετρήσιμη συνάρτηση  $f : \mathbb{Y} \rightarrow \mathbb{R}^n$ .

**Ιδιότητες** δεσμευμένης μέσης τιμής για  $X_1, X_2 : \Omega \rightarrow \mathbb{R}^n$ ,  $Y : \Omega \rightarrow \mathbb{Y}$ ,  $U : \Omega \rightarrow \mathbb{U}$  τυχαίες μεταβλητές:

1. **Γραμμικότητα:**  $E[\alpha X_1 + \beta X_2 | Y] = \alpha E[X_1 | Y] + \beta E[X_2 | Y]$ , για κάθε  $\alpha, \beta \in \mathbb{R}$ .
2. Αν  $X \geq 0$  σ.β., τότε  $E[X | Y] \geq 0$  σ.β.
3. **Τριγωνική Ανισότητα:**  $|E[X | Y]| \leq E[|X| | Y]$ .
4. **Μονοτονία:** Αν  $X_1 \leq X_2$  σ.β., τότε  $E[X_1 | Y] \leq E[X_2 | Y]$ , σ.β.
5.  $E[E[X | Y]] = E[X]$ .
6. Αν  $X, Y$  ανεξάρτητες, τότε  $E[X | Y] = E[X]$ .

7. Αν  $E[|f(Y)X|] < +\infty$ , τότε  $E[f(Y)X | Y] = f(Y)E[X | Y]$ .
8. **Ιδιότητα Πύργου:** Αν  $G(Y, U) = E[X | Y, U]$ , τότε  $E[G(Y, U) | Y] = E[X | Y] = g(Y)$ .

## 1.3 Χώροι $\mathcal{L}^p$ και Τρόποι Σύγκλισης Τυχαίων Μεταβλητών

### 1.3.1 Χώροι $\mathcal{L}^p$

**Ορισμός 1.36.** Έστω χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ ,  $X \in \mathcal{X}$ , με  $\mathcal{X} = \{X | X : \Omega \rightarrow [-\infty, +\infty]$  τυχαία μεταβλητή} και  $p \in [1, +\infty)$ .

Ορίζουμε

$$\|X\|_p := \{E[|X|^p]\}^{\frac{1}{p}} \quad (1.43)$$

και το σύνολο

$$\mathcal{L}^p(\Omega, \mathcal{F}, P) := \{X \in \mathcal{X} | \|X\|_p < +\infty\}. \quad (1.44)$$

το οποίο θα το συμβολίζουμε με σκέτο  $\mathcal{L}^p$  όταν εννοείται σε ποιόν χώρο πιθανότητας αναφερόμαστε. Επίσης, αποδεικνύεται πως η συνάρτηση  $\|\cdot\|_p : \mathcal{L}^p \rightarrow [0, +\infty)$  ικανοποιεί τις ιδιότητες:

1.  $\|\lambda X\|_p = |\lambda| \|X\|_p$  για κάθε  $\lambda \in \mathbb{R}$ .
2.  $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$ .

Συνεπώς, έπεται ότι ο  $\mathcal{L}^p$  είναι **Διανυσματικός Χώρος**.

**Ορισμός 1.37.** Παρατηρούμε ότι  $E[|X|^0] = 1 < +\infty$ , για κάθε  $X \in \mathcal{X}$ . Συνεπώς, μπορούμε στο ίδιο πλαίσιο να ορίσουμε  $\mathcal{L}^0 := \mathcal{X} = \{X | X : \Omega \rightarrow [-\infty, +\infty]$  τυχαία μεταβλητή}

**Ορισμός 1.38.** Θα λέμε ότι μια  $X \in \mathcal{L}^0$  είναι **φραγμένη με πιθανότητα 1** αν και μόνο αν  $\exists M \in \mathbb{R}$  ώστε  $P(|X| \leq M) = 1$ . Έπειτα, ορίζουμε

$$\|X\|_\infty := \inf\{M > 0 : P(|X| \leq M) = 1\} \quad (1.45)$$

και τέλος, το  $\mathcal{L}^\infty := \{X \in \mathcal{X} \mid \|X\|_\infty < +\infty\}$ .

Γενικώς, ισχύει ότι  $\mathcal{L}^\infty \subset \mathcal{L}^p \subset \mathcal{L}^0$ , για κάθε  $1 \leq p < +\infty$ , ενώ για  $p_1, p_2 \in [1, +\infty)$ , με  $p_1 < p_2$ , ισχύει ότι  $\mathcal{L}^{p_2} \subset \mathcal{L}^{p_1}$ .

**Ορισμός 1.39.** Έστω  $(\mathbb{X}, \mathcal{S}, \mu)$  χώρος μέτρου. Θα λέμε ότι μια συνάρτηση  $f : \mathbb{X} \rightarrow [-\infty, +\infty]$  είναι  $\mathcal{L}^p$  **Ολοκληρώσιμη**, αν και μόνο αν,  $f \in \mathcal{L}^p$ . Δηλαδή, η  $f$  είναι  $\mathcal{L}^p$  ολοκληρώσιμη, αν και μόνο αν,

$$\|f\|_p = \left( \int_{\mathbb{X}} |f|^p d\mu \right)^{\frac{1}{p}} < \infty. \quad (1.46)$$

**Ορισμός 1.40.** Έστω  $(\mathbb{X}, \mathcal{S}, \mu)$  χώρος θετικού μέτρου. Μία κλάση συναρτήσεων  $\Phi \subset \mathcal{L}^1$  θα λέγεται **Ομοιόμορφα Ολοκληρώσιμη**, αν και μόνο αν, για κάθε  $\epsilon > 0$  και  $f \in \Phi$ , υπάρχει  $\delta > 0$  που εξαρτάται μόνο από το  $\epsilon$  (δηλαδή  $\delta = \delta(\epsilon, f) = \delta(\epsilon)$ ), τέτοιο ώστε

$$\mu(E) < \delta \Rightarrow \int_E |f| d\mu < \epsilon. \quad (1.47)$$

### 1.3.2 Τρόποι Σύγκλισης Τυχαίων Μεταβλητών

**Ορισμός 1.41.** Έστω μια ακολουθία από τυχαίες μεταβλητές  $\{X_n\}_{n \in \mathbb{N}}$  ορισμένες στον χώρο πιθανότητας  $(\Omega, \mathcal{F}, P)$  και με τιμές στο  $[-\infty, +\infty]$  και τυχαία μεταβλητή  $X : \Omega \rightarrow [-\infty, +\infty]$ . Ορίζουμε τους εξής τρόπους σύγκλισης:

- Η  $\{X_n\}_{n \in \mathbb{N}}$  θα συγκλίνει στην  $X$  **σχεδόν βέβαια (ή με πιθανότητα 1)**, και θα συμβολίζεται με  $X_n \xrightarrow{\sigma.\beta.} X$ , αν και μόνο αν,

$$P(\lim_{n \rightarrow \infty} X_n = X) = 1, \quad (1.48)$$

δηλαδή

$$P\left(\left\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1. \quad (1.49)$$

- Για  $p \geq 1$  και  $X_n, X \in \mathcal{L}^p$ , για κάθε  $n \in \mathbb{N}$ , θα λέμε ότι η  $\{X_n\}_{n \in \mathbb{N}}$  συγκλίνει στη  $X$  στον  $\mathcal{L}^p$ , και θα συμβολίζεται με  $X_n \xrightarrow{\mathcal{L}^p} X$ , αν και μόνο αν,

$$\lim_{n \rightarrow \infty} E[|X_n - X|^p] = 0. \quad (1.50)$$

- Θα λέμε ότι η  $\{X_n\}_{n \in \mathbb{N}}$  συγκλίνει στη  $X$  κατά πιθανότητα, και θα συμβολίζεται με  $X_n \xrightarrow{P} X$ , αν και μόνο αν,

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0. \quad (1.51)$$

**Ορισμός 1.42.** Έστω  $\mu, \{\mu_n\}_{n \in \mathbb{N}}$ , μέτρα πιθανότητας στον  $\mathbb{R}$ . Θα λέμε ότι η  $\{\mu_n\}_{n \in \mathbb{N}}$  συγκλίνει ασθενώς στο  $\mu$ , και θα το συμβολίζουμε με  $\mu_n \Rightarrow \mu$ , αν και μόνο αν:

$$\mu_n((-\infty, x]) \rightarrow \mu((-\infty, x]), \quad (1.52)$$

καθώς  $n \rightarrow \infty$  για κάθε  $x \in \mathbb{R}$ , τέτοιο ώστε  $\mu(\{x\}) = 0$ . Δηλαδή, για κάθε σημείο συνέχειας του  $\mu$ .

**Ορισμός 1.43.** Έστω μια ακολουθία από τυχαίες μεταβλητές  $\{X_n\}_{n \in \mathbb{N}}$  ορισμένες για κάθε  $n \in \mathbb{N}$  στον χώρο πιθανότητας  $(\Omega_n, \mathcal{F}_n, P_n)$  και με τιμές στον  $\mathbb{R}$ . Επίσης, έστω τυχαία μεταβλητή  $X : \Omega \rightarrow \mathbb{R}$ . Θα λέμε ότι η  $\{X_n\}_{n \in \mathbb{N}}$

συγκλίνει στη  $X$  **κατά κατανομή** (ή **κατά νόμο** ή **ασθενώς**), και θα συμβολίζεται με  $X_n \xrightarrow{d} X$ , αν και μόνο αν,

$$P^{X_n} \Rightarrow P^X, \quad (1.53)$$

όπου  $\{P^{X_n}\}_{n \in \mathbb{N}}$  η ακολουθία κατανομών των  $X_n$  και  $P^X$  η κατανομή της  $X$ .

Δηλαδή, αν και μόνο αν,

$$F_{X_n} \Rightarrow F_X. \quad (1.54)$$

## 1.4 Χρήσιμα Αποτελέσματα & Εργαλεία

### 1.4.1 Παραδείγματα διακριτών Κατανομών

**Διακριτή Ομοιόμορφη Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί τη διακριτή Ομοιόμορφη στο  $K = \{\alpha, \alpha + 1, \alpha + 2, \dots, \beta - 2, \beta - 1, \beta\}$  και θα το συμβολίζουμε με  $X \sim \mathcal{U}\{\alpha, \beta\}$ , με  $n = \beta - \alpha + 1$ , όταν η  $X$  μπορεί να πάρει οποιαδήποτε τιμή στο  $K$  με πιθανότητα  $P(X = k) = \frac{1}{n}$  για κάθε  $k \in K$ . Για παράδειγμα, η  $X$  μπορεί να είναι το αποτέλεσμα μιας ρίψης ενός δίκαιου ζαριού.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \mathcal{U}\{\alpha, \beta\}$ ,  $\alpha, \beta \in \mathbb{Z}$ ,  $\beta \geq \alpha$ .
- Στήριγμα:  $k \in K = \{\alpha, \alpha + 1, \alpha + 2, \dots, \beta - 2, \beta - 1, \beta\}$
- Συνάρτηση Μάζας Πιθανότητας:  $P(X = k) = \frac{1}{n}$ , για κάθε  $k \in K$ .
- Μέση Τιμή:  $E[X] = \frac{\alpha + \beta}{2}$ .
- Διασπορά:  $\text{Var}[X] = \frac{(\beta - \alpha + 1)^2 - 1}{12}$ .



**Bernoulli Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί την κατανομή Bernoulli με παράμετρο  $p$  και θα το συμβολίζουμε με  $X \sim \text{Bernoulli}(p)$ , όταν η  $X$  αποτελεί το αποτέλεσμα ενός πειράματος με δύο πιθανές εκβάσεις, έστω 1 και 0, και αντίστοιχες πιθανότητες  $p$  και  $q = 1 - p$ . Για παράδειγμα, η μπορεί να  $X$  είναι το αποτέλεσμα μιας ρίψης ενός δίκαιου νομίσματος.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \text{Bernoulli}(p)$ ,  $p \in [0, 1]$ .
- Στήριγμα:  $k \in \{0, 1\}$ .
- Συνάρτηση Μάζας Πιθανότητας:  $P(X = 1) = p$ ,  $P(X = 0) = q$ .
- Μέση Τιμή:  $E[X] = p$ .
- Διασπορά:  $\text{Var}[X] = pq = p(1 - p)$ .

**Διωνυμική Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί τη Διωνυμική κατανομή (ή Binomial) με παραμέτρους  $n, p$  και θα το συμβολίζουμε με  $X \sim \text{Bin}(n, p)$ , όταν η  $X$  αποτελεί τον αριθμό των επιτυχιών σε  $n$  ανεξάρτητα πειράματα Bernoulli με παράμετρο  $p$ . Για παράδειγμα, η  $X$  μπορεί να είναι ο αριθμός των επιτυχιών σε  $n$  ρίψεις ενός δίκαιου νομίσματος.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \text{Bin}(n, p)$ ,  $n \in \mathbb{N}_0$ ,  $p \in [0, 1]$ .
- Στήριγμα:  $k \in \mathbb{N}_0$ .
- Συνάρτηση Μάζας Πιθανότητας:  $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ .
- Μέση Τιμή:  $E[X] = np$ .
- Διασπορά:  $\text{Var}[X] = np(1 - p)$ .

**Γεωμετρική Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί τη Γεωμετρική κατανομή όταν εκφράζει το πλήθος των ανεξάρτητων Bernoulli δοκιμών με παράμετρο  $p$  που χρειάζονται μέχρι την πρώτη επιτυχία και θα το συμβολίζουμε με

με  $X \sim \text{Geo}(p)$ . Για παράδειγμα, η  $X$  μπορεί να είναι οι βολές ενός παίχτη στο μπάσκετ, που γενικώς σκοράρει με πιθανότητα  $p$ , μέχρι να πετύχει την πρώτη.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \text{Geo}(p)$ ,  $p \in (0, 1]$ .
- Στήριγμα:  $k \in \mathbb{N}$ .
- Συνάρτηση Μάζας Πιθανότητας:  $P(X = k) = (1 - p)^{k-1}p$ .
- Μέση Τιμή:  $E[X] = \frac{1}{p}$ .
- Διασπορά:  $\text{Var}[X] = \frac{1-p}{p^2}$ .

**Poisson Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί την κατανομή **Poisson** με παράμετρο  $\lambda$  και θα το συμβολίζουμε με  $X \sim \text{Pois}(\lambda)$ , όταν η  $X$  εκφράζει τον αριθμό συγκεκριμένων και ανεξάρτητων συμβάντων σε ένα προκαθορισμένο χρονικό ή χωρικό διάστημα, με γνωστή μέση τιμή. Για παράδειγμα, η  $X$  μπορεί να είναι ο αριθμός των κλήσεων που λαμβάνει ένα τηλεφωνικό κέντρο κάθε μία ώρα, δεδομένου ότι η μέση τιμή είναι 14 κλήσεις την ώρα.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \text{Pois}(\lambda)$ ,  $\lambda \in \mathbb{R}^+ = [0, \infty)$ .
- Στήριγμα:  $k \in \mathbb{N}_0$ .
- Συνάρτηση Μάζας Πιθανότητας:  $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ .
- Μέση Τιμή:  $E[X] = \lambda$ .
- Διασπορά:  $\text{Var}[X] = \lambda$ .

## 1.4.2 Παραδείγματα Συνεχών Κατανομών

**Συνεχής Ομοιόμορφη Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί τη συνεχή Ομοιόμορφη κατανομή στο διάστημα  $K = [\alpha, \beta]$  (ή  $(\alpha, \beta)$  ή  $[\alpha, \beta)$  ή  $(\alpha, \beta]$ ) και θα το συμβολίζουμε με  $X \sim U(\alpha, \beta)$ , όταν η  $X$  μπορεί να ανήκει σε οποιαδήποτε διάστημα ίδιου μήκους στο  $K$  με ίδια πιθανότητα. Για παράδειγμα, η  $X$  μπορεί να είναι η γωνία που σχηματίζει ο μεγάλος δείκτης του ρολογιού ( $K = [0, 2\pi)$ ) όταν ξυπνήσεις μετά από έναν βαθύ ύπνο.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim U(\alpha, \beta)$ ,  $\alpha, \beta \in \mathbb{R}$ ,  $\beta > \alpha$ .
- Στήριγμα:  $x \in [\alpha, \beta]$ .
- Συνάρτηση Πυκνότητας Πιθανότητας:

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, & x \in [\alpha, \beta] \\ 0, & \text{αλλιώς.} \end{cases}$$

- Συνάρτηση Κατανομής Πιθανότητας:

$$F(x) = \begin{cases} 0, & x < \alpha \\ \frac{x - \alpha}{\beta - \alpha}, & x \in [\alpha, \beta] \\ 1, & x > \beta. \end{cases}$$

- Μέση Τιμή:  $E[X] = \frac{\alpha + \beta}{2}$ .
- Διασπορά:  $\text{Var}[X] = \frac{(\beta - \alpha)^2}{12}$ .

**Κανονική Κατανομή:** Η Κανονική κατανομή (ή κατανομή Gauss) είναι η κατανομή που εμφανίζεται πιο συχνά στη φύση. Έχει το σχήμα της καμπάνας και είναι η πρώτη πρόβλεψη που κάνει κάποιος όταν αναρωτιέται τι κατανομή ακολουθεί ένα δείγμα με πραγματικές και συνεχείς τιμές, για το οποίο δεν έχει καμμία άλλη πληροφορία. Ο λόγος για αυτό, είναι κυρίως το Κεντρικό

Οριακό Θεώρημα, το οποίο το παραθέτουμε παρακάτω, στην υποεπινότητα με τα Οριακά Θεωρήματα. Θα λέμε ότι η  $X$  ακολουθεί Κανονική κατανομή με παραμέτρους  $\mu$  και  $\sigma^2$ , όπου  $\mu = E[X]$  και  $\sigma^2 = \text{Var}[X]$ , και θα το συμβολίζουμε με  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Για παράδειγμα, η  $X$  μπορεί να είναι το ύψος των φοιτητριών σε ένα πανεπιστήμιο.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \mathcal{N}(\mu, \sigma^2)$ ,  $\mu \in \mathbb{R}$ ,  $\sigma^2 > 0$ .
- Στήριγμα:  $x \in \mathbb{R}$ .
- Συνάρτηση Πυκνότητας Πιθανότητας:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \quad (1.55)$$

- Συνάρτηση Κατανομής Πιθανότητας:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt. \quad (1.56)$$

- Μέση Τιμή:  $E[X] = \mu$ .
- Διασπορά:  $\text{Var}[X] = \sigma^2$ .

**Τυποποιημένη Κανονική Κατανομή:** Η Τυποποιημένη Κανονική κατανομή είναι η ειδική περίπτωση της κανονικής, όπου  $\mu = 0$  και  $\sigma^2 = 1$ . Δηλαδή, πρόκειται για την ίδια κατανομή, κεντραρισμένη γύρω από το 0 και με διασπορά ίση με τη μονάδα.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \mathcal{N}(0, 1)$ .
- Στήριγμα:  $x \in \mathbb{R}$ .
- Συνάρτηση Πυκνότητας Πιθανότητας:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \quad (1.57)$$

- Συνάρτηση Κατανομής Πιθανότητας:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt. \quad (1.58)$$

- Μέση Τιμή:  $E[X] = 0$ .
- Διασπορά:  $\text{Var}[X] = 1$ .

**Γάμμα Κατανομή:** Η κατανομή Γάμμα είναι μια σημαντική οικογένεια διπα-  
ραμετρικών συνεχών κατανομών που βασίζεται στη συνάρτηση Γάμμα, η οποία  
υπενθυμίζουμε ότι στους πραγματικούς αριθμούς έχει μορφή:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad (1.59)$$

ενώ ορίζονται και τα ατελή τμήματά της  $\Gamma(x, s) = \int_s^{\infty} t^{x-1} e^{-t} dt$  (άνω) και  
 $\gamma(x, s) = \int_0^s t^{x-1} e^{-t} dt$  (κάτω), έτσι ώστε  $\Gamma(x) = \gamma(x, s) + \Gamma(x, s)$ .

Η κατανομή Γάμμα δέχεται ως όρισμα δύο θετικές παραμέτρους,  $\alpha, \beta$ , που εκ-  
φράζουν το σχήμα της κατανομής και τον ρυθμό, αντίστοιχα, και όταν η  $X$   
ακολουθεί την κατανομή Γάμμα θα το συμβολίζουμε ως  $X \sim \text{Gamma}(\alpha, \beta)$ .  
Μια εφαρμογή της εν λόγω κατανομής θα ήταν για να εκτιμήσουμε την πι-  
θανότητα να πρέπει να περιμένουμε 2 με 4 ώρες ώστε να πιάσουμε 4 ψάρια,  
δεδομένου ότι πιάνουμε 1 ανά μισή ώρα.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \text{Gamma}(\alpha, \beta)$ ,  $\alpha, \beta > 0$ .
- Στήριγμα:  $x \in (0, \infty)$ .
- Συνάρτηση Πυκνότητας Πιθανότητας:

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

- Συνάρτηση Κατανομής Πιθανότητας:

$$F(x) = \begin{cases} \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \beta x), & x > 0 \\ 0, & x \leq 0. \end{cases}$$

- Μέση Τιμή:  $E[X] = \frac{\alpha}{\beta}$ .
- Διασπορά:  $\text{Var}[X] = \frac{\alpha}{\beta^2}$ .

**Εκθετική Κατανομή:** Θα λέμε ότι η  $X$  ακολουθεί την Εκθετική κατανομή με παράμετρο  $\lambda$  όταν εκφράζει τον χρόνο που μεσολαβεί ανάμεσα σε συγκεκριμένα γεγονότα που συμβαίνουν συνεχώς και ανεξάρτητα, με σταθερό ρυθμό  $\lambda$  (διαδικασίες Poisson) και θα το συμβολίζουμε με  $X \sim \text{Exp}(\lambda)$ . Είναι ειδική περίπτωση της κατανομής Γάμμα και το συνεχές ανάλογο της Γεωμετρικής κατανομής. Για παράδειγμα, η  $X$  μπορεί να είναι ο χρόνος που μεσολαβεί μέχρι την επόμενη κλήση που θα λάβουμε.

Χαρακτηριστικές ποσότητες της κατανομής:

- $X \sim \text{Exp}(\lambda)$ ,  $\lambda > 0$ .
- Στήριγμα:  $x \in [0, \infty)$ .
- Συνάρτηση Πυκνότητας Πιθανότητας:

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

- Συνάρτηση Κατανομής Πιθανότητας:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

- Μέση Τιμή:  $E[X] = \frac{1}{\lambda}$ .
- Διασπορά:  $\text{Var}[X] = \frac{1}{\lambda^2}$ .

### 1.4.3 Διάσημες Ανισότητες

**Πρόταση 1.44. Ανισότητα Hölder:** Έστω  $p, q \in (1, \infty)$  με  $\frac{1}{p} + \frac{1}{q} = 1$  και  $X, Y : \Omega \rightarrow [-\infty, +\infty]$  τυχαίες μεταβλητές με  $\|X\|_p < \infty$ ,  $\|Y\|_q < \infty$ . Τότε,  $XY \in \mathcal{L}_1$  και

$$\|XY\|_1 = |E[XY]| \leq \|X\|_p \|Y\|_q. \quad (1.60)$$

Ως συνέπεια της πρότασης, για  $X = |X|^r$ ,  $Y = 1$ ,  $p = \frac{s}{r}$  και  $q = \frac{s}{s-r}$  με  $1 \leq r < s$ , έπεται ότι

$$\|X\|_r \leq \|X\|_s. \quad (1.61)$$

Επίσης, ειδικότερα για  $p = q = 2$ , η ανισότητα είναι γνωστή ως Cauchy-Schwarz.

**Πρόταση 1.45. Ανισότητα Markov:** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ , τυχαία μεταβλητή  $X : \Omega \rightarrow [0, +\infty]$  και  $\alpha > 0$ , τότε:

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha}. \quad (1.62)$$

**Πρόταση 1.46. Ανισότητα Chebyshev:** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ , τυχαία μεταβλητή  $X : \Omega \rightarrow [-\infty, +\infty]$  με  $E[X] < \infty$  και  $\alpha > 0$ , τότε:

$$P(|X - E[X]| \geq \alpha) \leq \frac{\text{Var}[X]}{\alpha^2}. \quad (1.63)$$

**Πρόταση 1.47. Ανισότητα Jensen:** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$ , τυχαία μεταβλητή  $X : \Omega \rightarrow \mathbb{R}$  με  $E[X] < \infty$  και  $\Phi : I \rightarrow \mathbb{R}$  κυρτή συνάρτηση στο διάστημα  $I \subset \mathbb{R}$ , με  $P(\{X \in I\}) = 1$  και  $E[|\Phi(X)|] < \infty$ . Τότε:

$$\Phi(E[X]) \leq E[\Phi(X)]. \quad (1.64)$$

**Πρόταση 1.48. Τύπος Ολικής Πιθανότητας:** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$  και  $\{B_n : n = 1, 2, 3, \dots\}$  μια διαμέριση του  $\Omega$ , έτσι ώστε για κάθε  $n$  το  $B_n$  να είναι  $\mathcal{F}$ -μετρήσιμο. Τότε, για οποιοδήποτε μετρήσιμο  $A \subset \Omega$ , θα ισχύει ότι:

$$P(A) = \sum_n P(A \cap B_n) = \sum_n P(A|B_n)P(B_n). \quad (1.65)$$

**Λήμμα 1.49. Πρώτο Λήμμα Borel-Cantelli:** Έστω  $(\Omega, \mathcal{F}, P)$  χώρος πιθανότητας και  $\{A_n\}_{n \in \mathbb{N}}$  ακολουθία στοιχείων της  $\mathcal{F}$ . Αν  $\sum_{n=1}^{\infty} P(A_n) < \infty$ , τότε

$$P(\limsup_{n \geq 1} A_n) = 0. \quad (1.66)$$

**Λήμμα 1.50. Δεύτερο Λήμμα Borel-Cantelli:** Έστω  $(\Omega, \mathcal{F}, P)$  χώρος πιθανότητας και  $\{A_n\}_{n \in \mathbb{N}}$  ακολουθία ανεξάρτητων στοιχείων της  $\mathcal{F}$ . Αν  $\sum_{n=1}^{\infty} P(A_n) = \infty$ , τότε

$$P(\limsup_{n \geq 1} A_n) = 1. \quad (1.67)$$

**Θεώρημα 1.51. Θεώρημα Tonelli:** Έστω  $(X, \mathcal{S}, \mu)$ ,  $(Y, \mathcal{T}, \nu)$  χώροι πεπερασμένου μέτρου,  $(X \times Y, \mathcal{S} \otimes \mathcal{T}, \mu \otimes \nu)$  ο χώρος γινόμενο και  $f : X \times Y \rightarrow [0, \infty]$  μετρήσιμη συνάρτηση. Τότε, οι συναρτήσεις

$$x \mapsto \int f(x, y) d\nu(y), \quad x \mapsto \int f(x, y) d\mu(x), \quad (1.68)$$

είναι  $\mathcal{S}/\mathcal{B}([-\infty, +\infty])$ ,  $\mathcal{T}/\mathcal{B}([-\infty, +\infty])$  μετρήσιμες, αντίστοιχα, και



$$\int f(x, y) d(\mu \otimes \nu)(x, y) = \int \left( \int f(x, y) d\nu(y) \right) d\mu(x) = \int \left( \int f(x, y) d\mu(x) \right) d\nu(y). \quad (1.69)$$

**Θεώρημα 1.52. Θεώρημα Fubini:** Έστω  $(\mathbb{X}, \mathcal{S}, \mu)$ ,  $(\mathbb{Y}, \mathcal{T}, \nu)$  χώροι πεπερασμένου μέτρου,  $(\mathbb{X} \times \mathbb{Y}, \mathcal{S} \otimes \mathcal{T}, \mu \otimes \nu)$  ο χώρος γινόμενο και  $f: \mathbb{X} \times \mathbb{Y} \rightarrow [-\infty, +\infty]$  μετρήσιμη συνάρτηση. Αν επιπλέον έχουμε ότι  $\int |f(x, y)| d(\mu \otimes \nu)(x, y) < \infty$ , τότε θα ισχύουν οι προϋποθέσεις του Θεωρήματος Tonelli και άρα θα έπεται ότι:

$$\int |f(x, y)| d(\mu \otimes \nu)(x, y) = \int \left( \int |f(x, y)| d\nu(y) \right) d\mu(x) = \int \left( \int |f(x, y)| d\mu(x) \right) d\nu(y). \quad (1.70)$$

#### 1.4.4 Θεωρήματα Σύγκλισης

**Θεώρημα 1.53.** Έστω  $X, \{X_n\}_{n \in \mathbb{N}}$  τυχαίες μεταβλητές στον  $(\Omega, \mathcal{F}, P)$ , με τιμές στο  $[-\infty, +\infty]$  και  $p \geq 1$ . Τότε, θα ισχύουν τα εξής:

- Αν  $X_n \xrightarrow{\mathcal{L}^p} X$ , τότε  $X_n \xrightarrow{P} X$ .
- Αν  $X_n \xrightarrow{\sigma.\beta.} X$ , τότε  $X_n \xrightarrow{P} X$ .

Αν οι  $X, \{X_n\}_{n \in \mathbb{N}}$  παίρνουν τιμές στον  $\mathbb{R}$ , επιπλέον θα ισχύει ότι:

- Αν  $X_n \xrightarrow{P} X$ , τότε  $X_n \xrightarrow{d} X$ .

**Θεώρημα 1.54.** Έστω  $\{X_n\}_{n \in \mathbb{N}}$  ακολουθία τυχαίων μεταβλητών και  $X$  τυχαία μεταβλητή έτσι ώστε  $X_n \xrightarrow{P} X$ . Τότε, υπάρχει υπακολουθία  $\{X_{n_k}\}_{k \in \mathbb{N}}$ , έτσι ώστε  $X_{n_k} \xrightarrow{\sigma.\beta.} X$ .

**Θεώρημα 1.55.** Έστω  $\{X_n\}_{n \in \mathbb{N}}$  ακολουθία τυχαίων μεταβλητών και  $X$  τυχαία μεταβλητή έτσι ώστε  $X_n \xrightarrow{P} X$ . Έστω επίσης ότι υπάρχει  $Y \in \mathcal{L}^p$ , τέτοια ώστε  $|X_n| < Y$  για κάθε  $n \in \mathbb{N}$ . Τότε, θα ισχύει ότι  $X \in \mathcal{L}^p$  και  $X_n \xrightarrow{\mathcal{L}^p} X$ .

**Θεώρημα 1.56.** Έστω  $X, \{X_n\}_{n \in \mathbb{N}}$  τυχαίες μεταβλητές και  $f : \mathbb{R} \rightarrow \mathbb{R}$  συνεχής συνάρτηση. Τότε, θα ισχύουν τα εξής:

- Αν  $X_n \xrightarrow{P} X$ , τότε  $f(X_n) \xrightarrow{P} f(X)$ .
- Αν  $X_n \xrightarrow{\sigma.\beta.} X$ , τότε  $f(X_n) \xrightarrow{\sigma.\beta.} f(X)$ .

**Θεώρημα 1.57.** Έστω  $X, \{X_n\}_{n \in \mathbb{N}}$  τυχαίες μεταβλητές στον  $\mathbb{R}$ . Τότε,  $X_n \xrightarrow{d} X$ , αν και μόνο αν,

$$E[f(X_n)] \rightarrow E[f(X)], \quad (1.71)$$

για κάθε  $f : \mathbb{R} \rightarrow \mathbb{R}$  συνεχή και φραγμένη.

### 1.4.5 Θεωρήματα Σύγκλισης Μέσης Τιμής

**Λήμμα 1.58. Λήμμα Fatou:** Έστω χώρος μέτρου  $(X, \mathcal{S}, \mu)$  και  $\{f_n\}_{n \in \mathbb{N}}$  μια ακολουθία μετρήσιμων συναρτήσεων με  $f_n : X \rightarrow [0, +\infty]$ . Τότε,

$$\liminf_{n \rightarrow \infty} \int f_n d\mu \leq \int \liminf_{n \rightarrow \infty} f_n d\mu \quad (1.72)$$

**Θεώρημα 1.59. Θεώρημα Μονότονης Σύγκλισης:** Έστω χώρος μέτρου  $(X, \mathcal{S}, \mu)$  και  $\{f_n\}_{n \in \mathbb{N}}$  μια αύξουσα ακολουθία μετρήσιμων συναρτήσεων, έτσι ώστε  $f_n : X \rightarrow [0, +\infty]$  για κάθε  $n \in \mathbb{N}$ . Τότε θα ισχύει ότι:

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int \lim_{n \rightarrow \infty} f_n d\mu = \int f d\mu, \quad (1.73)$$

όπου τα όρια  $f = \lim_{n \rightarrow \infty} f_n$  και  $\lim_{n \rightarrow \infty} \int f_n d\mu$  υπάρχουν γιατί οι αντίστοιχες ακολουθίες είναι αύξουσες.

**Θεώρημα 1.60. Θεώρημα Κυριαρχημένης Σύγκλισης:** Έστω χώρος μέτρου  $(X, \mathcal{S}, \mu)$  και  $\{f_n\}_{n \in \mathbb{N}}$  μια ακολουθία μετρήσιμων συναρτήσεων, έτσι ώστε  $f_n : X \rightarrow \mathbb{R}$  για κάθε  $n \in \mathbb{N}$ , με  $f = \lim_{n \rightarrow \infty} f_n$ . Επίσης, έστω μετρήσιμη συνάρτηση  $g : X \rightarrow [0, \infty]$  με  $\int g d\mu < \infty$ . Τότε, αν επιπλέον για κάθε  $n \in \mathbb{N}$  ισχύει ότι  $|f_n| \leq g$  σχεδόν παντού, θα έχουμε ότι  $\int |f| d\mu < \infty$  και

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int \lim_{n \rightarrow \infty} f_n d\mu = \int f d\mu. \quad (1.74)$$

**Θεώρημα 1.61. Θεώρημα Φραγμένης Σύγκλισης:** Έστω χώρος πεπερασμένου μέτρου  $(X, \mathcal{S}, \mu)$  και  $\{f_n\}_{n \in \mathbb{N}}$  μια ακολουθία μετρήσιμων συναρτήσεων, έτσι ώστε  $f_n : X \rightarrow \mathbb{R}$  για κάθε  $n \in \mathbb{N}$ , με  $f = \lim_{n \rightarrow \infty} f_n$ . Επίσης, έστω σταθερά  $M \in \mathbb{R}^+ = [0, \infty)$ . Τότε, αν επιπλέον για κάθε  $n \in \mathbb{N}$  ισχύει ότι  $|f_n| \leq M$  σχεδόν παντού, θα έχουμε ότι  $\int |f| d\mu < \infty$  και

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int \lim_{n \rightarrow \infty} f_n d\mu = \int f d\mu. \quad (1.75)$$

Το πιθανοθεωρητικό ανάλογο στην περίπτωση που έχουμε χώρο πιθανότητας  $(\Omega, \mathcal{F}, P)$  (χώρο πεπερασμένου μέτρου), για τα παραπάνω θεωρήματα σύγκλισης, μας δίνει αποτελέσματα για τη μέση τιμή μιας τυχαίας μεταβλητής. Δηλαδή,

$$\lim_{n \rightarrow \infty} E[X_n] = E[\lim_{n \rightarrow \infty} X_n]. \quad (1.76)$$

**Θεώρημα 1.62. Θεώρημα Beppo-Levi:** Έστω χώρος μέτρου  $(X, \mathcal{S}, \mu)$  και  $\{f_n\}_{n \in \mathbb{N}}$  μια ακολουθία μετρήσιμων συναρτήσεων, έτσι ώστε  $f_n : X \rightarrow [0, +\infty]$  για κάθε  $n \in \mathbb{N}$ . Τότε, θα ισχύει ότι

$$\int \left( \sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int f_n d\mu. \quad (1.77)$$

**Λήμμα 1.63. Λήμμα Scheffé:** Έστω χώρος μέτρου  $(X, \mathcal{S}, \mu)$  και  $\{f_n\}_{n \in \mathbb{N}}$  μια ακολουθία ολοκληρώσιμων (κατά Lebesgue) συναρτήσεων που συγκλίνουν σχεδόν παντού σε μια επίσης ολοκληρώσιμη συνάρτηση  $f$ . Τότε θα ισχύει ότι:

$$\int |f_n - f| d\mu \rightarrow 0 \iff \int |f_n| d\mu \rightarrow \int |f| d\mu. \quad (1.78)$$

#### 1.4.6 Οριακά Θεωρήματα

**Θεώρημα 1.64. Ισχυρός Νόμος των Μεγάλων Αριθμών:** Έστω  $\{X_n\}_{n \in \mathbb{N}}$  ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών (δηλαδή οι  $X_n$  είναι ανεξάρτητες ανά δύο και όλες ακολουθούν την ίδια κατανομή) στον  $(\Omega, \mathcal{F}, P)$ , με τιμές στον  $\mathbb{R}$  έτσι ώστε  $E[X_1^2] < \infty$ . Θέτουμε  $E[X_1] = \mu$  και  $S_n = \sum_{k=1}^n X_k$  για κάθε  $n \in \mathbb{N}$ . Τότε,

$$\frac{S_n}{n} \xrightarrow{\text{σ.β.}} \mu. \quad (1.79)$$

**Θεώρημα 1.65. Ασθενής Νόμος των Μεγάλων Αριθμών:** Έστω  $\{X_n\}_{n \in \mathbb{N}}$  ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών (δηλαδή οι  $X_n$  είναι ανεξάρτητες ανά δύο και όλες ακολουθούν την ίδια κατανομή) στον  $(\Omega, \mathcal{F}, P)$ , με τιμές στον  $\mathbb{R}$ . Θέτουμε  $E[X_1] = \mu$  και  $S_n = \sum_{k=1}^n X_k$  για κάθε  $n \in \mathbb{N}$ . Τότε,

$$\frac{S_n}{n} \xrightarrow{P} \mu. \quad (1.80)$$

**Θεώρημα 1.66. Κεντρικό Οριακό Θεώρημα:** Έστω  $\{X_n\}_{n \in \mathbb{N}}$  ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, έτσι ώστε  $E[X_1] = \mu$  και  $\text{Var}[X_1] = \sigma^2 \in (0, \infty)$ . Θέτουμε  $S_n = \sum_{k=1}^n X_k$  για κάθε  $n \in \mathbb{N}$ . Τότε,

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z, Z \sim N(0, 1). \quad (1.81)$$

# Ενότητα 2

## Στοχαστικές Διαδικασίες

### 2.1 Γενικές Βασικές Έννοιες

Στην Ενότητα 2, θα παραθέσουμε ένα σημαντικό μέρος της θεωρίας των στοχαστικών διαδικασιών, όπως επίσης και διάφορα αποτελέσματα, τα οποία θα χρησιμεύσουν στη ροή της εργασίας. Για περισσότερη μελέτη των στοχαστικών διαδικασιών, παραπέμπουμε τον αναγνώστη στα [6], [7], [8], [9], [10], [11], [12] και [13].

#### 2.1.1 Έννοιες & Ορισμοί

**Ορισμός 2.1.** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, \mathbb{P})$ , ο μετρήσιμος χώρος  $(\mathbb{X}, \mathcal{S})$  και ένα σύνολο δεικτών  $T$ . Η οικογένεια τυχαίων μεταβλητών  $\{X_t\}_{t \in T}$  θα ονομάζεται **Στοχαστική Διαδικασία** ορισμένη στον  $T$  και με τιμές στον  $\mathbb{X}$ , αν και μόνον αν,

$$\forall t \in T, \quad X_t : \Omega \rightarrow \mathbb{X}. \quad (2.1)$$

Το σύνολο δεικτών  $T$  συνήθως αντιπροσωπεύει χρονικές στιγμές ή επαναλήψεις σε ένα πείραμα. Συνεπώς, αν το  $T$  είναι αριθμήσιμο (π.χ.  $T = \mathbb{N}_0 = \{0, 1, 2, \dots\}$ ), θα λέμε ότι η στοχαστική διαδικασία  $\{X_t\}_{t \in T}$  είναι μία **Στοχαστική Διαδικασία Διακριτού Χρόνου**. Αντίστοιχα, αν το  $T$  είναι υπεραριθμήσιμο (π.χ.  $T = [0, +\infty)$ ), θα λέμε ότι έχουμε μία **Στοχαστική Διαδικασία**

## Συνεχούς Χρόνου.

Επίσης, από εδώ και στο εξής, ο  $\mathbb{X}$  θα λέγεται **Χώρος Καταστάσεων** της στοχαστικής διαδικασίας.

**Ορισμός 2.2.** Έστω ο χώρος πιθανότητας  $(\Omega, \mathcal{F}, P)$  και ένα σύνολο δεικτών  $T$ . Επίσης, έστω ότι για κάθε  $t \in T$ , η  $\mathcal{F}_t$  είναι μία υπο-σ-άλγεβρα της  $\mathcal{F}$ . Αν ισχύει πως  $\mathcal{F}_k \subseteq \mathcal{F}_l \subseteq \mathcal{F}$  για κάθε  $k \leq l$ , θα λέμε ότι η οικογένεια  $(\mathcal{F}_t)_{t \in T}$  αποτελεί μία **διήθηση** (ή **φιλτράρισμα**) του χώρου πιθανότητας  $(\Omega, \mathcal{F}, P)$ .

**Ορισμός 2.3.** Η τετράδα  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, P)$  θα λέγεται **Χώρος με Διήθηση** ή **Διηθημένος Χώρος Πιθανότητας**.

**Ορισμός 2.4.** Έστω ο διηθημένος χώρος πιθανότητας  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, P)$  και η στοχαστική διαδικασία  $\{X_t\}_{t \in T}$ . Η διαδικασία θα λέγεται **προσαρμοσμένη** στη διήθηση  $(\mathcal{F}_t)_{t \in T}$ , αν και μόνο αν, η τυχαία μεταβλητή  $X_t$  είναι  $\mathcal{F}_t$ -μετρήσιμη για κάθε  $t \in T$ . Πολλές φορές μια  $\mathcal{F}_t$ -προσαρμοσμένη διαδικασία  $\{X_t\}_{t \in T}$  θα τη συμβολίζουμε ως  $(X_t, \mathcal{F}_t)_{t \in T}$ .

**Ορισμός 2.5.** Δύο στοχαστικές διαδικασίες, έστω  $\{X_t\}_{t \in T}$  και  $\{Y_t\}_{t \in T}$ , θα λέγονται **στοχαστικά μη διακριτές**, αν και μόνο αν, για κάθε  $t \in T$ ,

$$X_t(\omega) = Y_t(\omega), \text{ σ.β.} \quad (2.2)$$

**Ορισμός 2.6.** Πολλές φορές, μας διευκολύνει το να «βλέπουμε» μια στοχαστική διαδικασία  $\{X_t\}_{t \in T}$ , με τιμές στον  $(\mathbb{X}, \mathcal{S})$ , ως ένα τυχαίο στοιχείο, έστω  $\mathbf{X}$ , με τιμές στον  $\mathbb{X}^T$ , όπου  $T = [0, \infty)$ . Υπό αυτήν την οπτική, ορίζουμε ως **Τροχιά** της  $\{X_t\}_{t \in T}$  τη συνάρτηση  $t \mapsto X_t(\omega)$ , η οποία θα είναι στοιχείο του  $\mathbb{X}^T$  για κάθε  $\omega \in \Omega$ .

## 2.1.2 Συνέχεια & Ολοκληρωσιμότητα

**Ορισμός 2.7.** Αν για κάθε  $\omega \in \Omega$  οι τροχιές της  $\{X_t\}_{t \in T}$  είναι συνεχείς συναρτήσεις του  $t$  στο  $[0, \infty)$ , θα λέμε ότι η στοχαστική διαδικασία  $\mathbf{X} = \{X_t\}_{t \in T}$  έχει **Συνεχείς Τροχιές**. Για συντομία, ενδέχεται να το συμβολίζουμε ως  $\mathbf{X} \in C$ , όπου  $C$  είναι ο χώρος των συνεχών συναρτήσεων.

**Ορισμός 2.8.** Έστω  $f : [0, +\infty) \rightarrow \mathbb{R}$ . Η  $f$  θα λέγεται **Cádlág**, αν και μόνο αν, για κάθε  $t \in [0, +\infty)$ :

- Το αριστερό όριο  $f(t^-) := \lim_{s \uparrow t} f(s)$  υπάρχει.
- Το δεξί όριο  $f(t^+) := \lim_{s \downarrow t} f(s)$  υπάρχει και ισούται με  $f(t)$ .

Με άλλα λόγια, η  $f$  είναι δεξιά συνεχής και έχει αριστερό όριο.

**Ορισμός 2.9.** Αν για κάθε  $\omega \in \Omega$  οι τροχιές της  $\{X_t\}_{t \in T}$  είναι **cádlág** συναρτήσεις του  $t$  στο  $[0, \infty)$ , θα λέμε ότι η στοχαστική διαδικασία  $\mathbf{X} = \{X_t\}_{t \in T}$  έχει **Cádlág Τροχιές**. Για συντομία, ενδέχεται να το συμβολίζουμε ως  $\mathbf{X} \in D$ , όπου  $D$  είναι ο χώρος των **cádlág** συναρτήσεων.

**Ορισμός 2.10.** Έστω ο διηθημένος χώρος πιθανότητας  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, \mathbb{P})$  και η στοχαστική διαδικασία  $\{X_t\}_{t \in T}$ . Θα δώσουμε δύο διαφορετικούς ορισμούς για μια διαδικασία που θα λέγεται **προβλέψιμη**, ανάλογα αν η  $\{X_t\}_{t \in T}$  είναι διακριτού ή συνεχούς χρόνου.

- Στην περίπτωση που  $T = \mathbb{N}_0$ , η διαδικασία  $\{X_n\}_{n \in \mathbb{N}_0}$  θα λέγεται **προβλέψιμη**, αν και μόνο αν, η τυχαία μεταβλητή  $X_{n+1}$  είναι  $\mathcal{F}_n$ -μετρήσιμη για κάθε  $n \in \mathbb{N}_0$ .
- Στην περίπτωση που  $T = [0, \infty)$ , η διαδικασία  $\{X_t\}_{t \geq 0}$  θα λέγεται **προβλέψιμη**, αν και μόνο αν, είναι  $\mathcal{P}$ -μετρήσιμη, όπου  $\mathcal{P}$  είναι η  $\sigma$ -άλγεβρα στον  $\Omega \times [0, \infty)$  που παράγεται από όλες τις αριστερά συνεχείς (cág) προσαρμοσμένες διαδικασίες  $Y$ .



**Ορισμός 2.11.** Έστω  $(\Omega, \mathcal{F}, P)$  χώρος πιθανότητας και  $T$  ένα συνεχές σύνολο χρόνων, έστω  $[0, +\infty)$ . Επίσης, έστω η στοχαστική διαδικασία  $\{X_t\}_{t \in T}$ , ορισμένη στον  $\Omega \times T$  και με τιμές στον  $\mathbb{R}$ . Ορίζουμε τις εξής μορφές συνέχειας για τη στοχαστική διαδικασία  $\{X_t\}_{t \in T}$ :

- Έστω  $t \in T$ . Η  $\{X_t\}_{t \in T}$  θα λέγεται **Συνεχής με Πιθανότητα 1** (ή **Συνεχής Σχεδόν Βεβαίως**) στο  $t$ , αν και μόνο αν,

$$P\left(\left\{\omega \in \Omega \mid \lim_{s \rightarrow t} |X_s(\omega) - X_t(\omega)| = 0\right\}\right) = 1. \quad (2.3)$$

- Έστω  $t \in T$ . Η  $\{X_t\}_{t \in T}$  θα λέγεται **Συνεχής κατά Τετραγωνικό Μέσο** στο  $t$ , αν και μόνο αν,  $E[|X_t|^2] < +\infty$  και

$$\lim_{s \rightarrow t} E\left[|X_s - X_t|^2\right] = 0. \quad (2.4)$$

- Έστω  $t \in T$ . Η  $\{X_t\}_{t \in T}$  θα λέγεται **Συνεχής κατά Πιθανότητα** στο  $t$ , αν και μόνο αν, για κάθε  $\epsilon > 0$ ,

$$\lim_{s \rightarrow t} P\left(\left\{\omega \in \Omega \mid |X_s(\omega) - X_t(\omega)| \geq \epsilon\right\}\right) = 0. \quad (2.5)$$

Ισοδύναμα, Η  $\{X_t\}_{t \in T}$  θα λέγεται συνεχής κατά Πιθανότητα στο  $t$ , αν και μόνο αν,

$$\lim_{s \rightarrow t} E\left[\frac{|X_s - X_t|}{1 + |X_s - X_t|}\right] = 0. \quad (2.6)$$

- Έστω  $t \in T$ . Η  $\{X_t\}_{t \in T}$  θα λέγεται **Συνεχής κατά Κατανομή** στο  $t$ , αν και μόνο αν,

$$\lim_{s \rightarrow t} F_s(x) = F_t(x), \quad (2.7)$$

για κάθε  $x$  για το οποίο η  $F_t$  είναι συνεχής, όπου  $F_t$  είναι η συνάρτηση κατανομής πιθανότητας της τυχαίας μεταβλητής  $X_t$ .

**Ορισμός 2.12.** Μία στοχαστική διαδικασία συνεχούς χρόνου  $\{X_t\}_{t \in T}$  θα λέγεται **Ολοκληρώσιμη**, αν και μόνο αν, η τυχαία μεταβλητή  $X_t$  είναι ολοκληρώσιμη για κάθε  $t \in T$ . Αντίστοιχα, μια διαδικασία θα λέγεται  **$\mathcal{L}^p$ -Ολοκληρώσιμη**, αν για κάθε  $t \in T$  η τυχαία μεταβλητή  $X_t$  είναι  $\mathcal{L}^p$ -ολοκληρώσιμη.

**Ορισμός 2.13.** Μία διαδικασία που είναι  $\mathcal{L}^\infty$ -ολοκληρώσιμη θα λέγεται **Ομοιόμορφα Φραγμένη**.

**Ορισμός 2.14.** Έστω  $\Delta = \{0 = t_0 < t_1 < \dots < t_k = t\}$  μία διαμέριση του διαστήματος  $[0, t]$ . Για μια διαδικασία  $\mathbf{X} = \{X_t\}_{t \in T}$  ορίζουμε

$$S_t^\Delta = \sum_i |X_{t_{i+1}} - X_{t_i}|. \quad (2.8)$$

Θα λέμε ότι η  $\mathbf{X}$  είναι **πεπερασμένης μεταβολής** αν για κάθε  $t$

$$S_t = \sup_{\Delta} S_t^\Delta < \infty. \quad (2.9)$$

Επίσης, αν  $\lim_{t \rightarrow \infty} S_t < \infty$ , τότε θα λέμε ότι η  $\mathbf{X}$  είναι **φραγμένης μεταβολής**. Την κλάση με τις διαδικασίες φραγμένης μεταβολής θα τη συμβολίζουμε με  $\mathcal{V}$ . Εναλλακτικά, μπορούμε να σκεφτούμε ότι  $\mathbf{X} = \{X_t\}_{t \in T} \in \mathcal{V}$ , αν και μόνο αν,  $\int_0^t |dX_t(\omega)| < \infty$  για  $t > 0$  και  $\omega \in \Omega$ .

**Ορισμός 2.15.** Έστω  $\{\Delta_n\}_{n \geq 1}$  μία ακολουθία διαμερίσεων του διαστήματος  $[0, t]$ , της μορφής  $\{0 = t_0 < t_1 < \dots < t_k = t\}$ , με  $\|\Delta_n\| \rightarrow 0$ , όπου  $\|\Delta\| = \sup_i |t_i - t_{i-1}|$  είναι η **λεπτότητα** της διαμέρισης. Για μια τετραγωνικά ολοκληρώσιμη διαδικασία  $\mathbf{X} = \{X_t\}_{t \in T}$  ορίζουμε την **Τετραγωνική Μεταβολή** της, που θα τη συμβολίζουμε με  $[X]_t$ , ως το όριο κατά πιθανότητα:

$$[X]_t = \lim_{\|\Delta_n\| \rightarrow 0} \sum_i (X_{t_i} - X_{t_{i-1}})^2, \quad (2.10)$$

αν υπάρχει το όριο και είναι ίδιο για κάθε ακολουθία διαμερίσεων  $\{\Delta_n\}_{n \geq 1}$ , με  $\|\Delta_n\| \rightarrow 0$ .

### 2.1.3 Διαδικασίες Gauss, Wiener και Lévy

**Ορισμός 2.16. Τυχαίο Περίπατο** ονομάζουμε μια διαδικασία που σε κάθε βήμα κάνει μία τυχαία κίνηση προς μία κατεύθυνση σε έναν χώρο καταστάσεων. Ένα κλασικό παράδειγμα είναι ο τυχαίος περίπατος στους ακεραίους  $\mathbb{Z}$  που ξεκινάει από το 0 και σε κάθε βήμα πηγαίνει μία θέση δεξιά (+1) με πιθανότητα  $p$  ή μία θέση αριστερά (-1) με πιθανότητα  $q = 1 - p$ . Λόγω της «απλότητάς» του στο να κάνει αποκλειστικά βήματα μίας θέσης, ο περίπατος λέγεται **Απλός Τυχαίος Περίπατος**. Ένα ακόμα πιο κλασικό παράδειγμα είναι ο απλός τυχαίος περίπατος στους ακεραίους με «συμμετρική» πιθανότητα  $p = q = \frac{1}{2}$ . Σε αυτήν την περίπτωση, τον περίπατο τον λέμε **Απλό Συμμετρικό Τυχαίο Περίπατο**.

**Ορισμός 2.17.** Μία στοχαστική διαδικασία συνεχούς χρόνου  $\{X_t\}_{t \in T}$  θα λέγεται **Gauss**, αν και μόνο αν, για κάθε πεπερασμένο σύνολο δεικτών  $t_1, t_2, \dots, t_k \in T$ , η τυχαία μεταβλητή  $X_{t_1, t_2, \dots, t_k} = (X_{t_1}, X_{t_2}, \dots, X_{t_k})$  ακολουθεί την **Πολυδιάστατη Κανονική Κατανομή**. Δηλαδή, αν και μόνο αν, κάθε γραμμικός συνδυασμός των  $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$  ακολουθεί Κανονική κατανομή.

**Ορισμός 2.18.** Ορίζουμε ως **Τυπική Διαδικασία Wiener** (ή **Τυπική Κίνηση Brown**) τη στοχαστική διαδικασία συνεχούς χρόνου με πραγματικές τιμές,  $\{W_t\}_{t \in T}$ , για την οποία ισχύουν:

1.  $W_0 = 0$ .
2. Έχει ανεξάρτητες προσαυξήσεις. Δηλαδή, για κάθε  $t > 0$ , οι μελλοντικές προσαυξήσεις  $W_{t+u} - W_t$ ,  $u \geq 0$ , είναι ανεξάρτητες από τις παρελθοντικές τιμές  $W_s$ ,  $s \leq t$ .

3. Οι προσαυξήσεις της ακολουθούν κανονική κατανομή. Πιο συγκεκριμένα, για κάθε  $t \in T$  η προσαύξηση  $W_{t+u} - W_t$  ακολουθεί κανονική κατανομή με μέση τιμή 0 και διασπορά  $u$ ,  $u \in T$ . Δηλαδή,  $W_{t+u} - W_t \sim \mathcal{N}(0, u)$ .
4. Η  $W_t$  έχει συνεχείς τροχιές. Δηλαδή, για κάθε  $\omega \in \Omega$  η συνάρτηση  $t \mapsto W_t(\omega)$  είναι συνεχής συνάρτηση του  $t$  στο  $[0, +\infty)$ .

Ουσιαστικά, η τυπική κίνηση Brown είναι ένας τυχαίος περίπατος που ξεκινά από το 0 και κινείται στον χώρο των πραγματικών αριθμών  $\mathbb{R}$ , που είναι ένας συνεχής χώρος καταστάσεων. Δηλαδή, όταν το μήκος του βήματος ενός περιπάτου τείνει στο 0, μιλάμε για μία κίνηση Brown.

**Ορισμός 2.19.** Μία στοχαστική διαδικασία  $\{X_t\}_{t \in T}$  θα λέγεται Lévy, αν και μόνο αν, ικανοποιεί τις εξής συνθήκες:

1.  $X_0 = 0$ .
2. Έχει ανεξάρτητες προσαυξήσεις.
3. Οι προσαυξήσεις της είναι στάσιμες. Πιο συγκεκριμένα, για κάθε  $s < t \in T$  η προσαύξηση  $X_t - X_s$  ακολουθεί την ίδια κατανομή με την τυχαία μεταβλητή  $X_{t-s}$ .
4. Είναι συνεχής κατά πιθανότητα, έτσι ώστε για κάθε  $\epsilon > 0$  και  $t \geq 0$ , να ισχύει ότι  $\lim_{h \rightarrow 0} P(|X_{t+h} - X_t| > \epsilon) = 0$ .

## 2.2 Αλυσίδες Markov

### 2.2.1 Μαρκοβιανή Ιδιότητα

**Ορισμός 2.20.** Έστω  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, P)$  χώρος με διήθηση και  $\{X_n\}_{n \in \mathbb{N}_0}$  στοχαστική διαδικασία με τιμές στον μετρήσιμο και αριθμήσιμο χώρο καταστάσεων  $(X, \mathcal{S})$ . Θα λέμε ότι η  $X = (X_n, \mathcal{F}_n)_{n \geq 0}$  είναι μια **Αλυσίδα Markov** (ή **Διαδικασία Markov Διακριτού Χρόνου**), αν και μόνο αν,

1. Η τυχαία μεταβλητή  $X_n$  είναι  $\mathcal{F}_n$ -μετρήσιμη για κάθε  $n \in \mathbb{N}_0$ .
2.  $P(X_{n+1} \in B \mid \mathcal{F}_n)(\omega) = P(X_{n+1} \in B \mid X_n)(\omega)$ , P-σ.β., για κάθε  $n \geq 0$  και  $B \in \mathcal{S}$ .

Ουσιαστικά, η συνθήκη 2, γνωστή και ως **Μαρκοβιανή Ιδιότητα**, μας λέει ότι η πιθανότητα να βρεθούμε στο  $B$  στο βήμα  $n + 1$ , δεδομένης της προϊστορίας  $\mathcal{F}_n$ , είναι η ίδια με την πιθανότητα να βρεθούμε στο  $B$  τη χρονική στιγμή  $n + 1$ , γνωρίζοντας μόνο τη θέση της αλυσίδας τη χρονική στιγμή  $n$ . Με άλλα λόγια, η πιθανότητα να βρεθεί η αλυσίδα σε οποιαδήποτε θέση στο βήμα  $n + 1$  εξαρτάται μόνο από τη θέση της αλυσίδας στο βήμα  $n$  και καθόλου από το πώς έφτασε ως εκεί.

Με το συμβολισμό  $P(X_{n+1} \in B \mid X_n(\omega))$  θα εννοούμε  $P(X_{n+1} \in B \mid X_n)(\omega)$ .

**Ορισμός 2.21.** Από εδώ και στο εξής, θα θεωρούμε ότι ο  $\mathbb{X}$  είναι ένας χώρος Borel. Υπό αυτήν την προϋπόθεση, γνωρίζουμε ότι υπάρχει μια δεσμευμένη κατανομή  $P_n(x, B)$ , τέτοια ώστε (P-σ.β.):

$$P_n(x, B) = P(X_n \in B \mid X_{n-1}(\omega) = x), \quad B \in \mathcal{S}, \quad n \geq 1. \quad (2.11)$$

Οι πιθανότητες  $P_n(X_{n-1}(\omega), B)$  θα λέγονται **Πιθανότητες Μετάβασης** της μαρκοβιανής αλυσίδας.

Πιο απλά, μπορούμε να γράφουμε απλώς  $P_n(x, y)$  και θα εννοούμε

$$P_n(x, y) = P(X_n = y \mid X_{n-1} = x), \quad \forall x, y \in \mathbb{X}. \quad (2.12)$$

**Ορισμός 2.22.** Θα λέμε ότι μια αλυσίδα Markov είναι **Χρονικά Ομοιογενής**, αν και μόνο αν, οι πιθανότητες μετάβασής της δεν εξαρτώνται από το  $n$ , δηλαδή αν και μόνο αν

$$P_n(x, B) = P(x, B), \quad \forall n \geq 1. \quad (2.13)$$

**Ορισμός 2.23.** Για να ολοκληρωθεί η περιγραφή μιας μαρκοβιανής αλυσίδας, χρειάζεται και η **Αρχική Κατανομή** της, έστω  $\pi_0 = \pi_0(\mathbf{B})$ ,  $\mathbf{B} \in \mathcal{S}$ :

$$\pi_0(\mathbf{B}) = P(X_0 \in \mathbf{B}) = P(X_0 = x) = \pi_0(x). \quad (2.14)$$

Από εδώ και στο εξής, θα μιλάμε για  $(\pi_0, P_1, P_2, \dots)$ -μαρκοβιανές αλυσίδες ή  $(\pi_0, P)$ -μαρκοβιανές αλυσίδες σε περίπτωση χρονικής ομοιογένειας.

**Ορισμός 2.24.** Στην περίπτωση που ο  $\mathbb{X}$  είναι πεπερασμένος, έτσι ώστε  $\mathbb{X} = \{x_1, x_2, \dots, x_N\}$  και η αλυσίδα χρονικά ομοιογενής, ορίζουμε τον **Πίνακα Πιθανοτήτων Μετάβασης  $\mathbf{P}$**  έτσι ώστε  $\mathbf{P}_{i,j} = P(x_i, x_j) = P(X_n = x_j \mid X_{n-1} = x_i)$ .

**Ορισμός 2.25.** Αντίστοιχα, με  $\mathbf{P}^n$  θα ορίζουμε τον **Πίνακα Πιθανοτήτων Μετάβασης  $n$ -οστής τάξης**, έτσι ώστε  $\mathbf{P}_{i,j}^n = P^{(n)}(x_i, x_j)$ , όπου

$$P^{(n)}(x_i, x_j) = P(X_{n+k} = x_j \mid X_k = x_i), \quad \forall k \in \mathbb{N}_0, x_i, x_j \in \mathbb{X}, \quad (2.15)$$

οι **Πιθανότητες Μετάβασης  $n$ -οστής τάξης**.

**Ορισμός 2.26.** Από τους παραπάνω ορισμούς, έπεται ότι η **συνάρτηση μάζας πιθανότητας (κατανομή της αλυσίδας) τη χρονική στιγμή  $n$** , θα ορίζεται αναδρομικά, έτσι ώστε για κάθε  $y \in \mathbb{X}$ :

$$\begin{aligned} \pi_n(y) &= \sum_{x \in \mathbb{X}} P(X_{n-1} = x) P(X_n = y \mid X_{n-1} = x) \\ &= \sum_{x \in \mathbb{X}} \pi_{n-1}(x) P(x, y) \\ &= (\pi_{n-1} \mathbf{P})(y). \end{aligned} \quad (2.16)$$

Για κάθε  $n \in \mathbb{N}_0$ , η  $\pi_n$  θα μπορεί να εκφραστεί ως ένα διάνυσμα του  $[0, 1]^N$ , έτσι ώστε

$$\pi_n = (\pi_n(x_1), \pi_n(x_2), \dots, \pi_n(x_N)). \quad (2.17)$$

Τελικά, επαγωγικά προκύπτει ότι  $\pi_n = \pi_0 \mathbf{P}^n$  για κάθε  $n \in \mathbb{N}$ .

**Ορισμός 2.27.** Στάσιμη Κατανομή (ή Αναλλοίωτη Κατανομή) της αλυσίδας  $\{X_n\}_{n \in \mathbb{N}_0}$  με πίνακα μετάβασης  $\mathbf{P}$  θα λέμε την κατανομή  $\pi$  για την οποία ισχύει:

$$\pi = \pi \mathbf{P} \quad (2.18)$$

ή

$$\pi(y) = \sum_{x \in \mathbb{X}} \pi(x) p(x, y), \quad \forall y \in \mathbb{X}. \quad (2.19)$$

## 2.2.2 Χρόνοι Διακοπής

**Ορισμός 2.28.** Θα λέμε ότι ένα ενδεχόμενο  $A$  ανήκει στην  $\mathcal{F}_n$ , αν και μόνο αν, η δείκτρια συνάρτησή του είναι μια συνάρτηση των  $X_0, X_1, \dots, X_n$  και μόνο αυτών. Δηλαδή, αν και μόνο αν, για κάθε  $\omega \in \Omega$ :

$$\mathbb{1}_A(\omega) = f(X_0(\omega), X_1(\omega), \dots, X_n(\omega)). \quad (2.20)$$

**Ορισμός 2.29.** Έστω τυχαία μεταβλητή  $\tau: \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ . Η  $\tau$  θα λέγεται **Χρόνος Διακοπής** της  $\{X_n\}_{n \in \mathbb{N}_0}$ , αν και μόνο αν, για κάθε  $n \in \mathbb{N}_0$ , το ενδεχόμενο  $\{\tau = n\}$  ανήκει στην  $\mathcal{F}_n$ .

**Λήμμα 2.30.** Η τυχαία μεταβλητή  $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$  είναι χρόνος διακοπής, αν και μόνο αν, για κάθε  $n \in \mathbb{N}_0$ , το ενδεχόμενο  $\{\tau \leq n\}$  ανήκει στην  $\mathcal{F}_n$ .

**Πόρισμα 2.31.** Αν  $\tau_1, \tau_2$  είναι χρόνοι διακοπής, τότε οι  $\tau_1 \vee \tau_2$  και  $\tau_1 \wedge \tau_2$  θα είναι χρόνοι διακοπής, όπου:

$$(\tau_1 \vee \tau_2)(\omega) = \max\{\tau_1(\omega), \tau_2(\omega)\}, \quad (\tau_1 \wedge \tau_2)(\omega) = \min\{\tau_1(\omega), \tau_2(\omega)\}. \quad (2.21)$$

**Ορισμός 2.32.** Αν  $\tau$  είναι ένας χρόνος διακοπής, θα λέμε ότι το ενδεχόμενο  $A$  ανήκει στο παρελθόν του  $\tau$ , και θα το συμβολίζουμε  $A \in \mathcal{F}_\tau$ , αν και μόνο αν, για κάθε  $n \in \mathbb{N}_0$  το ενδεχόμενο  $A \cap \{\tau = n\}$  ανήκει στην  $\mathcal{F}_n$ .

**Θεώρημα 2.33. Ισχυρή Μαρκοβιανή Ιδιότητα:** Έστω  $\{X_n\}_{n \in \mathbb{N}_0}$  μαρκοβιανή αλυσίδα στον αριθμήσιμο χώρο καταστάσεων  $\mathbb{X}$ , με πίνακα πιθανοτήτων μετάβασης  $P$ . Αν ο  $\tau$  είναι ένας χρόνος διακοπής της  $\{X_n\}_{n \in \mathbb{N}_0}$  με  $\tau < \infty$  και  $X_\tau = x \in \mathbb{X}$ , τότε η στοχαστική διαδικασία  $\{Y_n\}_{n \in \mathbb{N}_0}$  με  $Y_n = X_{\tau+n}$  είναι μια μαρκοβιανή αλυσίδα που ξεκινά από τη  $x$ , έχει και αυτή πίνακα μετάβασης  $P$  και είναι ανεξάρτητη από κάθε ενδεχόμενο  $A \in \mathcal{F}_\tau$ .

**Ορισμός 2.34.** Έστω στοχαστική διαδικασία  $\mathbf{X} = \{X_t\}_{t \geq 0}$  και  $T$  ένας χρόνος διακοπής. Ορίζουμε τη σταματημένη διαδικασία  $\mathbf{X}^\tau = \{X_{t \wedge \tau}\}_{t \geq 0}$ , όπου  $X_{t \wedge \tau} = X_{t \wedge \tau(\omega)}(\omega)$ . Άμεσα έπεται ότι  $\mathbf{X}^\tau = \mathbf{X}$  στο σύνολο  $\{\omega : \tau(\omega) = \infty\}$ .

### 2.2.3 Χαρακτηρισμοί Αλυσίδας

**Ορισμός 2.35.** Θα λέμε ότι μία κατάσταση  $y \in \mathbb{X}$  είναι προσβάσιμη από την κατάσταση  $x \in \mathbb{X}$ , αν και μόνο αν, υπάρχει διαδρομή από τη  $x$  στην  $y$  με θετική πιθανότητα. Δηλαδή, αν και μόνο αν, υπάρχει  $n \in \mathbb{N}_0$  τέτοιο ώστε  $P^{(n)}(x, y) > 0$  και θα το συμβολίζουμε  $x \rightarrow y$ .



**Ορισμός 2.36.** Θα λέμε ότι οι καταστάσεις  $x, y$  **επικοινωνούν**, και θα το συμβολίζουμε  $x \leftrightarrow y$ , αν και μόνο αν,  $x \rightarrow y$  και  $y \rightarrow x$ .

**Πρόταση 2.37.** Η σχέση  $\leftrightarrow$  είναι μια σχέση ισοδυναμίας, δηλαδή είναι:

1. Αυτοπαθή:  $x \leftrightarrow x$ , για κάθε  $x \in \mathbb{X}$ .
2. Συμμετρική:  $x \leftrightarrow y \Leftrightarrow y \leftrightarrow x$ , για κάθε  $x, y \in \mathbb{X}$ .
3. Μεταβατική:  $(x \leftrightarrow u \text{ και } u \leftrightarrow y) \Rightarrow x \leftrightarrow y$ , για κάθε  $x, y, u \in \mathbb{R}$ .

Ως εκ τούτου, η  $\leftrightarrow$  διαμερίζει τον χώρο σε κλάσεις  $\mathcal{C}$ , που τις ονομάζουμε **Κλάσεις Επικοινωνίας**.

**Ορισμός 2.38.** Μια κλάση επικοινωνίας  $\mathcal{C}$  θα λέγεται **ανοιχτή**, αν υπάρχουν καταστάσεις  $x \in \mathcal{C}$  και  $y \notin \mathcal{C}$ , τέτοιες ώστε  $x \rightarrow y$ . Αν μία κλάση δεν είναι ανοιχτή, τότε θα λέγεται **κλειστή**.

**Ορισμός 2.39.** Μία μαρκοβιανή αλυσίδα θα λέγεται **μη υποβιβάσιμη**, αν ολόκληρος ο χώρος είναι μία κλάση, η οποία τότε θα είναι αναγκαστικά κλειστή.

**Ορισμός 2.40.** Έστω μαρκοβιανή αλυσίδα  $\{X_n\}_{n \in \mathbb{N}_0}$  με τιμές σε έναν αριθμήσιμο χώρο καταστάσεων  $\mathbb{X}$ . Για κάθε  $x \in \mathbb{X}$ , ορίζουμε το συνολικό πλήθος των επισκέψεων της αλυσίδας στην κατάσταση  $x$  ως

$$V(X) = \sum_{k=0}^{\infty} \mathbb{1}\{X_k = x\}. \quad (2.22)$$

Αντίστοιχα, ορίζουμε το πλήθος των επισκέψεων στην κατάσταση  $x$ , πριν τη χρονική στιγμή  $n$  ως

$$V_n(X) = \sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\}. \quad (2.23)$$

**Ορισμός 2.41.** Μία κατάσταση  $x \in \mathbb{X}$  θα λέγεται **επαναληπτική** ή **παροδική**, αν και μόνο αν:

- $x$  επαναληπτική  $\Leftrightarrow P_x(V(x) < \infty) = 0$ .
- $x$  παροδική  $\Leftrightarrow P_x(V(x) < \infty) = 1$ .

Με τον συμβολισμό  $P_x$  εννοούμε ότι η αλυσίδα μας ξεκινά από την κατάσταση  $x$ .

**Ορισμός 2.42.** Ορίζουμε **Χρόνο Πρώτης Επιστροφής** στη  $x$  ως  $\tau_x^+ = \inf\{k > 0 : X_k = x\}$ . Επίσης, ορίζουμε αναδρομικά τους χρόνους της  $n$ -οστής επιστροφής της αλυσίδας  $\{X_n\}_{n \in \mathbb{N}_0}$  στην κατάσταση  $x$  ως εξής:

$$\tau_x^{+,0} = 0, \quad \tau_0^{+,n+1} = \inf\{k > \tau_x^{+,n} : X_k = x\}, \quad n \in \mathbb{N}_0. \quad (2.24)$$

Για τη συνέπεια των ορισμών μας, εννοείται ότι  $\tau_x^+ = \tau_x^{+,1}$ .

**Θεώρημα 2.43.** Αν  $P_x(\tau_x^+ < \infty) = 1$ , η κατάσταση  $x$  θα είναι επαναληπτική. Αν  $P_x(\tau_x^+ < \infty) < 1$ , η  $x$  θα είναι παροδική. Ειδικότερα, κάθε κατάσταση  $x \in \mathbb{X}$  θα είναι επαναληπτική ή παροδική.

**Λήμμα 2.44.** Η επαναληπτικότητα και η παροδικότητα είναι ιδιότητες κλάσης. Δηλαδή, αν  $x$  επαναληπτική/παροδική, τότε κάθε κατάσταση  $y$  που επικοινωνεί με τη  $x$  θα είναι αντίστοιχα επαναληπτική/παροδική.

**Λήμμα 2.45.** Αν μια κλάση είναι  $\mathcal{C}$  είναι παροδική, τότε για κάθε  $x, y \in \mathcal{C}$  θα ισχύει ότι  $P_x(V(y) < \infty) = 1$ . Αντίστοιχα, αν μια κλάση είναι  $\mathcal{C}$  είναι παροδική, τότε για κάθε  $x, y \in \mathcal{C}$  θα ισχύει ότι  $P_x(V(y) < \infty) = 0$ .

**Ορισμός 2.46.** Μία κατάσταση  $x \in \mathbb{X}$  θα λέγεται **γνησίως επαναληπτική**, αν και μόνο αν,  $E_x[\tau_x^+] < +\infty$ .

Η γνήσια επαναληπτικότητα είναι μια έννοια ισχυρότερη της απλής επαναληπτικότητας, αφού αν  $E_x[\tau_x^+] < +\infty$ , τότε αναγκαστικά θα έχουμε και ότι  $P_x(\tau_x^+ < \infty) = 1$ . Επίσης, είναι κι αυτή μια ιδιότητα κλάσης.

**Πόρισμα 2.47.** Η γνήσια επαναληπτικότητα είναι ιδιότητα κλάσης.

**Πόρισμα 2.48.** Μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα σε έναν πεπερασμένο χώρο καταστάσεων θα είναι γνησίως επαναληπτική.

**Θεώρημα 2.49.** Έστω  $x \in \mathbb{X}$  μια γνησίως επαναληπτική κατάσταση. Ορίζουμε

$$\pi_x(y) = \frac{1}{E_x[\tau_x^+]} E_x \left[ \sum_{k=1}^{\tau_x^+} \mathbf{1}\{X_k = y\} \right]. \quad (2.25)$$

Τότε, η  $\pi_x$  θα είναι μια αναλλοίωτη κατανομή της αλυσίδας  $\{X_n\}_{n \in \mathbb{N}_0}$  και

$$\pi_x(x) = \frac{1}{E_x[\tau_x^+]}. \quad (2.26)$$

**Ορισμός 2.50.** Για μια κατάσταση  $x \in \mathbb{X}$  ορίζουμε το **Σύνολο των Δυνατών Χρόνων Επιστροφής** στη  $x$  ως

$$R(x) = \{n \in \mathbb{N} : P^{(n)}(x, x) > 0\}. \quad (2.27)$$

Θέτουμε ως  $d(x)$  τον μέγιστο κοινό διαιρέτη του  $R(x)$  και τον ορίζουμε ως **Περίοδο** της κατάστασης  $x$ .

**Ορισμός 2.51.** Μία κατάσταση  $x \in \mathbb{X}$  θα λέγεται **απεριοδική**, αν και μόνο αν,  $d(x) = 1$ .

**Θεώρημα 2.52.** Η περιοδικότητα είναι ιδιότητα κλάσης.

**Πόρισμα 2.53.** Αν μια μη υποβιβάσιμη αλυσίδα έχει μια κατάσταση  $x$  για την οποία  $P(x, x) > 0$ , τότε η αλυσίδα είναι απεριοδική.

## 2.2.4 Ασυμπτωτική Κατανομή & Εργοδικό Θεώρημα

**Θεώρημα 2.54.** Έστω  $\{X_n\}_{n \in \mathbb{N}_0}$  μια μη υποβιβάσιμη, γνησίως επαναληπτική και απεριοδική μαρκοβιανή αλυσίδα στον χώρο καταστάσεων  $\mathbb{X}$ . Αν  $\pi_n(x) = P(X_n = x)$ , για  $x \in \mathbb{X}$ , είναι η κατανομή της αλυσίδας μετά από  $n \in \mathbb{N}_0$  βήματα και  $\pi$  είναι η αναλλοίωτη κατανομή της αλυσίδας, τότε θα έχουμε ότι:

$$\lim_{n \rightarrow \infty} \pi_n = \pi. \quad (2.28)$$

Και ειδικότερα,

$$\lim_{n \rightarrow \infty} P^{(n)}(x, y) = \pi(y), \quad \forall x, y \in \mathbb{X}. \quad (2.29)$$

**Θεώρημα 2.55. Εργοδικό Θεώρημα:** Έστω  $\{X_n\}_{n \in \mathbb{N}_0}$  μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα σε έναν διακριτό χώρο καταστάσεων  $\mathbb{X}$ . Για οποιαδήποτε αρχική κατανομή της αλυσίδας, θα έχουμε ότι:

$$P\left(\frac{V_n(x)}{n} \rightarrow \frac{1}{E_x[\tau_x^+]}, \quad \forall x \in \mathbb{X}\right) = 1 \quad (2.30)$$

Αν επιπλέον η αλυσίδα είναι και μη υποβιβάσιμη, τότε θα ισχύει ότι  $\pi(x) = \frac{1}{E_x[\tau_x^+]}$ , για κάθε  $x \in \mathbb{X}$  και άρα σε αυτήν την περίπτωση το Εργοδικό Θεώρημα μας λέει ότι:

$$P\left(\frac{V_n(x)}{n} \rightarrow \pi(x), \forall x \in \mathbb{X}\right) = 1 \quad (2.31)$$

**Θεώρημα 2.56. Εργοδικό Θεώρημα II:** Έστω  $\{X_n\}_{n \in \mathbb{N}_0}$  μια μη υποβιβάσιμη, γνησίως επαναληπτική μαρκοβιανή αλυσίδα σε έναν διακριτό χώρο καταστάσεων  $\mathbb{X}$ . Αν  $f: \mathbb{X} \rightarrow \mathbb{R}$  είναι μια φραγμένη συνάρτηση, τότε για οποιαδήποτε αρχική κατανομή της αλυσίδας έχουμε:

$$P\left(\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \rightarrow E^\pi[f]\right) = 1, \quad (2.32)$$

όπου  $E^\pi[f] = \sum_{x \in \mathbb{X}} f(x)\pi(x)$  είναι η μέση τιμή της τυχαίας μεταβλητής  $f(X)$ , αν η τυχαία μεταβλητή ακολουθεί την αναλλοίωτη κατανομή της αλυσίδας  $\pi$ .

## 2.3 Διαδικασίες Martingale

### 2.3.1 Ορισμοί & Ειδικές Περιπτώσεις

**Ορισμός 2.57.** Έστω  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$  διηθημένος χώρος πιθανότητας,  $T = [0, +\infty)$  και  $\mathbf{X} = \{X_t\}_{t \in T}$  στοχαστική διαδικασία προσαρμοσμένη στην  $\mathcal{F}_t$ . Τότε, η διαδικασία θα λέγεται **Martingale**, αν και μόνο αν,  $\mathbf{X} \in D$  (cádlág) και

1.  $E[|X_t|] < \infty$ , για  $t \geq 0$ .
2.  $E[X_t | \mathcal{F}_s] = X_s$ , για  $s \geq t$ .

Η συνθήκη 2 λέγεται και ιδιότητα **Martingale**. Επίσης, ορίζουμε τις διαδικασίες **Submartingale** και **Supermartingale**, αλλάζοντας την ιδιότητα **Martingale** σε  $E[X_t | \mathcal{F}_s] \geq X_s$  και  $E[X_t | \mathcal{F}_s] \leq X_s$ , αντίστοιχα.

**Ορισμός 2.58.** Έστω  $\mathcal{X}$  μία κλάση στοχαστικών διαδικασιών. Θα λέμε ότι η διαδικασία  $\mathbf{X}$  ανήκει στην **Τοπική Κλάση**  $\mathcal{X}_{loc}$ , αν και μόνο αν, υπάρχει αύξουσα ακολουθία χρόνων διακοπής  $\{\tau_n\}_{n \geq 1}$  της  $\mathbf{X}$ , τέτοια ώστε  $\lim_{n \rightarrow \infty} \tau_n = \infty$ , σ.β. και κάθε σταματημένη διαδικασία  $\mathbf{X}^{\tau_n}$  ανήκει στην  $\mathcal{X}$ . Η  $\{\tau_n\}_{n \geq 1}$  θα λέγεται **Τοπική Ακολουθία** για την  $\mathbf{X}$  ως προς την κλάση  $\mathcal{X}$ .

**Ορισμός 2.59.** Ονομάζουμε  $\mathcal{M}$ ,  $\text{sup}\mathcal{M}$  και  $\text{sub}\mathcal{M}$  την κλάση των martingales, supermartingales και submartingales, αντίστοιχα. Προκύπτει άμεσα ότι αν  $\mathbf{X} \in \text{sup}\mathcal{M}$ , τότε  $-\mathbf{X} \in \text{sub}\mathcal{M}$ . Οι διαδικασίες που ανήκουν στις κλάσεις  $\mathcal{M}_{loc}$ ,  $(\text{sup}\mathcal{M})_{loc}$  και  $(\text{sub}\mathcal{M})_{loc}$ , θα λέγονται **Τοπική Martingale**, **Τοπική Supermartingale** και **Τοπική Submartingale**, αντίστοιχα.

**Ορισμός 2.60.** Έστω  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, \mathbb{P})$  διηθημένος χώρος πιθανότητας,  $T = [0, +\infty)$  και  $\mathbf{X} = \{X_t\}_{t \in T}$  στοχαστική διαδικασία προσαρμοσμένη στην  $\mathcal{F}_t$ . Η διαδικασία θα λέγεται **Semimartingale** ( $\mathbf{X} \in \text{semi}\mathcal{M}$ ) αν μπορεί να αναπαρασταθεί στην (όχι μοναδική απαραίτητα) μορφή:

$$\mathbf{X} = X_0 + \mathbf{M} + \mathbf{A}. \quad (2.33)$$

όπου  $X_0$  είναι μία  $\mathcal{F}_0$ -μετρήσιμη τυχαία μεταβλητή με πεπερασμένες τιμές,  $\mathbf{M} = \{M_t\}_{t \in T}$  είναι μία  $\mathcal{F}_t$ -προσαρμοσμένη τοπική martingale διαδικασία ( $\mathbf{M} \in \mathcal{M}_{loc}$ ) και  $\mathbf{A} = \{A_t\}_{t \in T}$  μια  $\mathcal{F}_t$ -προσαρμοσμένη διαδικασία φραγμένης μεταβολής ( $\mathbf{A} \in \mathcal{V}$ ).

**Ορισμός 2.61.** Θα λέμε ότι μια στοχαστική διαδικασία  $\mathbf{X}$  ανήκει στην **κλάση Dirichlet**, την οποία κλάση τη συμβολίζουμε με  $\mathcal{D}$ , αν και μόνο αν, η οικογένεια τυχαίων μεταβλητών  $\{X_\tau : \tau \text{ είναι ένας πεπερασμένος χρόνος διακοπής}\}$  είναι ομοιόμορφα ολοκληρώσιμη.

### 2.3.2 Στοιχειώδη Θεωρήματα

**Θεώρημα 2.62. Αποσύνθεση Doob-Meyer:**

1. Κάθε submartingale  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μπορεί να αποσυντεθεί και να γραφεί στη μορφή:

$$X_t = X_0 + M_t + A_t, \quad t \geq 0, \quad (2.34)$$

όπου  $\mathbf{M} \in \mathcal{M}_{loc}$  και  $\mathbf{A} = (A_t, \mathcal{F}_t)_{t \geq 0}$  είναι μια αύξουσα και προβλέψιμη διαδικασία, που ανήκει στην τοπική κλάση των ολοκληρώσιμων διαδικασιών. Δηλαδή, αν  $\mathcal{A}$  είναι η κλάση των ολοκληρώσιμων διαδικασιών, τότε  $\mathcal{A}_{loc}$  θα είναι η τοπική της κλάση,  $\mathcal{A}^+$  θα είναι η κλάση των αυξουσών και ολοκληρώσιμων διαδικασιών και τελικά,  $\mathbf{A} \in \mathcal{P} \cap \mathcal{A}_{loc}^+$ .

2. Κάθε submartingale  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  που ανήκει στην κλάση Dirichlet ( $\mathcal{D}$ ) μπορεί να αποσυντεθεί και να γραφεί στη μορφή:

$$X_t = X_0 + M_t + A_t, \quad t \geq 0, \quad (2.35)$$

όπου  $\mathbf{M}$  είναι μία ομοιόμορφα ολοκληρώσιμη martingale διαδικασία και  $\mathbf{A}$  μία αύξουσα, προβλέψιμη και ολοκληρώσιμη διαδικασία ( $\mathbf{A} \in \mathcal{P} \cap \mathcal{A}^+$ ).

Οι παραπάνω αποσυνθέσεις είναι μοναδικές στο βαθμό που αν υπήρχε και άλλη, έστω  $X_t = X_0 + M'_t + A'_t$ , τότε οι διαδικασίες  $M_t$  και  $A_t$  θα ήταν στοχαστικά μη διακριτές με τις  $M'_t$  και  $A'_t$ , αντίστοιχα.

**Θεώρημα 2.63. Θεώρημα Επιλεκτικής Διακοπής του Doob:** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία submartingale (ή martingale) διαδικασία και  $\tau$  ένας χρόνος διακοπής. Τότε, η σταματημένη διαδικασία  $\mathbf{X}^\tau = (X_{t \wedge \tau}, \mathcal{F}_t)_{t \geq 0}$  είναι επίσης submartingale (ή martingale).

**Θεώρημα 2.64. Θεώρημα Επιλεκτικής Διακοπής του Hunt:** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία submartingale διαδικασία. Επίσης, έστω  $\tau(\omega)$  και  $\sigma(\omega)$  δύο φραγμένοι χρόνοι διακοπής, με  $\sigma(\omega) \geq \tau(\omega)$ ,  $\omega \in \Omega$ . Τότε,

$$X_\sigma \leq E[X_\tau | \mathcal{F}_\sigma], \quad \mathbb{P} - \sigma.β. \quad (2.36)$$

Το ίδιο Θεώρημα ισχύει και στην περίπτωση που η  $\mathbf{X}$  είναι μια martingale διαδικασία, που τότε θα έχουμε ότι

$$X_\sigma = E[X_\tau | \mathcal{F}_\sigma], \quad \mathbb{P} - \sigma.β. \quad (2.37)$$

Η προϋπόθεση των φραγμένων χρόνων διακοπής μπορεί να αντικατασταθεί από την προϋπόθεση ότι η οικογένεια των τυχαίων μεταβλητών  $\{X_t : t \geq 0\}$  είναι ομοιόμορφα ολοκληρώσιμη.

**Θεώρημα 2.65.** Τα δύο τελευταία Θεωρήματα είναι ειδικές περιπτώσεις για το ακόλουθο, γενικότερο **Θεώρημα Επιλεκτικής Διακοπής:** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία submartingale διαδικασία. Επίσης, έστω  $\tau$  και  $\sigma$  δύο χρόνοι διακοπής για τους οποίους ισχύει:

$$E[|X_\tau|] < \infty, \quad E[|X_\sigma|] < \infty, \quad \liminf_{t \rightarrow \infty} E[\mathbf{1}_{\{\tau > t\}} | X_t] = 0. \quad (2.38)$$

Τότε, στο σύνολο  $\{\tau \geq \sigma\}$  θα ισχύει ότι

$$E[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma, \quad \sigma.β. \quad (2.39)$$

Αν επιπλέον,  $\mathbb{P}(\tau \geq \sigma) = 1$ , τότε

$$E[X_\tau] \geq E[X_\sigma]. \quad (2.40)$$



Το Θεώρημα ισχύει και στην περίπτωση που η  $\mathbf{X}$  είναι martingale και απλά τότε θα έχουμε ότι

$$E[X_\tau | \mathcal{F}_\sigma] = X_\sigma, \text{ σ.β.} \quad (2.41)$$

και

$$E[X_\tau] = E[X_\sigma]. \quad (2.42)$$

**Θεώρημα 2.66. Θεώρημα Σύγκλισης του Doob:** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία submartingale διαδικασία με  $\sup_t E[|X_t|] < \infty$  (ισοδύναμα  $\sup_t E[X_t^+] < \infty$ ). Τότε, υπάρχει μία  $\mathcal{F}_\infty$ -μετρήσιμη τυχαία μεταβλητή  $X_\infty$  (όπου  $\mathcal{F}_\infty \equiv \sigma(\bigcup_{t \geq 0} \mathcal{F}_t)$ ), με  $E[|X_\infty|] < \infty$ , τέτοια ώστε

$$X_t \xrightarrow{t \rightarrow \infty} X_\infty, \text{ P-σ.β.} \quad (2.43)$$

Αν στις προϋποθέσεις του Θεωρήματος προσθέσουμε και την προϋπόθεση της ομοιόμορφης ολοκληρωσιμότητας, τότε η σχεδόν βέβαιη σύγκλιση θα συμβαίνει και στον  $\mathcal{L}^1$ . Συνεπώς, έπεται το ακόλουθο Θεώρημα:

**Θεώρημα 2.67.** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία submartingale διαδικασία με  $\sup_t E[|X_t|] < \infty$  (ισοδύναμα  $\sup_t E[X_t^+] < \infty$ ). Επίσης, η οικογένεια τυχαίων μεταβλητών  $\{X_t : t \geq 0\}$  είναι ομοιόμορφα ολοκληρώσιμη. Τότε, υπάρχει μία  $\mathcal{F}_\infty$ -μετρήσιμη τυχαία μεταβλητή  $X_\infty$ , με  $E[|X_\infty|] < \infty$ , τέτοια ώστε

$$E[|X_t - X_\infty|] \xrightarrow{t \rightarrow \infty} 0. \quad (2.44)$$

**Θεώρημα 2.68. Θεώρημα Σύγκλισης του Lévy:** Έστω  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  ένας διηθημένος χώρος πιθανότητας και  $\xi$  μία ολοκληρώσιμη  $\mathcal{F}$ -μετρήσιμη τυχαία μεταβλητή. Θέτουμε  $\mathcal{F}_\infty = \sigma(\bigcup_{t \geq 0} \mathcal{F}_t)$ . Τότε, P-σ.β. και κατά  $\mathcal{L}^1$ ,

$$E[\xi | \mathcal{F}_t] \xrightarrow{t \rightarrow \infty} E[\xi | \mathcal{F}_\infty]. \quad (2.45)$$

### 2.3.3 Μεγιστικές Ανισότητες

Οι ακόλουθες δύο κλασικές ανισότητες του Kolmogorov και το Khintchine συνέβαλαν στην ανάπτυξη του κλάδου των **Ανισοτήτων Martingale** για τυχαίες μεταβλητές της μορφής

$$\sup_{t \leq T} X_t, \quad \sup_{t \leq T} |X_t|, \quad \sup_{t \geq 0} X_t, \quad \sup_{t \geq 0} |X_t|. \quad (2.46)$$

**Θεώρημα 2.69. Ανισότητες Kolmogorov:** Έστω  $S_n = \xi_1 + \xi_2 + \dots + \xi_n$ ,  $n \geq 1$ , όπου  $\xi_1, \xi_2, \dots$  ανεξάρτητες τυχαίες μεταβλητές με  $E[\xi_k] = 0$ ,  $E[\xi_k^2] < \infty$ ,  $k \geq 1$ . Τότε, για κάθε  $\epsilon > 0$  και τυχαίο  $n \geq 1$ ,

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq \epsilon\right) \leq \frac{E[S_n^2]}{\epsilon^2}. \quad (2.47)$$

Αν επιπλέον,  $P(|\xi_k| \leq c) = 1$ ,  $k \geq 1$ , τότε

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq \epsilon\right) \geq 1 - \frac{(c + \epsilon)^2}{E[S_n^2]}. \quad (2.48)$$

Παρατηρούμε ότι η ακολουθία  $\{S_n\}_{n \geq 1}$  είναι μια martingale στοχαστική διαδικασία, καθώς

1. Για κάθε  $n \geq 1$  έχουμε ότι

$$E[|S_n|] = E\left[\left|\sum_{k=1}^n \xi_k\right|\right] \leq E\left[\sum_{k=1}^n |\xi_k|\right] \leq E\left[n \sup_{1 \leq k \leq n} |\xi_k|\right] = nE\left[\sup_{1 \leq k \leq n} |\xi_k|\right]. \quad (2.49)$$

Δηλαδή,

$$E[|S_n|] < \infty, \quad \forall n \geq 1, \quad (2.50)$$

αφού  $E[\xi_k^2] < \infty$  για κάθε  $k \geq 1$  και άρα  $E[|\xi_k|] < \infty$  για κάθε  $k \geq 1$  και συνεπώς,  $E\left[\sup_{1 \leq k \leq n} \xi_k\right] < \infty$ .

2. Για κάθε  $n \geq 1$  έχουμε ότι

$$E[S_{n+1}|\mathcal{F}_n] = E[S_n + \xi_{n+1}|\mathcal{F}_n] = S_n + E[\xi_{n+1}|\mathcal{F}_n] = S_n. \quad (2.51)$$

**Θεώρημα 2.70. Ανισότητες Khintchine:** Έστω  $\xi_1, \xi_2, \dots$  ανεξάρτητες Bernoulli τυχαίες μεταβλητές, με  $P(\xi_j = 1) = P(\xi_j = -1) = 1/2$ ,  $j \geq 1$  και  $\{c_j\}_{j \geq 1}$  ακολουθία πραγματικών αριθμών. Τότε, για κάθε  $p > 0$  και  $n \geq 1$ ,

$$A_p \left( \sum_{j=1}^n c_j^2 \right)^{\frac{p}{2}} \leq E \left[ \left| \sum_{j=1}^n c_j \xi_j \right|^p \right] \leq B_p \left( \sum_{j=1}^n c_j^2 \right)^{\frac{p}{2}}, \quad (2.52)$$

όπου  $A_p$  και  $B_p$  είναι κάποιες σταθερές που εξαρτώνται από το  $p$ . Παρατηρούμε ότι η ακολουθία  $\{C_n\}_{n \geq 1}$  με  $C_n = \sum_{j=1}^n c_j \xi_j$  είναι μία martingale στοχαστική διαδικασία, καθώς

1. Για κάθε  $n \geq 1$  έχουμε ότι

$$E \left[ \left| \sum_{j=1}^n c_j \xi_j \right| \right] \leq E \left[ \sum_{j=1}^n |c_j \xi_j| \right] = \sum_{j=1}^n (|c_j| E[|\xi_j|]) = \sum_{j=1}^n |c_j| < \infty. \quad (2.53)$$

Δηλαδή,

$$E \left[ |C_n| \right] < \infty, \quad \forall n \geq 1. \quad (2.54)$$

2. Για κάθε  $n \geq 1$  έχουμε ότι

$$E[C_{n+1}|\mathcal{F}_n] = E[C_n + c_{n+1}\xi_{n+1}|\mathcal{F}_n] = C_n + E[c_{n+1}\xi_{n+1}|\mathcal{F}_n] = C_n. \quad (2.55)$$

**Θεώρημα 2.71. Ανισότητες Doob (κατά πιθανότητα):** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία submartingale διαδικασία. Τότε, για κάθε  $\epsilon > 0$  και  $T > 0$ , θα ισχύει ότι:

$$P\left(\sup_{t \leq T} X_t \geq \epsilon\right) \leq \frac{1}{\epsilon} E\left[X_T^+ \mathbb{1}_{\left\{\sup_{t \leq T} X_t \geq \epsilon\right\}}\right] \leq \frac{1}{\epsilon} E\left[X_T^+\right] \quad (2.56)$$

και

$$P\left(\sup_{t \leq T} |X_t| \geq \epsilon\right) \leq \frac{1}{\epsilon} \sup_{t \leq T} E\left[|X_t|\right]. \quad (2.57)$$

Αν  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  είναι martingale, τότε για κάθε  $p \geq 1$ , θα ισχύει ότι:

$$P\left(\sup_{t \leq T} X_t \geq \epsilon\right) \leq \frac{1}{\epsilon^p} E\left[|X_T|^p\right] \quad (2.58)$$

και άρα, συγκεκριμένα για  $p = 2$ ,

$$P\left(\sup_{t \leq T} X_t \geq \epsilon\right) \leq \frac{1}{\epsilon^2} E\left[|X_T|^2\right]. \quad (2.59)$$

**Θεώρημα 2.72. Ανισότητες Doob (κατά μέση τιμή):** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία μη αρνητική submartingale διαδικασία. Τότε, για κάθε  $p > 0$  και  $T > 0$ , θα ισχύει ότι:

$$\mathbb{E} \left[ X_T^p \right] \leq \mathbb{E} \left( \sup_{t \leq T} X_t \right)^p \leq \left( \frac{p}{p-1} \right)^p \mathbb{E} \left[ X_T^p \right] \quad (2.60)$$

ενώ για  $p = 1$  θα έχουμε ότι:

$$\mathbb{E} \left[ X_T \right] \leq \mathbb{E} \left[ \sup_{t \leq T} X_t \right] \leq \frac{e}{e-1} \left[ 1 + \mathbb{E} \left[ X_T \log^+ (X_T) \right] \right]. \quad (2.61)$$

Πιο συγκεκριμένα, αν  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  είναι μία τετραγωνικά ολοκληρώσιμη martingale, τότε θα έχουμε ότι:

$$\mathbb{E} \left[ \sup_{t \leq T} X_t^2 \right] \leq 4 \mathbb{E} \left[ X_T^2 \right]. \quad (2.62)$$

**Θεώρημα 2.73. Ανισότητες Burkholder-Davis-Gundy:** Έστω  $\mathbf{X} = (X_t, \mathcal{F}_t)_{t \geq 0}$  μία martingale. Τότε, για κάθε  $p \geq 1$  υπάρχουν σταθερές  $A_p^*$  και  $B_p^*$ , τέτοιες ώστε για κάθε χρόνο διακοπής  $\tau$ , να ισχύει ότι:

$$A_p^* \mathbb{E} \left[ |X|_{\tau}^{p/2} \right] \leq \mathbb{E} \left[ \sup_{t \leq \tau} |X_t|^p \right] \leq B_p^* \mathbb{E} \left[ |X|_{\tau}^{p/2} \right], \quad (2.63)$$

όπου  $|X|_t$  είναι η τετραγωνική μεταβολή της  $\mathbf{X}$ .

Στην περίπτωση όπου  $p > 1$ , οι παραπάνω ανισότητες είναι ισοδύναμες (λόγω των ανισοτήτων του Doob κατά μέση τιμή για  $p > 1$ ) με τις ακόλουθες ανισότητες:

$$A_p \mathbb{E} \left[ |X|_{\tau}^{p/2} \right] \leq \mathbb{E} \left[ |X_{\tau}|^p \right] \leq B_p \mathbb{E} \left[ |X|_{\tau}^{p/2} \right], \quad (2.64)$$

όπου  $A_p$  και  $B_p$  είναι κάποιες σταθερές και  $\mathbb{E} \left[ |X|_{\tau}^{p/2} \right] < \infty$ .

## Ενότητα 3

### Βέλτιστη Διακοπή

Η θεωρία της Βέλτιστης Διακοπής ([14], [15], [16], [17]) αφορά στο πρόβλημα της επιλογής μιας κατάλληλης στιγμής προκειμένου να πραγματοποιηθεί μια προκαθορισμένη ενέργεια, βασισμένη σε μια ακολουθία από τυχαίες μεταβλητές που έχουν παρατηρηθεί, με σκοπό τη μεγιστοποίηση του εκάστοτε κέρδους ή την ελαχιστοποίηση του κόστους. Προβλήματα τέτοιου τύπου παρουσιάζονται στον κλάδο της στατιστικής, όπου η προαναφερθείσα ενέργεια θα μπορούσε να είναι ένας στατιστικός έλεγχος ή η εκτίμηση μιας παραμέτρου, καθώς και στην επιχειρησιακή έρευνα, όπου η ενέργεια ενδεχομένως θα ήταν η αντικατάσταση κάποιας μηχανής, η πρόσληψη ενός νέου υπαλλήλου ή η πραγματοποίηση μιας νέας παραγγελίας για αναπλήρωση του αποθέματος. Επίσης, η θεωρία της Βέλτιστης Διακοπής εφαρμόζεται στα χρηματοοικονομικά, όπου ο ενδιαφερόμενος αναζητά τη βέλτιστη στιγμή άσκησης ενός δικαιώματος αμερικανικού τύπου, καθώς και σε διάφορους, ακόμη, κλάδους, επιστημονικούς ή καθημερινούς.

Η πρώτη προσέγγιση στο πρόβλημα θα γίνει για διαδικασίες πεπερασμένου χρόνου, μέσω της θεωρίας των Martingales και της μεθόδου «επαγωγή προς τα πίσω» (backwards induction) για πεπερασμένο ορίζοντα, την οποία θα επεκτείνουμε σε άπειρο ορίζοντα. Στη συνέχεια, θα προσεγγίσουμε το πρόβλημα για περιπτώσεις μαρκοβιανών αλυσίδων και στο τέλος θα δούμε τα ανάλογα αποτελέσματα για διαδικασίες συνεχούς χρόνου.

## 3.1 Το Πρόβλημα σε Διακριτό Χρόνο

### 3.1.1 Η περίπτωση των Martingales

Έστω  $G = \{G_n\}_{n \geq 0}$  μια ακολουθία από τυχαίες μεταβλητές, ορισμένη και προσαρμοσμένη στη διήθηση του χώρου πιθανότητας  $(\Omega, \mathcal{F}, \mathcal{F}_n, P)$ . Την  $G_n$  μπορούμε να τη φανταστούμε ως το κέρδος που θα έχουμε (gain) αν σταματήσουμε να παρατηρούμε την  $G$  τη χρονική στιγμή  $n$ .

Συμβολίζουμε με  $\mathcal{K}$  την οικογένεια των φραγμένων χρόνων διακοπής, έτσι ώστε η οικογένεια με όλους τους χρόνους διακοπής να είναι το  $\bar{\mathcal{K}}$ . Συνεπώς, ορίζουμε την υπο-οικογένεια:

$$\mathcal{K}_n^N = \{\tau \in \mathcal{K} : n \leq \tau \leq N\}, \quad (3.1)$$

όπου  $0 \leq n \leq N$ . Για λόγους ευκολίας, θέτουμε  $\mathcal{K}^N = \mathcal{K}_0^N$  και  $\mathcal{K}_n = \mathcal{K}_n^\infty$ .

Το πρόβλημα της *Βέλτιστης Διακοπής* που μελετάμε είναι το:

$$V_* = \sup_{\tau} E[G_\tau], \quad (3.2)$$

όπου το  $\tau$  ανήκει σε μια οικογένεια φραγμένων χρόνων διακοπής. Το πρόβλημα, ουσιαστικά, χωρίζεται σε δύο μέρη:

1. Να υπολογίσουμε την τιμή της συνάρτησης  $V_*$ .
2. Να παρουσιάσουμε έναν χρόνο βέλτιστης διακοπής,  $\tau_*$ , για τον οποίο επιτυγχάνεται το  $V_*$ .

Για να εξασφαλίσουμε την ύπαρξη του  $E[G_\tau]$  θα πρέπει να επιβάλουμε κάποιες περαιτέρω συνθήκες στην  $G$  και στον  $\tau$ . Αν η ακόλουθη συνθήκη ικανοποιείται (με  $G_N \equiv 0$ , όταν  $N = \infty$ ):

$$E \left[ \sup_{n \leq k \leq N} |G_k| \right] < \infty, \quad (3.3)$$

τότε η  $E[G_\tau]$  είναι καλά ορισμένη για κάθε  $\tau \in \mathcal{K}_n^N$ . Ενίοτε, αρκεί στη θέση του  $|G_k|$  να έχουμε απλώς  $G_k^+$  ή  $G_k^-$  (ή ακόμα και να θεωρήσουμε το σύνολο των  $\tau$  για τα οποία η  $E[G_\tau]$  είναι καλά ορισμένη), αλλά για λόγους απλότητας γενικώς θα θεωρούμε ότι η παραπάνω συνθήκη ικανοποιείται.

Στην ίδια λογική, ορίζουμε τις συναρτήσεις

$$V_n^N = \sup_{\tau \in \mathcal{K}_n^N} E[G_\tau], \quad (3.4)$$

με  $0 \leq n \leq N$ . Ξανά, για λόγους απλότητας, θέτουμε  $V^N = V_0^N$  και  $V_n = V_n^\infty$ , ενώ με  $V = V_0^\infty$  θα εννοούμε ότι το supremum λαμβάνεται πάνω σε όλα τα  $\tau$  του  $\mathcal{K}$ .

**1. Η μέθοδος της «επαγωγής προς τα πίσω».** Η πρώτη μέθοδος επίλυσης του προβλήματος, στην περίπτωση που  $N < \infty$ , πρώτα χρησιμοποιεί επαγωγή προς τα πίσω για να κατασκευάσει μια ακολουθία τυχαίων μεταβλητών  $\{S_n^N\}_{0 \leq n \leq N}$  η οποία λύνει το πρόβλημα υπό μία στοχαστική έννοια. Πιο απλά, αρκεί να θυμηθούμε το πρόβλημά μας:

$$V_n^N = \sup_{n \leq \tau \leq N} E[G_\tau]. \quad (3.5)$$

Για  $n = N$ , θα πρέπει να σταματήσουμε αναγκαστικά και τότε το κέρδος  $S_N^N$  θα είναι ίσο με  $G_N$ .

Για  $n = N - 1$ , θα έχουμε την επιλογή είτε να σταματήσουμε και να λάβουμε κέρδος  $S_{N-1}^N = G_{N-1}$ , είτε να συνεχίσουμε, λαμβάνοντας ιδανικά  $S_{N-1}^N = E[S_N^N | \mathcal{F}_{N-1}]$ .

Το τελευταίο συμπέρασμα αντικατοπτρίζει το γεγονός ότι η απόφασή μας εξαρτάται αποκλειστικά από την πληροφορία που έχουμε μέχρι εκείνη τη χρονική στιγμή. Δηλαδή, πληροφορία που ανήκει στην  $\mathcal{F}_{N-1}$ . Έπεται ότι, αν



$G_{N-1} \geq E[S_N^N | \mathcal{F}_{N-1}]$ , τότε θα πρέπει να σταματήσουμε τη χρονική στιγμή  $n = N - 1$ , ενώ, αν  $G_{N-1} < E[S_N^N | \mathcal{F}_{N-1}]$ , τότε θα πρέπει να συνεχίσουμε μέχρι τη στιγμή  $N$ .

Ανάλογη λογική ακολουθούμε και για  $n = N - 2, N - 3, \dots, 0$ . Τελικά, η μέθοδος της επαγωγής προς τα πίσω γεννά μια ακολουθία  $\{S_n^N\}_{0 \leq n \leq N}$ , της οποίας ο τύπος δίνεται αναδρομικά ως:

$$S_n^N = G_N, \quad n = N \quad (3.6)$$

$$S_n^N = \max \left( G_n, E[S_{n+1}^N | \mathcal{F}_n] \right), \quad n = N - 1, \dots, 0 \quad (3.7)$$

Επίσης, η μέθοδος μαρτυρά ότι θα έπρεπε να θεωρήσουμε τον εξής χρόνο διακοπής:

$$\tau_n^N = \inf \{ n \leq k \leq N : S_k^N = G_k \}, \quad (3.8)$$

για  $0 \leq n \leq N$ . Παρατηρούμε ότι το infimum επιτυγχάνεται πάντα, καθώς στη χειρίστη περίπτωση, θα το πετύχουμε τη χρονική στιγμή  $N$ .

**Θεώρημα 3.1. (Πεπερασμένος Ορίζοντας)** Θεωρούμε το πρόβλημα βέλτιστης διακοπής (3.5) και θεωρούμε ότι η συνθήκη (3.3) ισχύει. Τότε, για κάθε  $0 \leq n \leq N$  έχουμε ότι:

$$(i) \quad S_n^N \geq E[G_\tau | \mathcal{F}_n] \quad \forall \tau \in \mathcal{K}_n^N.$$

$$(ii) \quad S_n^N = E[G_{\tau_n^N} | \mathcal{F}_n].$$

Επιπλέον, για δοθέν  $0 \leq n \leq N$ , θα ισχύει ότι:

$$(iii) \quad \text{Ο χρόνος διακοπής } \tau_n^N \text{ είναι βέλτιστος για το (3.5).}$$

$$(iv) \quad \text{Η ακολουθία } \{S_k^N\}_{n \leq k \leq N} \text{ είναι η μικρότερη supermartingale που κυριαρχεί της } \{G_k\}_{n \leq k \leq N}.$$

- (v) Αν ο  $\tau_*$  είναι ένας βέλτιστος χρόνος διακοπής για το (3.5), τότε  $\tau_n^N \leq \tau_*$  P-σ.β.
- (vi) Η σταματημένη διαδικασία  $\{S_{k \wedge \tau_n^N}^N\}_{n \leq k \leq N}$  είναι martingale.

*Απόδειξη.* Για τις δύο πρώτες αποδείξεις, (i) και (ii), θα χρησιμοποιήσουμε την επαγωγή για  $n = N, N-1, \dots, 0$ . Παρατηρούμε ότι αμφότερες οι σχέσεις επαληθεύονται με τετριμμένο τρόπο για  $n = N$  λόγω της σχέσης (3.6) που γράψαμε παραπάνω. Θα θεωρήσουμε ότι οι (i), (ii) ισχύουν για  $n = N, N-1, \dots, k$ , με  $k \geq 1$ , και θα δείξουμε ότι θα πρέπει να ισχύουν και για  $n = k-1$ .

(i) Παίρνουμε τυχαίο  $\tau \in \mathcal{K}_{k-1}^N$  και θέτουμε  $\bar{\tau} = \tau \vee k$ . Τότε,  $\bar{\tau} \in \mathcal{K}_k^N$  και δεδομένου ότι  $\{\tau \geq k\} \in \mathcal{F}_{k-1}$ , ως χρόνος διακοπής, θα έπεται ότι:

$$\begin{aligned} E[G_\tau | \mathcal{F}_{k-1}] &= E[I(\tau = k-1)G_{k-1} | \mathcal{F}_{k-1}] + E[I(\tau \geq k)G_{\bar{\tau}} | \mathcal{F}_{k-1}] \\ &= I(\tau = k-1)G_{k-1} + I(\tau \geq k)E[E[G_{\bar{\tau}} | \mathcal{F}_k] | \mathcal{F}_{k-1}], \end{aligned} \quad (3.9)$$

έχοντας χρησιμοποιήσει μόνο ιδιότητες της δεσμευμένης μέσης τιμής. Σημειώνουμε ότι με  $I$  συμβολίσαμε τη δείκτηρα συνάρτηση για λόγους ευκολίας. Στη συνέχεια, παρατηρούμε ότι η ανισότητα της σχέσης (i) ισχύει για  $n = k$  από υπόθεση επαγωγής. Συνεπώς, αφού  $\bar{\tau} \in \mathcal{K}_k^N$ , θα έχουμε ότι  $E[G_{\bar{\tau}} | \mathcal{F}_{k-1}] \leq S_k^N$ . Παράλληλα, από τη σχέση (3.7), γνωρίζουμε ότι  $G_{k-1} \leq S_{k-1}^N$  και  $E[S_k^N | \mathcal{F}_{k-1}] \leq S_{k-1}^N$ . Εφαρμόζοντας τις τρεις ανισότητες στο δεξί μέλος της (3.9), καταλήγουμε στο:

$$\begin{aligned} E[G_\tau | \mathcal{F}_{k-1}] &\leq I(\tau = k-1)S_{k-1}^N + I(\tau \geq k)E[S_k^N | \mathcal{F}_{k-1}] \\ &\leq I(\tau = k-1)S_{k-1}^N + I(\tau \geq k)S_{k-1}^N = S_{k-1}^N. \end{aligned} \quad (3.10)$$

Δηλαδή, αποδείξαμε τον ισχυρισμό μας, ότι η σχέση (i) ισχύει και για  $n = k-1$  και άρα ισχύει για κάθε  $0 \leq n \leq N$ .

(ii) Θα δείξουμε ότι η (ii) ισχύει για  $n = k - 1$ :

$$\begin{aligned} E[G_{\tau_{k-1}^N} | \mathcal{F}_{k-1}] &= E[I(\tau_{k-1}^N = k-1)G_{k-1} | \mathcal{F}_{k-1}] + E[I(\tau_{k-1}^N \geq k)G_{\tau_k^N} | \mathcal{F}_{k-1}] \\ &= I(\tau_{k-1}^N = k-1)G_{k-1} + I(\tau_{k-1}^N \geq k)E[E[G_{\tau_k^N} | \mathcal{F}_k] | \mathcal{F}_{k-1}] \end{aligned} \quad (3.11)$$

Παρατηρούμε ότι, στο ενδεχόμενο  $\{\tau_{k-1}^N = k-1\}$ , από την (3.8) θα έχουμε ότι  $G_{k-1} = S_{k-1}^N$ . Επίσης, από επαγωγική υπόθεση για  $n = k$ , θα ισχύει ότι  $E[G_{\tau_k^N} | \mathcal{F}_k] = S_k^N$  και τέλος, στο ενδεχόμενο  $\{\tau_{k-1}^N \geq k\}$ , θα έχουμε ότι  $E[S_k^N | \mathcal{F}_{k-1}] = S_{k-1}^N$  από την (3.8) και την (3.7). Συνεπώς, η (3.11) γράφεται ως:

$$\begin{aligned} E[G_{\tau_{k-1}^N} | \mathcal{F}_{k-1}] &= I(\tau_{k-1}^N = k-1)G_{k-1} + I(\tau_{k-1}^N \geq k)E[S_k^N | \mathcal{F}_{k-1}] \\ &= I(\tau_{k-1}^N = k-1)S_{k-1}^N + I(\tau_{k-1}^N \geq k)S_{k-1}^N = S_{k-1}^N. \end{aligned} \quad (3.12)$$

Δηλαδή, αποδείξαμε τον ισχυρισμό μας, ότι η σχέση (ii) ισχύει και για  $n = k-1$  και άρα ισχύει για κάθε  $0 \leq n \leq N$ .

(iii) Για να αποδείξουμε το (iii), αρχικά παίρνουμε τη μέση τιμή στη σχέση (i) και προκύπτει πως  $E[S_n^N] \geq E[G_\tau]$  για κάθε  $\tau \in \mathcal{K}_n^N$ . Στη συνέχεια, παίρνουμε το supremum ως προς  $\tau \in \mathcal{K}_n^N$  και βλέπουμε ότι  $E[S_n^N] \geq V_n^N$ . Από την άλλη, παίρνουμε τη μέση τιμή στη σχέση (ii) και προκύπτει ότι  $E[S_n^N] = E[G_{\tau_n^N}]$  που συνεπάγεται ότι  $E[S_n^N] \leq V_n^N$ . Από τις δύο ανισότητες, καταλήγουμε στην ισότητα, δηλαδή  $V_n^N = E[S_n^N]$  και, αφού  $E[S_n^N] = E[G_{\tau_n^N}]$ , έπεται το ζητούμενο. Δηλαδή, ότι  $V_n^N = E[G_{\tau_n^N}]$ .

(iv) Από την (3.7) έπεται ότι

$$S_k^N \geq E[S_{k+1}^N | \mathcal{F}_k] \quad (3.13)$$

για κάθε  $n \leq k \leq N-1$ , γεγονός που καθιστά την  $\{S_k^N\}_{n \leq k \leq N}$  supermartingale. Από την (3.6) και την (3.7) προκύπτει ότι  $S_k^N \geq G_k$  P-σ.β. για κάθε

$n \leq k \leq N$  που σημαίνει ότι η  $\{S_k^N\}_{n \leq k \leq N}$  κυριαρχεί έναντι της  $\{G_k\}_{n \leq k \leq N}$ . Επιπλέον, αν η  $\{\tilde{S}_k\}_{n \leq k \leq N}$  είναι μια άλλη supermartingale που κυριαρχεί έναντι της  $\{G_k\}_{n \leq k \leq N}$ , τότε ο ισχυρισμός ότι η  $\tilde{S}_k \geq S_k^N$  P-σ.β. μπορεί να επαληθευθεί με χρήση επαγωγής για  $k = N, N-1, \dots, l$ . Πράγματι, αν  $k = N$ , τότε ο ισχυρισμός προκύπτει άμεσα από την (3.6), αφού  $S_N^N = G_N$  και  $\tilde{S}_N \geq G_N$ . Υποθέτουμε ότι  $S_k \geq S_k^N$  P-σ.β. για  $k = N, N-1, \dots, l$ , με  $l \geq n+1$  και τότε, από την (3.7) προκύπτει ότι  $S_{l-1}^N = \max(G_{l-1}, E[S_l^N | \mathcal{F}_{l-1}]) \leq \max(G_{l-1}, E[\tilde{S}_l^N | \mathcal{F}_{l-1}])$ . Τελικά, επειδή η  $\{\tilde{S}_k\}_{n \leq k \leq N}$  είναι supermartingale, καταλήγουμε στο ότι  $\tilde{S}_{l-1} \geq S_{l-1}^N$  P-σ.β. και δηλαδή αποδείξαμε τον αρχικό ισχυρισμό.

(v) Για την απόδειξη του (v), ξεκινάμε με τον ισχυρισμό ότι το γεγονός πως ο  $\tau_*$  είναι βέλτιστος, συνεπάγεται ότι  $S_{\tau_*}^N = G_{\tau_*}$  P-σ.β. Πράγματι, αν δεν ίσχυε αυτό, τότε χρησιμοποιώντας το ότι  $S_k^N \geq G_k$  από τις (3.6), (3.7), θα παίρναμε ότι  $S_{\tau_*}^N \geq G_{\tau_*}$ , με  $P(S_{\tau_*}^N > G_{\tau_*}) > 0$ . Από το τελευταίο, προκύπτει ότι  $E[G_{\tau_*}] < E[S_{\tau_*}^N]$ . Επίσης, από την supermartingale ιδιότητα της  $\{S_k^N\}_{n \leq k \leq N}$  και το Θεώρημα Επιλεκτικής Διακοπής, έπεται ότι  $E[S_{\tau_*}^N] \leq E[S_n^N]$ , ενώ από την απόδειξη του (iii) είδαμε ότι  $E[S_n^N] = V_n^N$ . Συνεπώς, έπεται ότι  $E[G_{\tau_*}] < V_n^N$ , που έρχεται σε αντίθεση με το γεγονός ότι ο  $\tau_*$  είναι βέλτιστος και άρα ισχύει ο αρχικός ισχυρισμός, ότι δηλαδή  $S_{\tau_*}^N = G_{\tau_*}$  P-σ.β. Τέλος, από ορισμό της (3.8), θα έχουμε ότι  $\tau_n^N \leq \tau_*$  P-σ.β.

(vi) Για να επιβεβαιώσουμε την martingale ιδιότητα

$$E[S_{(k+1) \wedge \tau_n^N}^N | \mathcal{F}_k] = S_{k \wedge \tau_n^N}^N \quad (3.14)$$

για δοθέν  $n \leq k \leq N-1$ , αρκεί να παρατηρήσουμε ότι:

$$\begin{aligned} E[S_{(k+1) \wedge \tau_n^N}^N | \mathcal{F}_k] &= \\ &= E[I(\tau_n^N \leq k) S_{k \wedge \tau_n^N}^N | \mathcal{F}_k] + E[I(\tau_n^N \geq k+1) S_{k+1}^N | \mathcal{F}_k] \\ &= I(\tau_n^N \leq k) S_{\tau_n^N}^N + I(\tau_n^N \geq k+1) E[S_{k+1}^N | \mathcal{F}_k] \\ &= I(\tau_n^N \leq k) S_{\tau_n^N}^N + I(\tau_n^N \geq k+1) S_k^N = S_{k \wedge \tau_n^N}^N \end{aligned} \quad (3.15)$$

με το δεύτερο μέρος της τελευταίας ισότητας να προκύπτει από το γεγονός ότι  $E[S_{k+1}^N | \mathcal{F}_k] = S_k^N$  στο ενδεχόμενο  $\{\tau_n^N \geq k+1\}$ , ενώ  $\{\tau_n^N \geq k+1\} \in \mathcal{F}_k$ , επειδή ο  $\tau_n^N$  είναι χρόνος διακοπής. Έτσι, ολοκληρώθηκε η απόδειξη του (vi) και μαζί της ολοκληρώθηκε και η απόδειξη του Θεωρήματος 3.1 (Πεπερασμένος Ορίζοντας).

□

Παρατηρούμε ότι με τη χρήση του Θεωρήματος Επιλεκτικής Διακοπής, η σχέση (i) που αποδείξαμε στο Θεώρημα 3.1 θα μπορούσε να εξαχθεί άμεσα από τη supermartingale ιδιότητα του (iv), ενώ θα μπορούσαμε να συμπεράνουμε το (ii) από τη martingale ιδιότητα του (vi).

**2. Η μέθοδος του ουσιώδους supremum (essential supremum).** Η μέθοδος της επαγωγής προς τα πίσω εκ φύσεως απαιτεί πεπερασμένο ορίζοντα ( $N > 0$ ). Ωστόσο, φαίνεται ότι η επέκταση στον άπειρο ορίζοντα μπορεί να επιτευχθεί.

Από τα (i) και (ii) του Θεωρήματος 3.1 συμπεραίνουμε ότι γενικώς ισχύει:

$$S_n^N = \sup_{\tau \in \mathcal{K}_n^N} E[G_\tau | \mathcal{F}_n]. \quad (3.16)$$

Αμφότερα τα (i) και (ii), όμως, ισχύουν μόνο P-σ.β. και το P-κενό σύνολο μπορεί να εξαρτάται από το  $\tau \in \mathcal{K}_n^N$ . Οπότε, αν διαλέξουμε το supremum της (3.16) από ένα σύνολο με υπεραριθμήσιμα πολλά  $\tau$ , τότε το δεξί μέλος δεν ορίζει αναγκαστικά μετρήσιμη συνάρτηση και ως εκ τούτου, όλη η σχέση (3.16) μπορεί να μην ισχύει. Για να ξεπεράσουμε αυτή τη δυσκολία, θα μας φανεί χρήσιμη η έννοια του essential supremum.

**Λήμμα 3.2. (Essential Supremum)** Έστω  $\{Z_\alpha : \alpha \in I\}$ , μία οικογένεια τυχαίων μεταβλητών στον  $(\Omega, \mathcal{F}, P)$ , όπου ο δείκτης  $I$  μπορεί να είναι αυθαίρετος. Τότε, υπάρχει ένα αριθμήσιμο υποσύνολο  $J$  του  $I$ , τέτοιο ώστε η τυχαία μεταβλητή  $Z^* : \Omega \rightarrow \overline{\mathbb{R}}$ , που ορίζεται ως

$$Z^* = \sup_{\alpha \in J} Z_\alpha, \quad (3.17)$$

να ικανοποιεί τις εξής δύο ιδιότητες:

$$(i) \quad P(Z_\alpha \leq Z^*) = 1, \quad \forall \alpha \in I. \quad (3.18)$$

(ii) Αν  $\tilde{Z} : \Omega \rightarrow \overline{\mathbb{R}}$  είναι μια άλλη τυχαία μεταβλητή που ικανοποιεί την (3.18) στη θέση της  $Z^*$ , τότε

$$P(Z^* \leq \tilde{Z}) = 1. \quad (3.19)$$

Η τυχαία μεταβλητή  $Z^*$  ονομάζεται **essential supremum (ουσιώδες supremum)** της  $\{Z_\alpha : \alpha \in I\}$  ως προς  $P$ , συμβολίζεται με  $Z^* = \text{ess sup}_{\alpha \in I} Z_\alpha$  και καθορίζεται με μοναδικό τρόπο από τις σχέσεις (3.18) και (3.19).

Επιπλέον, αν η οικογένεια  $\{Z_\alpha : \alpha \in I\}$  έχει κατεύθυνση προς τα πάνω, εννοώντας ότι για κάθε  $\alpha$  και  $\beta$  στο  $I$  υπάρχει ένα  $\gamma$  στο  $I$  τέτοιο ώστε

$$Z_\alpha \vee Z_\beta \leq Z_\gamma \quad P - \sigma. \beta. \quad (3.20)$$

τότε το αριθμήσιμο σύνολο  $J = \{\alpha_n : n \geq 1\}$  μπορεί να επιλεγεί με τρόπο τέτοιο ώστε

$$Z^* = \lim_{n \rightarrow \infty} Z_{\alpha_n} \quad P - \sigma. \beta. \quad (3.21)$$

όπου  $Z_{\alpha_1} \leq Z_{\alpha_2} \leq \dots$ ,  $P$ -σ.β.

Έχοντας την έννοια του essential supremum μπορούμε να ξαναγράψουμε τις σχέσεις (i) και (ii) του Θεωρήματος 3.1 ως εξής:

$$S_n^N = \text{ess sup}_{n \leq \tau \leq N} E[G_\tau | \mathcal{F}_n] \quad (3.22)$$

για κάθε  $0 \leq n \leq N$ . Αυτή η ιδιότητα παρέχει ένα περαιτέρω χαρακτηριστικό για την ακολουθία τυχαίων μεταβλητών  $\{S_n^N\}_{0 \leq n \leq N}$  που παρουσιάστηκαν αρχικά ως παράγωγο των αναδρομικών σχέσεων (3.6) και (3.7). Το πλεονέκτημα της εν λόγω ιδιότητας συγκριτικά με τις αναδρομικές σχέσεις έγκειται στο γεγονός ότι η (3.22) μπορεί φυσικά να επακταθεί στην περίπτωση του άπειρου ορίζονται  $N$ .

Θεωρούμε το πρόβλημα βέλτιστης διακοπής (3.5) για  $N = \infty$ . Δηλαδή, ξαναγράφουμε το πρόβλημα προσαρμοσμένο στο  $N = \infty$  ως:

$$V_n^N = \sup_{\tau \geq n} E[G_\tau], \quad (3.23)$$

όπου ο  $\tau$  είναι ένας χρόνος διακοπής και  $n \geq 0$ . Για να λύσουμε το πρόβλημα θα θεωρήσουμε την ακολουθία τυχαίων μεταβλητών  $\{S_n\}_{n \geq 0}$ , ορισμένη ως:

$$S_n = \text{ess sup}_{\tau \geq n} E[G_\tau | \mathcal{F}_n], \quad (3.24)$$

όπως επίσης και τον εξής χρόνο διακοπής:

$$\tau_n = \inf\{k \geq n : S_k = G_k\}, \quad (3.25)$$

για  $n \geq 0$ , όπου  $\inf \emptyset = \infty$ , εξ ορισμού. Η ακολουθία  $\{S_n\}_{n \geq 0}$  συχνά αναφέρεται ως ο φάκελος Snell της  $G$  (Snell envelope of  $G$ ).

Το πρώτο μέρος του ακόλουθου Θεωρήματος δείχνει ότι η  $\{S_n\}_{n \geq 0}$  ικανοποιεί τις ίδιες αναδρομικές σχέσεις με την  $\{S_n^N\}_{0 \leq n \leq N}$ . Το δεύτερο μέρος δείχνει ότι η  $S_n$  και ο  $\tau_n$  λύνουν το πρόβλημα υπό τη στοχαστική έννοια. Το τρίτο μέρος δείχνει ότι αυτό οδηγεί στη λύση του αρχικού προβλήματος (3.23). Τέλος, το τέταρτο μέρος μας παρέχει έναν supermartingale χαρακτηρισμό για τη λύση.

**Θεώρημα 3.3. (Άπειρος Ορίζοντας)** Θεωρούμε το πρόβλημα βέλτιστης διακοπής (3.23), με την προϋπόθεση (3.3) να ισχύει. Τότε:

(i) Θα ισχύουν οι εξής αναδρομικές σχέσεις:

$$S_n = \max(G_n, E[S_{n+1} | \mathcal{F}_n]) \quad (3.26)$$

για κάθε  $n \geq 0$ .

Επιπλέον, θεωρούμε ότι ισχύει η προϋπόθεση

$$P(\tau_n < \infty) = 1 \quad (3.27)$$

για  $n \geq 0$ . Τότε, για κάθε  $n \geq 0$  έχουμε ότι:

(ii)

$$S_n \geq E[G_\tau | \mathcal{F}_n], \quad \forall \tau \in \mathcal{K}_n, \quad (3.28)$$

$$S_n = E[G_{\tau_n} | \mathcal{F}_n]. \quad (3.29)$$

Επίσης, για δοθέν  $n \geq 0$ , θα έχουμε ότι:

(iii) Ο χρόνος διακοπής  $\tau_n$  είναι βέλτιστος για το (3.23) και αν  $\tau_*$  είναι ένας άλλος βέλτιστος χρόνος διακοπής για το (3.23), τότε  $\tau_n \leq \tau_*$  P-σ.β.

(iv) Η ακολουθία  $\{S_k\}_{k \geq n}$  είναι η ελάχιστη supermartingale που κυριαρχεί της  $\{G_k\}_{k \geq n}$ .

Η σταματημένη διαδικασία  $\{S_{k \wedge \tau_n}\}_{k \geq n}$  είναι martingale.



Τέλος, αν η συνθήκη (3.27) δεν ισχύει, έτσι ώστε  $P(\tau_n = \infty) > 0$ , τότε με πιθανότητα 1 δε θα υπάρξει χρόνος βέλτιστης διακοπής στο (3.23).

**3. Από τον πεπερασμένο ορίζοντα στον άπειρο.** Στο τελευταίο μέρος αυτής της υποενότητας θα εξερευνήσουμε τη σύνδεση ανάμεσα στις δύο μεθόδους, για την περίπτωση που το  $N$  τείνει στο άπειρο στην πρώτη μέθοδο. Για τον σκοπό αυτόν, από την (3.22) παρατηρούμε ότι οι απεικονίσεις  $N \mapsto S_n^N$  και  $N \mapsto \tau_n^N$  είναι αύξουσες, έτσι ώστε τα όρια

$$S_n^\infty = \lim_{N \rightarrow \infty} S_n^N, \quad (3.30)$$

$$\tau_n^\infty = \lim_{N \rightarrow \infty} \tau_n^N, \quad (3.31)$$

να υπάρχουν P-σ.β. για κάθε  $n \geq 0$ . Αντίστοιχα, από την (3.5) παρατηρούμε ότι η  $N \mapsto V_n^N$  είναι αύξουσα, έτσι ώστε το όριο

$$V_n^\infty = \lim_{N \rightarrow \infty} V_n^N, \quad (3.32)$$

να υπάρχει για κάθε  $n \geq 0$ . Επίσης, από τις (3.22) και (3.24) βλέπουμε ότι

$$S_n^\infty \leq S_n, \quad (3.33)$$

$$\tau_n^\infty \leq \tau_n, \quad (3.34)$$

P-σ.β. για κάθε  $n \geq 0$ . Ομοίως, από το (ii) του Θεωρήματος 3.1 και τη σχέση (3.29) συμπεραίνουμε ότι

$$V_n^\infty \leq V_n, \quad (3.35)$$

για κάθε  $n \geq 0$ .

Σημειώνουμε ότι εν απουσία της συνθήκης (3.3), οι ανισότητες (3.33)-(3.35) μπορεί να γίνουν *αυστηρές*.

**Θεώρημα 3.4.** (Από τον πεπερασμένο ορίζοντα στον άπειρο) Θεωρούμε τα προβλήματα βέλτιστης διακοπής (3.5) και (3.23), ενώ η συνθήκη (3.3) ισχύει. Τότε, για κάθε  $n \geq 0$  οι ανισότητες (3.33)-(3.35) ισχύουν ως *ισότητες*.

### 3.1.2 Η περίπτωση των Μαρκοβιανών Αλυσίδων

Σε αυτήν την υποενότητα θα παρουσιάσουμε βασικά αποτελέσματα της βέλτιστης διακοπής όταν ο χρόνος είναι διακριτός και η διαδικασία, Μαρκοβιανή.

Θα θεωρούμε μια χρονικά ομοιογενή Μαρκοβιανή αλυσίδα  $\mathbf{X} = \{X_n\}_{n \geq 0}$ , ορισμένη σε έναν διηθημένο χώρο πιθανότητας  $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n \geq 0}, P_x)$  και με τιμές στον μετρήσιμο χώρο  $(E, \mathcal{B})$ , όπου για λόγους απλότητας θεωρούμε ότι  $E = \mathbb{R}^d$  για κάποιο  $d \geq 1$  και  $\mathcal{B} = \mathcal{B}(\mathbb{R}^d)$  είναι η Borel σ-άλγεβρα στον  $\mathbb{R}^d$ . Επίσης, έστω ότι η αλυσίδα  $\mathbf{X}$  ξεκινά από το  $x$  υπό το μέτρο  $P_x$  για κάποιο  $x \in E$  και παράλληλα, η απεικόνιση  $x \mapsto P_x(F)$  είναι μετρήσιμη για κάθε  $F \in \mathcal{F}$ . Έτσι, θα έπεται ότι η απεικόνιση  $x \mapsto E_x[Z]$  είναι και αυτή μετρήσιμη για κάθε τυχαία μεταβλητή  $Z$ . Τέλος, χωρίς βλάβη της γενικότητας, θεωρούμε ότι ο  $(\Omega, \mathcal{F})$  ισούται με τον κανονικό χώρο  $(E^{\mathbb{N}_0}, \mathcal{B}^{\mathbb{N}_0})$ , ούτως ώστε ο τελεστής μετατόπισης  $\theta_n : \Omega \rightarrow \Omega$  να ορίζεται καλά από τη σχέση  $\theta_n(\omega)(k) = \omega(n+k)$  για  $\omega = \{\omega(k)\}_{k \geq 0} \in \Omega$  και  $n, k \geq 0$ . Υπενθυμίζουμε πως  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ .

**1.** Δεδομένης μιας μετρήσιμης συνάρτησης  $G : E \rightarrow \mathbb{R}$  που ικανοποιεί την εξής συνθήκη (με  $G(X_N) = 0$ , αν  $N = \infty$ ):

$$E_x \left[ \sup_{0 \leq n \leq N} |G(X_n)| \right] < \infty, \quad (3.36)$$

για κάθε  $x \in E$ , θεωρούμε το πρόβλημα βέλτιστης διακοπής:

$$V^N(x) = \sup_{0 \leq \tau \leq N} E_x[G(X_\tau)], \quad (3.37)$$

όπου  $x \in E$  και το supremum λαμβάνεται πάνω από όλους τους χρόνους διακοπής  $\tau$  του  $\mathbf{X}$ . Το τελευταίο σημαίνει ότι ο  $\tau$  είναι ένας χρόνος διακοπής ως προς τη φυσική διήθηση του  $\mathbf{X}$ , που δίνεται από τον τύπο  $\mathcal{F}_n^{\mathbf{X}} = \sigma(X_k : 0 \leq k \leq n)$  για  $n \geq 0$ . Επειδή ισχύουν τα ίδια αποτελέσματα αν πάρουμε το supremum στην (3.37) πάνω από τους χρόνους διακοπής  $\tau$  ως προς την  $\{\mathcal{F}_n\}_{n \geq 0}$  και αυτή η προϋπόθεση κάνει τα τελικά συμπεράσματα πιο ισχυρά, θα θεωρήσουμε στη συνέχεια ότι το supremum της (3.37) λαμβάνεται πάνω από αυτή τη μεγαλύτερη κλάση χρόνων διακοπής. Παρατηρούμε, επίσης, ότι στην (3.37) αφήνουμε να εννοηθεί πως το  $N$  θα μπορούσε να είναι  $+\infty$ . Σε αυτήν την περίπτωση, ωστόσο, θα θεωρήσουμε πως το supremum επιτυγχάνεται πάνω σε πεπερασμένους χρόνους διακοπής, δηλαδή  $\tau < \infty$ , P-σ.β.

**2.** Για να λύσουμε το πρόβλημα (3.37) στην περίπτωση που  $N < \infty$ , μπορούμε να παρατηρήσουμε ότι θέτοντας

$$G_n = G(X_n), \quad (3.38)$$

για  $n \geq 0$ , το πρόβλημα (3.37) μετατρέπεται στο πρόβλημα (3.5), μόνο που πλέον αντί για  $P$  και  $E$  έχουμε  $P_x$  και  $E_x$ , αντίστοιχα, με  $x \in E$ .

Έχοντας αναγνωρίσει το πρόβλημα (3.37) ως το (3.5), μπορούμε να εφαρμόσουμε τη μέθοδο της προς τα πίσω επαγωγής (3.6)-(3.7) η οποία καταλήγει σε μια ακολουθία τυχαίων μεταβλητών  $\{S_n^N\}_{0 \leq n \leq N}$  και έναν χρόνο διακοπής  $\tau_n^N$ , ο οποίος προκύπτει από την (3.8). Η ιδιότητα που θα παίξει κύριο ρόλο στην ανάλυσή μας είναι η

$$S_n^N = V^{N-n}(X_n) \quad (3.39)$$

για  $0 \leq n \leq N$ , την οποία θα εδραιώσουμε στην απόδειξη του επόμενου Θεωρήματος. Άπαξ και γνωρίζουμε την (3.39), τα αποτελέσματα του Θεωρήματος (3.1) μεταφράζονται άμεσα στις τρέχουσες ρυθμίσεις και αποκτούν μια πιο ξεκάθαρη μορφή όπως θα δούμε ακολούθως:

Θέτουμε

$$C_n = \{x \in E : V^{N-n}(x) > G(x)\}, \quad (3.40)$$

$$D_n = \{x \in E : V^{N-n}(x) = G(x)\}, \quad (3.41)$$

για  $0 \leq n \leq N$ . Επίσης, ορίζουμε

$$\tau_D = \inf\{0 \leq n \leq N : X_n \in D\}. \quad (3.42)$$

Τέλος, ο τελεστής μετάβασης  $T$  του  $\mathbf{X}$  ορίζεται ως

$$TF(x) = E_x[F(X_1)], \quad (3.43)$$

για  $x \in E$ , όπου η  $F : E \rightarrow \mathbb{R}$  είναι μια μετρήσιμη συνάρτηση έτσι ώστε η  $F(X_1)$  να είναι ολοκληρώσιμη ως προς το μέτρο  $P_x$  για κάθε  $x \in E$ .

**Θεώρημα 3.5. (Πεπερασμένος Ορίζοντας: Η χρονικά ομοιογενής περίπτωση)** Θεωρούμε το πρόβλημα (3.37), αφού υποθέσουμε πως ισχύει η (3.36). Τότε, η συνάρτηση-αξίας  $V^n$  ικανοποιεί τις εξισώσεις *Wald-Bellman*:

$$(i) \quad V^n(x) = \max(G(x), TV^{n-1}(x)), \quad x \in E \quad (3.44)$$

για  $n = 1, \dots, N$ , όπου  $V^0 = G$ . Επιπλέον, έχουμε ότι:

- (ii) Ο χρόνος διακοπής  $\tau_D$  είναι βέλτιστος για το (3.37).
- (iii) Αν  $\tau_*$  είναι ένας άλλος βέλτιστος χρόνος διακοπής για το (3.37), τότε  $\tau_D \leq \tau_*$  P-σ.β. για κάθε  $x \in E$ .

(iv) Η ακολουθία  $\{V^{N-n}(X_n)\}_{0 \leq n \leq N}$  είναι η μικρότερη supermartingale που κυριαρχεί της  $\{G(X_n)\}_{0 \leq n \leq N}$  υπό το μέτρο  $P_x$  για δοθέν  $x \in E$ .

(v) Η σταματημένη διαδικασία  $\{V^{N-n \wedge \tau_D}(X_{n \wedge \tau_D})\}_{0 \leq n \leq N}$  είναι martingale υπό το μέτρο  $P_x$  για κάθε  $x \in E$ .

*Απόδειξη.* Αρχικά, για να αποδείξουμε την (3.39), θα θυμηθούμε από τη σχέση (ii) του Θεωρήματος 3.1 ότι

$$S_n^N = E_x[G(X_{\tau_n^N}) | \mathcal{F}_n], \quad (3.45)$$

για κάθε  $0 \leq n \leq N$ . Αφού  $S_k^{N-n} \circ \theta_n = S_{n+k}^N$ , έπεται ότι ο  $\tau_n^N$  ικανοποιεί τη σχέση

$$\tau_n^N = \inf\{n \leq k \leq N : S_k^N = G(X_k)\} = n + \tau_0^{N-n} \circ \theta_n, \quad (3.46)$$

για  $0 \leq n \leq N$ . Επίσης, παρατηρούμε ότι

$$S_{\tau_n^N}^N = S_{\tau_0^{N-n}}^{N-n} \circ \theta_n = S_{\tau_0^{N-n}}^{N-n} \circ \theta_n = G(X_{\tau_0^{N-n}}) \circ \theta_n \quad (3.47)$$

και έπειτα, αντικαθιστούμε την (3.47) στην (3.45), χρησιμοποιώντας την ιδιότητα Markov, και παίρνουμε ότι

$$\begin{aligned} S_n^N &= E_x[G(X_{\tau_n^N}) | \mathcal{F}_n] = E_x[S_{\tau_n^N}^N | \mathcal{F}_n] \\ &= E_x[G(X_{\tau_0^{N-n}}) \circ \theta_n | \mathcal{F}_n] \\ &= E_{x_n}[G(X_{\tau_0^{N-n}})] = V^{N-n}(X_n), \end{aligned} \quad (3.48)$$

$$(3.49)$$

όπου η τελευταία ισότητα προκύπτει από τις (i)-(ii) του Θεωρήματος (3.1) μέσα από τις οποίες βλέπουμε ότι

$$E_x[S_0^{N-n}] = E_x[G(X_{\tau_0^{N-n}})] = \sup_{0 \leq \tau \leq N-n} E_x[G(X_\tau)] = V^{N-n}(x) \quad (3.50)$$

για  $0 \leq n \leq N$  και  $x \in E$ . Συνεπώς, η (3.39) ισχύει όπως ισχυριστήκαμε.

Για να αποδείξουμε την (3.44), παρατηρούμε ότι η (3.7), χρησιμοποιώντας την (3.39) και τη μαρκοβιανή ιδιότητα, γράφεται ως εξής:

$$\begin{aligned} V^{N-n}(X_n) &= \max(G(X_n), E_x[V^{N-n-1}(X_{n+1}) | \mathcal{F}_n]) \\ &= \max(G(X_n), E_x[V^{N-n-1}(X_1) \circ \theta_n | \mathcal{F}_n]) \\ &= \max(G(X_n), E_{x_n}[V^{N-n-1}(X_1)]) \\ &= \max(G(X_n), TV^{N-n-1}(X_1)), \end{aligned} \quad (3.51)$$

για κάθε  $0 \leq n \leq N$ . Επιτρέποντας το  $n$  να γίνει 0 και χρησιμοποιώντας ότι  $X_0 = x$  υπό το  $P_x$ , βλέπουμε ότι η (3.51) μας δίνει τη ζητούμενη (3.44).

Τα υπόλοιπα μέρη του Θεωρήματος 3.5 προκύπτουν άμεσα από το Θεώρημα 3.1 και η απόδειξη ολοκληρώθηκε. □

**3.** Οι εξισώσεις **Wald-Bellman** (3.44) μπορούν να γραφούν σε μια πιο συμπαγή μορφή. Για τον σκοπό αυτόν, ορίζουμε τον τελεστή  $Q$ , θέτοντας:

$$QF(x) = \max(G(x), TF(x)), \quad (3.52)$$

για  $x \in E$ , όπου  $F: E \rightarrow \mathbb{R}$  είναι μια μετρήσιμη συνάρτηση για την οποία ισχύει ότι  $F(X_1) \in \mathcal{L}^1(P_x)$  για  $x \in E$ . Τότε, η (3.44) γράφεται ως

$$V^n(x) = Q^n G(x), \quad (3.53)$$

για  $1 \leq n \leq N$ , όπου με  $Q^n$  συμβολίζουμε τη  $n$ -οστή δύναμη του  $Q$ .

4. Έστω τώρα ότι βρισκόμαστε στην περίπτωση που έχουμε μια χρονικά ανομοιογενή Μαρκοβιανή αλυσίδα. Θέτοντας  $Z_n = (n, X_n)$  για  $n \geq 0$ , ξέρουμε ότι η  $Z = \{Z_n\}_{n \geq 0}$  θα είναι μια χρονικά ομοιογενής Μαρκοβιανή αλυσίδα. Έστω μια μετρήσιμη συνάρτηση  $G : \{0, 1, \dots, N\} \times E \rightarrow \mathbb{R}$  που ικανοποιεί την εξής συνθήκη:

$$E_{n,x} \left[ \sup_{0 \leq k \leq N-n} |G(n+k, X_{n+k})| \right] < \infty, \quad (3.54)$$

για κάθε  $0 \leq n \leq N$  και  $x \in E$ . Τότε, το πρόβλημα βέλτιστης διακοπής (3.37) με φυσικό τρόπο επεκτείνεται ως:

$$V^N(n, x) = \sup_{0 \leq \tau \leq N-n} E_{n,x} [G(n+\tau, X_{n+\tau})], \quad (3.55)$$

όπου το supremum λαμβάνεται πάνω από τους χρόνους διακοπής  $\tau$  του  $\mathbf{X}$  και  $X_n = x$  υπό το μέτρο  $P_{n,x}$  με  $0 \leq n \leq N$  και δοθέν  $x \in E$ .

Όπως και παραπάνω, μπορούμε να διαπιστώσουμε ότι

$$S_{n+k}^N = V^N(n+k, X_{n+k}), \quad (3.56)$$

υπό το μέτρο  $P_{n,x}$  για  $0 \leq n \leq N-n$ . Επιπλέον, χρησιμοποιώντας την (3.7) και τη μαρκοβιανή ιδιότητα, έπεται ότι:

$$\begin{aligned} V^N(n+k, X_{n+k}) &= \\ &= \max \left( G(n+k, X_{n+k}), E_{n,x} [V^N(n+k+1, X_{n+k+1}) | \mathcal{F}_{n+k}] \right) \\ &= \max \left( G(Z_{n+k}), E_z [V^N(Z_{n+k+1}) | \mathcal{F}_{n+k}] \right) \\ &= \max \left( G(Z_{n+k}), E_z [V^N(Z_1) \circ \theta_{n+k} | \mathcal{F}_{n+k}] \right) \\ &= \max \left( G(Z_{n+k}), E_{Z_{n+k}} [V^N(Z_1)] \right) \end{aligned} \quad (3.57)$$

$$(3.58)$$

για  $0 \leq k \leq N-n-1$ , όπου  $z = (n, x)$  με  $0 \leq n \leq N$  και  $x \in E$ . Επιτρέποντας  $k = 0$  και χρησιμοποιώντας ότι  $Z_n = z = (n, x)$  υπό το μέτρο  $P_z$ , παίρνουμε ότι

$$V^N(n, x) = \max(G(n, x), TV^N(n, x)), \quad (3.59)$$

για  $n = N-1, \dots, 1, 0$  όπου  $TV^N(N-1, x) = E_{N-1, x} G(N, X_N)$  και  $T$  είναι ο τελεστής μετάβασης του  $Z$ , που δίνεται από τον τύπο

$$TF(n, x) = E_{n, x} F(n+1, X_{n+1}), \quad (3.60)$$

για  $0 \leq n \leq N$  και  $x \in E$ , όταν το δεξί μέλος της (3.60) είναι καλά ορισμένο (δηλαδή πεπερασμένο).

Οι αναδρομικές σχέσεις της (3.59) είναι οι εξισώσεις *Wald-Bellman* για το χρονικά ανομοιογενές πρόβλημα (3.55). Παρατηρούμε ότι όταν η  $\mathbf{X}$  είναι χρονικά ομοιογενής (και  $G = G(x)$  μόνο), έχουμε  $V^N(n, x) = V^{N-n}(x)$  και η (3.59) ανάγεται στην (3.44). Προκειμένου να παρουσιάσουμε μια αναδιατύπωση της ιδιότητας (iii) του Θεωρήματος 3.5, θα χρειαστεί να ορίσουμε μερικές έννοιες.

**5.** Ο ακόλουθος ορισμός παίζει στοιχειώδη ρόλο στην εύρεση λύσης για το πρόβλημα (3.55).

**Ορισμός 3.6.** Μια μετρήσιμη συνάρτηση  $F : \{0, 1, \dots, N\} \times E \rightarrow \mathbb{R}$  θα λέγεται **Υπεραρμονική (superharmonic)** αν και μόνο αν

$$TF(n, x) \leq F(n, x) \quad (3.61)$$

για κάθε  $(n, x) \in \{0, 1, \dots, N\} \times E$ .

Θεωρείται ως δεδομένο στην (3.61) ότι το  $TF(n, x)$  είναι καλά ορισμένο και δηλαδή  $F(n+1, X_{n+1}) \in \mathcal{L}^1(P_{n, x})$ , για κάθε  $(n, x)$ . Επιπλέον, αν  $F(n+1$



$k, X_{n+k}) \in \mathcal{L}^1(P_{n,x})$  για κάθε  $0 \leq k \leq N - n$  και κάθε  $(n, x)$ , τότε μπορούμε να διαπιστώσουμε τη μαρκοβιανή ιδιότητα που έχει η παρακάτω πρόταση για τη στοχαστική μορφή των υπεραρμονικών συναρτήσεων:

**Πρόταση 3.7.** Η  $F$  θα είναι υπεραρμονική, αν και μόνο αν, η  $\{F(n+k, X_{n+k})\}_{0 \leq k \leq N-n}$  είναι supermartingale υπό το μέτρο  $P_{n,x}$  για κάθε  $(n, x) \in \{0, 1, \dots, N-1\} \times E$ .

Η απόδειξη της παραπάνω πρότασης είναι αρκετά απλή και θα δοθεί σε μια πιο γενική περίπτωση, αυτή της πρότασης 3.9, παρακάτω.

Ορίζουμε το σύνολο που συμβολίζει το «πότε συνεχίζουμε»

$$C = \{(n, x) \in \{0, 1, \dots, N\} \times E : V(n, x) > G(n, x)\} \quad (3.62)$$

και το σύνολο που συμβολίζει το «πότε σταματάμε»

$$D = \{(n, x) \in \{0, 1, \dots, N\} \times E : V(n, x) = G(n, x)\}. \quad (3.63)$$

Ορίζουμε τη στιγμή πρώτης άφιξης  $\tau_D$  στο  $D$

$$\tau_D = \inf\{n \leq k \leq N : (n+k, X_{n+k}) \in D\}, \quad (3.64)$$

υπό το μέτρο  $P_{n,x}$ , όπου  $(n, x) \in \{0, 1, \dots, N\} \times E$ .

**Θεώρημα 3.8. (Πεπερασμένος Ορίζοντας: Η χρονικά ανομοιογενής περίπτωση)** Θεωρούμε το πρόβλημα βέλτιστης διακοπής (3.55) αφού υποθέσουμε ότι ισχύει η συνθήκη (3.54). Τότε, θα ισχύουν τα εξής:

(i) Η συνάρτηση-αξίας ικανοποιεί τις εξισώσεις *Wald-Bellman*

$$V^N(n, x) = \max(G(n, x), TV^N(n, x)) \quad (3.65)$$

για  $n = N-1, \dots, 1, 0$  όπου  $TV^N(N-1, x) = E_{N-1, x}[G(N, X_N)]$  και  $x \in E$ .

- (ii) Ο χρόνος διακοπής  $\tau_D$  είναι βέλτιστος για το (3.55).
- (iii) Αν  $\tau_*$  είναι ένας άλλος βέλτιστος χρόνος διακοπής για το (3.55), τότε  $\tau_D \leq \tau_*$  P-σ.β. για κάθε  $(n, x) \in \{0, 1, \dots, N\} \times E$ .
- (iv) Η συνάρτηση-αξίας  $V^N$  είναι η μικρότερη υπεραρμονική συνάρτηση που κυριαρχεί τη συνάρτηση κέρδους  $G$  στο  $\{0, 1, \dots, N\} \times E$ .
- (v) Η σταματημένη διαδικασία  $\left\{ V^N((n+k) \wedge \tau_D, X_{(n+k) \wedge \tau_D}) \right\}_{0 \leq k \leq N-n}$  είναι martingale υπό το μέτρο  $P_{n, x}$  για κάθε  $(n, x) \in \{0, 1, \dots, N\} \times E$ .

**6.** Στη συνέχεια, θα μελετήσουμε το πρόβλημα (3.37) για  $N = \infty$ , το οποίο υπενθυμίζουμε ότι γράφεται ως:

$$V(x) = \sup_{\tau} E_x[G(X_{\tau})], \quad (3.66)$$

όπου ο  $\tau$  είναι ένας χρόνος διακοπής της  $\mathbf{X}$  και  $P_x(X_0 = x) = 1$  για  $x \in E$ .

Ορίζουμε το σύνολο που συμβολίζει το «πότε συνεχίζουμε»

$$C = \{x \in E : V(x) > G(x)\} \quad (3.67)$$

και το σύνολο που συμβολίζει το «πότε σταματάμε»

$$D = \{x \in E : V(x) = G(x)\}. \quad (3.68)$$

Επίσης, ορίζουμε τον χρόνο πρώτης άφιξης  $\tau_D$  στο  $D$ , θέτοντας

$$\tau_D = \inf\{t \geq 0 : X_t \in D\}. \quad (3.69)$$

7. Ο ακόλουθος ορισμός παίζει στοιχειώδη ρόλο στην εύρεση λύσης για το πρόβλημα (3.66). Παρατηρούμε ότι ο Ορισμός 3.6 είναι μια ειδική περίπτωση του παρακάτω.

**Ορισμός 3.9.** Μια μετρήσιμη συνάρτηση  $F : E \rightarrow \mathbb{R}$  θα λέγεται **Υπεραρμονική (superharmonic)** αν και μόνο αν

$$TF(x) \leq F(x) \quad (3.70)$$

για κάθε  $x \in E$ .

Θεωρείται ως δεδομένο στην (3.70) ότι το  $TF(x)$  είναι καλά ορισμένο και δηλαδή  $F(X_1) \in \mathcal{L}^1(P_x)$ , για κάθε  $x \in E$ . Επιπλέον, αν  $F(X_n) \in \mathcal{L}^1(P_x)$  για κάθε  $n \geq 0$  και κάθε  $x \in E$ , τότε μπορούμε να διαπιστώσουμε τη μαρκοβιανή ιδιότητα που έχει η παρακάτω πρόταση για τη στοχαστική μορφή των υπεραρμονικών συναρτήσεων:

**Πρόταση 3.10.** Η  $F$  θα είναι υπεραρμονική, αν και μόνο αν, η  $\{F(X_n)\}_{n \geq 0}$  είναι supermartingale υπό το μέτρο  $P_x$  για κάθε  $x \in E$ .

*Απόδειξη.* Η απόδειξη της σχέσης ισοδυναμίας είναι απλή. Έστω ότι η  $F$  είναι υπεραρμονική. Τότε, ισχύει η (3.70) για κάθε  $x \in E$  και άρα από τη μαρκοβιανή ιδιότητα θα έχουμε ότι:

$$\begin{aligned} TF(X_n) &= E_{X_n}[F(X_1)] = E_x[F(X_1) \circ \theta_n \mid \mathcal{F}_n] \\ &= E_x[F(X_{n+1}) \mid \mathcal{F}_n] \leq F(X_n), \end{aligned} \quad (3.71)$$

για κάθε  $n \geq 0$ , αποδεικνύοντας την supermartingale ιδιότητα της  $\{F(X_n)\}_{n \geq 0}$  υπό το μέτρο  $P_x$  για κάθε  $x \in E$ . Από την άλλη, αν η  $\{F(X_n)\}_{n \geq 0}$  είναι supermartingale υπό το μέτρο  $P_x$  για κάθε  $x \in E$ , τότε η τελευταία ανισότητα της (3.71) θα ισχύει για κάθε  $n \geq 0$ , άρα και για  $n = 0$ . Παίρνοντας και τη μέση τιμή  $E_x$  στα δύο μέλη, προκύπτει η (3.70). Συνεπώς, η  $F$  θα είναι supermartingale.

□

8. Στην περίπτωση του άπειρου ορίζοντα δεν είναι ανάγκη να χειριζόμαστε τη χρονικά ανομοιογενή περίπτωση διαφορετικά από τη χρονικά ομοιογενή, σε αντίθεση με την περίπτωση του πεπερασμένου ορίζοντα. Αυτό συμβαίνει διότι ο χώρος καταστάσεων  $E$  μπορεί να είναι ούτως ή άλλως πιο γενικός (π.χ. δύο διαστάσεων) και η μετάβαση από την χρονικά ανομοιογενή  $\{X_n\}_{n \geq 0}$ , στη χρονικά ομοιογενή  $\{(n, X_n)\}_{n \geq 0}$ , δε θα επηρεάσει το σύνολο χρόνων πάνω στο οποίο λαμβάνουν τις τιμές τους οι χρόνοι διακοπής της  $\mathbf{X}$ .

**Θεώρημα 3.11. (Άπειρος Ορίζοντας)** Θεωρούμε το πρόβλημα βέλτιστης διακοπής (3.66), ενώ η συνθήκη (3.36) ισχύει. Τότε, θα ισχύουν τα εξής:

(i) Η συνάρτηση-αξίας  $V$  ικανοποιεί την εξίσωση *Wald-Belman*

$$V(x) = \max(G(x), TV(x)), \quad (3.72)$$

για  $x \in E$ . Επίσης, υποθέτουμε όπου χρειάζεται παρακάτω πως

$$P_x(\tau_D < \infty) = 1, \quad (3.73)$$

για κάθε  $x \in E$ . Τότε, θα έχουμε ότι

(ii) Ο χρόνος διακοπής  $\tau_D$  είναι βέλτιστος για το (3.66).

(iii) Αν  $\tau_*$  είναι ένας βέλτιστος χρόνος διακοπής για το (3.66), τότε  $\tau_D \leq \tau_*$   $P_x$ -σ.β. για κάθε  $x \in E$ .

- (iv) Η συνάρτηση-αξίας  $V$  είναι η ελάχιστη υπεραρμονική συνάρτηση που κυριαρχεί της συνάρτησης κέρδους  $G$  στον  $E$ .
- (v) Η σταματημένη διαδικασία  $\{V(X_{n \wedge \tau_D})\}_{n \geq 0}$  είναι martingale υπό το μέτρο  $P_x$  για κάθε  $x \in E$ .

**Πόρισμα 3.12. (Επαναληπτική Μέθοδος)** Υπό την αρχική υπόθεση του Θεωρήματος (3.11), θα έχουμε ότι

$$V(x) = \lim_{n \rightarrow \infty} Q^n G(x), \quad (3.74)$$

για κάθε  $x \in E$ .

*Απόδειξη.* Η απόδειξη του πορίσματος προκύπτει από την (3.44) και το Θεώρημα 3.4.  $\square$

Η σχέση (3.74) μάς δίνει μία κατασκευαστική μέθοδο εύρεσης της συνάρτησης-αξίας  $V$ . (Παρατηρούμε ότι η  $n \mapsto Q^n G(x)$  είναι αύξουσα στο  $\{0, 1, \dots\}$  για κάθε  $x \in E$ .)

**9.** Είδαμε στο Θεώρημα 3.5 και στο Θεώρημα 3.8 ότι οι εξισώσεις *Wald-Bellman* (3.44) και (3.65) χαρακτηρίζουν τη συνάρτηση-αξίας  $V^N$  όταν ο ορίζοντας  $N$  είναι πεπερασμένος (δηλαδή, αυτές οι εξισώσεις δεν μπορεί να έχουν άλλη λύση). Αυτό συμβαίνει λόγω του ότι η  $V^N$  ισούται της  $G$  στο «τέλος του χρόνου»  $N$ . Ωστόσο, όταν ο ορίζοντας είναι άπειρος ( $N = \infty$ ), κάτι τέτοιο δεν ισχύει για την *Wald-Bellman* εξίσωση (3.72). Για παράδειγμα, αν η  $G$  είναι σταθερή και ίση με  $c \in \mathbb{R}$ , τότε οποιαδήποτε άλλη σταθερά  $C \in \mathbb{R}$  με  $C > c$  θα αποτελεί μια λύση για το (3.72). Από την άλλη, η (3.72) μάς λέει ότι κάθε λύση αυτής της εξίσωσης είναι υπεραρμονική και κυριαρχεί της  $G$ . Από την (iv) του Θεωρήματος 3.11 ξέρουμε ότι μια ελαχιστική λύση της (3.72) θα συμπίπτει με τη συνάρτηση-αξίας. Αυτή η «ελαχιστική συνθήκη» μπορεί να αντικατασταθεί από μία εναλλακτική συνθήκη, όπως φαίνεται στο παρακάτω Θεώρημα.

**Θεώρημα 3.13.** Υπό την υπόθεση του Θεωρήματος 3.11, έστω ότι  $F : E \rightarrow \mathbb{R}$  είναι μια συνάρτηση που λύνει τις εξισώσεις *Wald-Bellman*

$$F(x) = \max(G(x), TF(x)), \quad (3.75)$$

για  $x \in E$  (επίσης, έχουμε υποθέσει ότι η  $F$  είναι μετρήσιμη και  $F(X_1) \in \mathcal{L}^1(P_x)$  για κάθε  $x \in E$ ). Επιπλέον, υποθέτουμε ότι η  $F$  ικανοποιεί

$$E \left[ \sup_{n \geq 0} F(X_n) \right] < \infty. \quad (3.76)$$

Τότε, η  $F$  θα ισούται με τη συνάρτηση-αξίας  $V$ , αν και μόνο αν η ακόλουθη «οριακή συνθήκη στο άπειρο» ισχύει:

$$\limsup_{n \rightarrow \infty} F(X_n) = \limsup_{n \rightarrow \infty} G(X_n), \quad P_x - \sigma.β. \quad (3.77)$$

για κάθε  $x \in E$ . Σημειώνουμε, ότι σε αυτήν την περίπτωση το  $\limsup$  του αριστερού μέλους της (3.77) ισούται με το  $\liminf$ . Δηλαδή, η ακολουθία  $\{F(X_n)\}_{n \geq 0}$  συγκλίνει σχεδόν βεβαίως ως προς το μέτρο  $P_x$  για κάθε  $x \in E$ .

**10.** Στο τέλος αυτής της υποενότητας θα μελετήσουμε ένα πιο γενικό πρόβλημα, όπου σε κάθε βήμα  $n$  θα πρέπει να «πληρώνουμε» ένα κόστος  $c(X_n)$  προκειμένου να μπορέσουμε να μεταβούμε στην επόμενη χρονική στιγμή και τελικά να αποφασίσουμε αν σταματάμε ή όχι.

Έστω  $\alpha \in (0, 1]$  και οι (φραγμένες) μετρήσιμες συναρτήσεις  $g : E \rightarrow \mathbb{R}$  και  $c : E \rightarrow \mathbb{R}_+$ . Τότε, θεωρούμε το πρόβλημα βέλτιστης διακοπής

$$V(x) = \sup_{\tau} E_x \left[ \alpha^\tau g(X_\tau) - \sum_{k=1}^{\tau} \alpha^{k-1} c(X_{k-1}) \right], \quad (3.78)$$

όπου ο  $\tau$  είναι χρόνος διακοπής της  $\mathbf{X}$  και  $P_x(X_0 = x) = 1$ .

Η τιμή της  $c(x)$ , όπως είπαμε και παραπάνω, ερμηνεύεται ως το κόστος που πρέπει να πληρώσουμε για να κάνουμε ένα ακόμη βήμα στην  $\mathbf{X}$ , όταν η  $\mathbf{X}$  είναι ίση με  $x$ . Το άθροισμα στην (3.78) εξ ορισμού ισούται με 0 όταν  $\tau = 0$ .

Θα περιγράψουμε συντόμως πώς το πρόβλημα (3.78) μπορεί να αναλυθεί χρησιμοποιώντας το Θεώρημα 3.11.

Για τον σκοπό αυτόν, ορίζουμε την  $\tilde{\mathbf{X}} = \{\tilde{X}_n\}_{n \geq 0}$ , η οποία συμβολίζει τη Μαρκοβιανή αλυσίδα  $\mathbf{X}$  που φθίνει με ρυθμό  $\alpha$ . Αυτό συνεπάγεται ότι ο τελεστής μετάβασης της  $\tilde{\mathbf{X}}$  θα δίνεται από τον τύπο

$$\tilde{T}F(x) = \alpha TF(x), \quad (3.79)$$

για  $x \in E$ , στις περιπτώσεις που  $F(X_1) \in \mathcal{L}^1(P_x)$ .

Χωρίς να εισέλθουμε σε περαιτέρω λεπτομέρειες, δίνουμε στον αναγνώστη ότι η εξίσωση *Wald-Bellman* (τελικώς) θα λάβει την ακόλουθη μορφή:

$$V(x) = \max(g(x), \alpha TV(x) - c(x)), \quad (3.80)$$

για  $x \in E$ . Επίσης, ο ορισμός για τις υπεραρμονικές συναρτήσεις τώρα εφαρμόζεται ως εξής:

$$\alpha TF(x) - c(x) \leq F(x), \quad (3.81)$$

για  $x \in E$ . Αποδεικνύεται ότι η  $F$  ικανοποιεί την (3.81), αν και μόνο αν, η  $(x, \alpha) \mapsto F(x) - \alpha$  είναι υπεραρμονική ως προς τη Μαρκοβιανή αλυσίδα  $\tilde{\mathbf{Z}} = (\tilde{\mathbf{X}}, \tilde{\mathbf{I}})$ , όπου η  $\tilde{\mathbf{I}} = \{I_n\}_{n \geq 0}$  δίνεται από τον τύπο

$$\tilde{I}_n = \alpha + \sum_{k=1}^n c(\tilde{X}_{k-1}). \quad (3.82)$$

## 3.2 Το Πρόβλημα σε Συνεχή Χρόνο

Για το πρόβλημα σε συνεχή χρόνο, θα αναφέρουμε μόνο την περίπτωση των Martingales (και όχι των Μαρκοβιανών Αλυσίδων) για να δείξουμε τη βασική ιδέα της γενίκευσης από τον διακριτό χρόνο στον συνεχή. Για περισσότερες σχετικές πληροφορίες με το πρόβλημα βέλτιστης διακοπής σε συνεχή χρόνο, αλλά και για τις μεθόδους επίλυσης ως ένα πρόβλημα ελευθέρου συνόρου, παραπέμπουμε τον αναγνώστη στα [14], [15] και [16].

### 3.2.1 Η περίπτωση των Martingales

Έστω  $G = \{G_t\}_{t \geq 0}$  μια στοχαστική διαδικασία, ορισμένη και προσαρμοσμένη στη διήθηση του χώρου πιθανότητας  $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ . Την  $G_t$  μπορούμε να τη φανταστούμε ως το κέρδος που θα έχουμε (gain) αν σταματήσουμε να παρατηρούμε την  $G$  τη χρονική στιγμή  $t$ . Θα θεωρούμε ότι όλοι οι χρόνοι διακοπής είναι πεπερασμένοι ( $\tau < \infty$  P-σ.β.).

Θεωρούμε ότι η διαδικασία  $G$  είναι δεξιά συνεχής και αριστερά συνεχής ως προς τους χρόνους διακοπής (δηλαδή, αν  $\tau_n$  και  $\tau$  είναι χρόνοι διακοπής, τέτοιοι ώστε  $\tau_n \uparrow \tau$ , όταν  $n \rightarrow \infty$ , τότε  $G_{\tau_n} \rightarrow G_\tau$  P-σ.β., όταν  $n \rightarrow \infty$ ). Επίσης, θεωρούμε ότι η ακόλουθη συνθήκη θα ισχύει πάντα (με  $G_T \equiv 0$ , όταν  $T = \infty$ ):

$$E \left[ \sup_{0 \leq t \leq T} |G_t| \right] < \infty. \quad (3.83)$$

Όπως και στην περίπτωση του διακριτού χρόνου, η παραπάνω συνθήκη μπορεί να αντικατασταθεί από ασθενέστερες συνθήκες και στη θεωρία και στην πράξη. Ωστόσο, κανένα από τα αποτελέσματα δε διαφέρει ουσιωδώς και ως εκ τούτου θα δουλέψουμε με την (3.83).

Επιπλέον, προκειμένου να επικαλεστούμε ένα Θεώρημα ύπαρξης για μια δεξιά συνεχή τροποποίηση μιας δοσμένης supermartingale, θα θεωρούμε ότι η διήθηση  $\{\mathcal{F}_t\}_{t \geq 0}$  είναι δεξιά συνεχής και ότι κάθε  $\mathcal{F}_t$  περιέχει όλα τα P-κενά σύνολα από την  $\mathcal{F}$ . Αυτή είναι μια καθαρά τεχνική προϋπόθεση και η επιβολή της δεν ε-



πηρεάζει κάπως τις ερμηνείες μας σχετικά με το πρόβλημα επιλεκτικής διακοπής.

1. Έστω το πρόβλημα επιλεκτικής διακοπής

$$V_t^T = \sup_{t \leq \tau \leq T} E[G_\tau], \quad (3.84)$$

όπου ο  $\tau$  είναι ένας χρόνος διακοπής με  $0 \leq \tau \leq T$ . Στην (3.84) επιτρέπουμε στο  $T$  να είναι  $+\infty$ , αλλά παρόλα αυτά το supremum το παίρνουμε πάνω σε πεπερασμένους χρόνους διακοπής. Σε αυτήν την περίπτωση θέτουμε  $V_t = V_t^\infty$  για  $t \geq 0$ .

2. Αντίστοιχα με τα αποτελέσματα της Υποενότητας 3.1 (Το Πρόβλημα σε Διακριτό Χρόνο), υπάρχουν δύο τρόποι να αντιμετωπίσουμε το πρόβλημα (3.57). Η πρώτη μέθοδος μας λέει να αντικαταστήσουμε το χρονικό διάστημα  $[0, T]$  με τα σύνολα  $\mathcal{J}_n = \{t_0^n, t_1^n, \dots, t_n^n\}$ , όπου  $\mathcal{J}_n \uparrow \mathcal{J}$  όταν  $n \rightarrow \infty$  και  $\mathcal{J}$  είναι ένα αριθμήσιμο και πυκνό υποσύνολο του  $[0, T]$ , εφαρμόζοντας τα αποτελέσματα της προς τα πίσω επαγωγής για διακριτό χρόνο σε κάθε  $G^n = \{G_{t_i^n}\}_{0 \leq i \leq n}$  και παίρνοντας το όριο για  $n \rightarrow \infty$ . Για την ανάλυση αυτήν, μας είναι χρήσιμο να γνωρίζουμε ότι κάθε χρόνος διακοπής  $\tau$  μπορεί να γραφεί ως το φθίνον όριο μιας ακολουθίας διακριτών χρόνων διακοπής  $\tau_n = \sum_{i=1}^n t_i^n I(t_{i-1}^n \leq \tau < t_i^n)$ , όταν το  $n \rightarrow \infty$ . Αυτή η μέθοδος είναι χρήσιμη όταν θέλουμε να λάβουμε αριθμητικές προσεγγίσεις για τη λύση του προβλήματος, αλλά δε θα κινηθούμε περαιτέρω προς αυτήν την κατεύθυνση. Η δεύτερη μέθοδος σκοπεύει άμεσα στην επέκταση της μεθόδου του essential supremum στην περίπτωση του συνεχούς χρόνου και είναι η μέθοδος με την οποία θα συνεχίσουμε την ανάλυσή μας.

3. Επειδή δεν υπάρχει ουσιαστική διαφορά στο πώς χειριζόμαστε τις περιπτώσεις πεπερασμένου και άπειρου ορίζοντα  $T$ , για λόγους απλούστευσης του συμβολισμού, θέτουμε

$$V_t = V_t^T. \quad (3.85)$$

Για να λύσουμε το (3.84), θα θεωρήσουμε (κατά αναλογία με τα αποτελέσματα για την περίπτωση των martingales σε διακριτό χρόνο) τη διαδικασία  $\mathbf{S} = \{S_t\}_{t \geq 0}$ , που ορίζεται ως

$$S_t = \operatorname{ess\,sup}_{\tau \geq t} E[G_\tau | \mathcal{F}_t], \quad (3.86)$$

όπου ο  $\tau$  είναι χρόνος διακοπής. Στην περίπτωση του πεπερασμένου ορίζοντα απαιτούμε, επίσης,  $\tau \leq T$ . Αποδεικνύεται πως η παραδοχή ότι η διαδικασία  $S$  είναι δεξιά συνεχής δε βλάπτει τη γενικότητα.

Για τους ίδιους λόγους, θεωρούμε τον εξής χρόνο διακοπής:

$$\tau_t = \inf\{s \geq t : S_s = G_s\}, \quad (3.87)$$

για  $t \geq 0$ , όπου  $\inf \emptyset = \infty$  εξ' ορισμού. Στην περίπτωση του πεπερασμένου ορίζοντα  $T$ , στην (3.87) απαιτούμε επίσης  $s \leq T$ .

Όσον αφορά το αρχικό μέρος των Θεωρημάτων 3.1 και 3.3 (την εξίσωση Wald-Bellman), οφείλουμε να παρατηρήσουμε ότι, παρακάτω, το Θεώρημα 3.14 θα συνεπάγεται πως

$$S_t \geq \max(G_t, E[S_s | \mathcal{F}_t]), \quad (3.88)$$

για  $s \geq t$ . Η αντίστροφη ανισότητα, ωστόσο, δεν ισχύει καθολικά. Με απλά λόγια, ο λόγος έγκειται στο ότι, σε αντίθεση με την περίπτωση του διακριτού χρόνου, στον συνεχή χρόνο δεν υφίσταται ελάχιστη μονάδα χρόνου. Συνεπώς, όσο κοντά κι αν βρίσκεται το  $s$  στο  $t$ , οι τιμές  $S_u$  μπορούν να απέχουν πολύ από το  $S_t$  για  $u \in (t, s)$ . Παρόλα αυτά, παρατηρούμε ότι από το παρακάτω Θεώρημα 3.14 συνεπάγεται ότι εξακολουθεί να ισχύει το εξής «ραφινάρισμα» της εξίσωσης Wald-Bellman:

$$S_t = \max(G_t, E[S_{\sigma \wedge \tau_t} | \mathcal{F}_t]), \quad (3.89)$$

για κάθε χρόνο διακοπής  $\sigma$  μεγαλύτερο ή ίσο του  $t$ . Το  $\tau_t$  είναι ο χρόνος διακοπής της σχέσης (3.87).

Τα άλλα τρία μέρη των Θεωρημάτων 3.1 και 3.3 επεκτείνονται στην παρούσα περίπτωση χωρίς σημαντικές διαφοροποιήσεις. Ως εκ τούτου, το πρώτο μέρος του ακόλουθου Θεωρήματος μας λέει ότι τα  $\{S_s\}_{s \geq t}$  και  $\tau_t$  λύνουν το πρόβλημα υπό τη στοχαστική έννοια. Το δεύτερο μέρος του Θεωρήματος, δείχνει ότι αυτό οδηγεί σε λύση του αρχικού προβλήματος (3.84). Το τρίτο μέρος του Θεωρήματος μας παρέχει τη supermartingale ιδιότητα της λύσης.

**Θεώρημα 3.14.** Θεωρούμε το πρόβλημα βέλτιστης διακοπής (3.84), ενώ η συνθήκη (3.83) ισχύει. Επίσης, θεωρούμε ως δεδομένο (όπου χρειάζεται) πως  $P(\tau_t < \infty) = 1$ , όπου  $t \geq 0$  (η συνθήκη ισχύει πάντα στην περίπτωση του πεπερασμένου ορίζοντα  $T$ ). Τότε, για κάθε  $t \geq 0$ , θα έχουμε ότι:

$$S_t \geq E[G_\tau | \mathcal{F}_t], \quad \forall \tau \in \mathfrak{M}_t, \quad (3.90)$$

$$S_t = E[G_{\tau_t} | \mathcal{F}_t], \quad (3.91)$$

όπου με  $\mathfrak{M}_t$  συμβολίζουμε την οικογένεια όλων των χρόνων διακοπής  $\tau$  που είναι μεγαλύτεροι ή ίσοι του  $t$  (και ταυτόχρονα μικρότεροι ή ίσοι του  $T$  για πεπερασμένο ορίζοντα). Επιπλέον, για  $t \geq 0$  δοθέν και σταθερό, θα έχουμε ότι:

- (i) Ο χρόνος διακοπής  $\tau_t$  είναι βέλτιστος για το (3.84).
- (ii) Αν  $\tau_*$  είναι ένας βέλτιστος χρόνος διακοπής για το (3.84), τότε  $\tau_t \leq \tau_*$   $P_x$ -σ.β..
- (iii) Η διαδικασία  $\{S_s\}_{s \geq t}$  είναι η ελάχιστη δεξιά συνεχής supermartingale συνάρτηση που κυριαρχεί της συνάρτησης  $\{G_s\}_{s \geq t}$ .
- (iv) Η σταματημένη διαδικασία  $\{S_{s \wedge \tau_t}\}_{s \geq t}$  είναι μία δεξιά συνεχής martingale.

Τέλος, αν δεν υπήρχε ο περιορισμός  $P(\tau_t < \infty) = 1$  και προέκυπτε ότι  $P(\tau_t = 0) > 0$ , τότε με πιθανότητα 1 δε θα υπήρχε λύση στο πρόβλημα (3.84).

## Ενότητα 4

# Το Πρόβλημα του Ληστή με τα Πολλά Χέρια (Multi-Armed Bandit Problem)

Τα προβλήματα των ληστών με τα πολλά χέρια (multi-armed bandit problems, [18], [19], [20], [21]) υπάγονται στην κατηγορία της ακολουθιακής επιλογής πειραμάτων και σχετίζονται με τα προβλήματα βέλτιστης διακοπής μέσα από το Θεώρημα των Gittins & Jones (1974) [22]. Με λίγα λόγια, το Θεώρημα μας λέει ότι σε κάθε χρονική στιγμή είναι βέλτιστο να διαλέγεις τη δράση με το μεγαλύτερο δείκτη Gittins, είτε το δίλημμα είναι «συνεχίζω ή σταματάω;», που παραπέμπει σε πρόβλημα βέλτιστης διακοπής, είτε είναι «ποιό απ' όλα πρέπει να διαλέξω;», που παραπέμπει σε πρόβλημα multi-armed bandit. Στην παρούσα εργασία, θα αναφέρουμε κάποιες βασικές έννοιες για το πρόβλημα και το εν λόγω Θεώρημα, αλλά κατά κύριο λόγο θα εστιάσουμε σε αλγοριθμικές μεθόδους επίλυσης, όπως είναι ο αλγόριθμος  $\epsilon$ -Greedy και ο UBC (Upper Confidence Bounds). Απώτερος σκοπός θα είναι να καταλήξουμε στον αλγόριθμο Thompson Sampling ο οποίος θα λύσει τις πραγματικές εφαρμογές της επόμενης ενότητας.

## 4.1 Εισαγωγή στο Πρόβλημα

Θεωρούμε ένα πρόβλημα ακολουθιακών αποφάσεων, στο οποίο σε κάθε βήμα υπάρχουν  $K$  πιθανές δράσεις ή επιλογές πειραμάτων. Η επιλογή της δράσης  $j$  έχει ως αποτέλεσμα τη λήψη μιας παρατήρησης από το πείραμα  $j$ , η οποία επιστρέφει την αντίστοιχη αμοιβή. Κάθε παρατήρηση μας δίνει χρήσιμη πληροφορία για τις μελλοντικές δράσεις και επιλογές. Ο σκοπός είναι η μεγιστοποίηση της τρέχουσας αξίας των επιλογών μας, με τρόπο τέτοιο ώστε από ένα σημείο και μετά να επιλέγουμε μόνο εκείνη τη δράση, έστω  $j^*$ , η οποία έχει τη μεγαλύτερη αμοιβή.

Ο όρος 'bandit' («ληστής») προέρχεται από τη μοντελοποίηση τέτοιου είδους προβλημάτων θεωρώντας έναν κουλοχέρη που έχει  $K$  χέρια και το κάθε χέρι έχει μια διαφορετική κατανομή αποδόσεων. Δε γνωρίζουμε ποιο από τα χέρια έχει μεγαλύτερη πιθανότητα επιτυχίας (ή μεγαλύτερη μέση αμοιβή), αλλά δοκιμάζοντας τα διαφορετικά χέρια του κουλοχέρη, λαμβάνουμε επαρκή πληροφορία ώστε να καταλήξουμε στο βέλτιστο. Ωστόσο, η πληροφορία που χρησιμοποιούμε είναι και η αμοιβή μας. Συνεπώς, πρέπει να οριστεί μια καλή ισορροπία ανάμεσα στη λήψη πληροφορίας και στη λήψη καλών αμοιβών. Για παράδειγμα, δεν είναι πάντα βέλτιστο να διαλέγουμε το χέρι που είχε την καλύτερη απόδοση στο παρελθόν, γιατί μπορεί απλά να ήμασταν άτυχοι και να μη βρήκαμε ποτέ το πραγματικό βέλτιστο. Αν έχουμε αρκετές δοκιμές στον ορίζοντά μας και παράλληλα χρειάζονται μόνο λίγες για να ξεκαθαρίσουμε το ζήτημα, αξίζει να αφιερώσουμε μερικές από αυτές τις προσπάθειες στην αναζήτηση του βέλτιστου χεριού. Συνήθως, σε αυτού του είδους τα προβλήματα, υπάρχει μια πρώτη περίοδος λήψης πληροφορίας, μία δεύτερη περίοδος απόρριψης χεριών και μία τρίτη περίοδος που έχουμε καταλήξει στο χέρι και παίζουμε αποκλειστικά αυτό, ώστε να λάβουμε το μέγιστο κέρδος. Το μεγάλο, αυτό, δίλημμα που περιγράψαμε στην παράγραφο, είναι γνωστό και ως "exploration vs exploitation".



Σχήμα 4.1: Πραγματικό παράδειγμα του exploration vs exploitation: Πού θα φάμε;

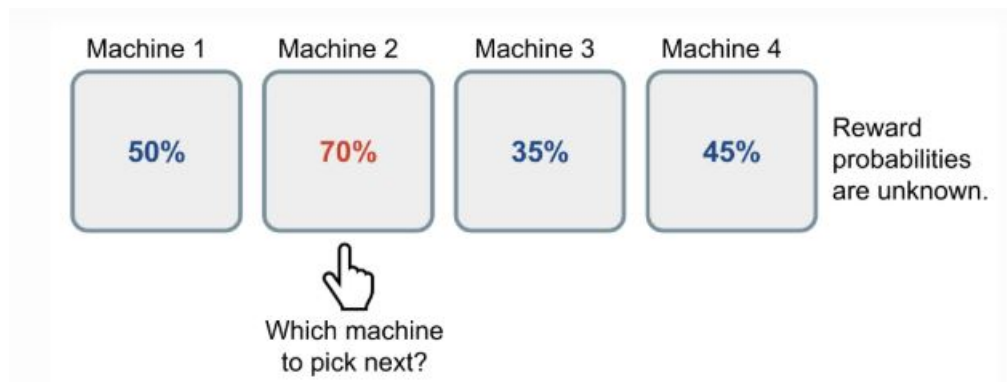
Μια σημαντική εφαρμογή του προβλήματος multi-armed bandit (από εδώ και στο εξής θα το αναφέρουμε ως MAB για συντομία) βρίσκεται στην ιατρική και τις κλινικές δοκιμές, όπου υπάρχουν  $K$  θεραπείες για μια ασθένεια. Οι ασθενείς εμφανίζονται ακολουθιακά και τους δίνεται άμεσα μία από τις θεραπείες. Θεωρείται ως δεδομένο ότι το αποτέλεσμα της θεραπείας (θετικό ή αρνητικό) είναι άμεσο και χρησιμοποιείται ως γνώση για τη θεραπεία του επόμενου ασθενή. Δεν είναι γνωστό ποιά θεραπεία είναι η βέλτιστη και σκοπός μας είναι να θεραπεύσουμε όσο το δυνατόν περισσότερους ασθενείς. Στα αρνητικά τής εν λόγω μεθόδου ανήκει το γεγονός ότι ενδέχεται να χορηγήσουμε μια θεραπεία η οποία δε φαίνεται να είναι βέλτιστη, με μοναδικό σκοπό τη λήψη πληροφορίας που θα χρησιμεύσει σε μετέπειτα ασθενείς.

## 4.2 Bernoulli Bandit

Έστω ότι υπάρχουν  $K$  επιλογές (δράσεις, χέρια στον κουλοχέρη) και κάθε επιλογή επιστρέφει είτε επιτυχία, είτε αποτυχία (όπως το παράδειγμα με τις θεραπείες που αναφέραμε στην προηγούμενη παράγραφο). Η δράση  $k \in \{1, 2, \dots, K\}$  επιτυγχάνει με πιθανότητα  $\theta_k \in [0, 1]$ . Οι πιθανότητες  $(\theta_1, \theta_2, \dots, \theta_K)$  είναι άγνωστες στον παίχτη, αλλά σταθερές ως προς τον χρόνο και συνεπώς, ο παίχτης μπορεί να τις μάθει πειραματίζοντας. Ο σκοπός είναι να μεγιστοποιήσει το άθροισμα των επιτυχιών σε έναν χρονικό ορίζοντα  $T$ , όπου το  $T$  είναι αρκετά

μεγαλύτερο του  $K$ .

Τα χέρια, σε αυτό το πρόβλημα, μπορεί να αναπαριστούν διαφορετικές διαφημίσεις που εμφανίζονται σε μία ιστοσελίδα. Οι χρήστες που επισκέπτονται τη σελίδα, έρχονται σε επαφή με κάποια από τις διαφημίσεις για ένα προϊόν. Η επιτυχία ορίζεται είτε ως το "click" στη διαφήμιση, είτε ως πώληση του προϊόντος. Αντίστοιχα, τα  $\theta_k$  αναπαριστούν το ποσοστό των χρηστών που είδε τη διαφήμιση και είτε πάτησε "click", είτε αγόρασε το προϊόν. Η ιστοσελίδα ελπίζει στην εύρεση της βέλτιστης διαφήμισης και της μεγιστοποίησης των πωλήσεων.



Σχήμα 4.2: Απεικόνιση του πώς δουλεύει το Bernoulli MAB. Οι πιθανότητες επιτυχίας είναι άγνωστες στον παίχτη.

Μια «αφελής» προσέγγιση στο πρόβλημα θα ήταν να ορίσουμε κάποιες προκαθορισμένες χρονικές περιόδους για εξερεύνηση και σε κάθε περίοδο να διαλέγαμε ένα χέρι από την ομοιόμορφη κανονική. Στις επόμενες περιόδους, θα σκοπεύαμε να διαλέξουμε τα επιτυχημένα χέρια, χρησιμοποιώντας τη γνώση που αποκομίσαμε από την εξερεύνηση. Ωστόσο, μία τέτοια προσέγγιση μπορεί να γίνει εξαιρετικά «σπάταλη» ακόμα και για το πιο απλό Bernoulli MAB και να αποτύχει εντελώς για πιο σύνθετα προβλήματα.

### 4.2.1 Ορισμός του Προβλήματος

Το Bernoulli MAB μπορεί να περιγραφεί ως μία δυάδα  $\langle \mathcal{A}, \mathcal{R} \rangle$ , όπου:

- Έχουμε  $K$  χέρια με πιθανότητες επιτυχίας  $\{\theta_1, \theta_2, \dots, \theta_K\}$ .
- Σε κάθε βήμα  $t$ , διαλέγουμε ένα χέρι και λαμβάνουμε την αμοιβή  $r$ .
- Με  $\mathcal{A}$  συμβολίζουμε το σύνολο των δράσεων, όπου η κάθε μία δράση αντιστοιχεί στην επιλογή ενός χεριού. Η αξία της δράσης  $\alpha$  δίνεται από την δεσμευμένη μέση τιμή:

$$Q(\alpha) = E[r | \alpha] = \theta. \quad (4.1)$$

Αν τη χρονική στιγμή  $t$  επιλέξαμε το χέρι  $i$ , τότε  $Q(\alpha_t) = \theta_i$ .

- Με  $\mathcal{R}$  συμβολίζουμε τη συνάρτηση αμοιβής. Στην περίπτωση του Bernoulli MAB, η αμοιβή παρατηρείται πιθανοτικά. Τη χρονική στιγμή  $t$ , η αμοιβή  $r_t = \mathcal{R}(\alpha_t)$  θα είναι ίση με 1, με πιθανότητα  $Q(\alpha_t)$  και 0, διαφορετικά.

Σκοπός μας είναι να μεγιστοποιήσουμε την αθροιστική αμοιβή  $\sum_{t=1}^T r_t$ . Εάν γνωρίζουμε τη βέλτιστη επιλογή (χέρι) με τη βέλτιστη αμοιβή, το πρόβλημα είναι ισοδύναμο με το να ελαχιστοποιήσουμε την ενδεχόμενη ζημιά που θα προέκυπτε αν δεν επιλέγαμε το σωστό χέρι.

Η βέλτιστη πιθανότητα επιτυχίας, έστω  $\theta^*$ , της βέλτιστης επιλογής, έστω  $\alpha^*$ , θα είναι:

$$\theta^* = Q(\alpha^*) = \max_{\alpha \in \mathcal{A}} Q(\alpha) = \max_{1 \leq i \leq K} \theta_i. \quad (4.2)$$

Η συνάρτηση ζημίας θα είναι η συνολική «μετάνοια» ("regret") που θα είχαμε αν δεν επιλέγαμε τα βέλτιστα χέρια μέχρι και τη χρονική στιγμή  $T$ :

$$L_T = E \left[ \sum_{t=1}^T (\theta^* - Q(\alpha_t)) \right]. \quad (4.3)$$



Οι στρατηγικές για το MAB βασίζονται στο πώς θα προσεγγίσουμε την εξερεύνησή μας:

- Καμμία εξερεύνηση. Η πιο αφελής και κακή στρατηγική.
- Τυχαία εξερεύνηση.
- Έξυπνη εξερεύνηση με προτίμηση στην αβεβαιότητα.

## 4.2.2 Ο Αλγόριθμος $\epsilon$ -Greedy

Ο αλγόριθμος  $\epsilon$ -Greedy επιλέγει πάντα το χέρι που -μέχρι τώρα- ξέρει ότι είναι βέλτιστο και ενίοτε επιτρέπει τυχαία εξερεύνηση. Η αξία της εκάστοτε δράσης εκτιμάται με βάση την προϊστορία, ως ο μέσος όρος των αμοιβών που έχει λάβει η εν λόγω δράση μέχρι τη χρονική στιγμή  $t$ . Πιο συγκεκριμένα,

$$\hat{Q}_t(\alpha) = \frac{1}{N_t(\alpha)} \sum_{s=1}^t r_s \mathbb{1}[\alpha_s = \alpha], \quad (4.4)$$

όπου με  $\mathbb{1}[\alpha_s = \alpha]$  συμβολίζουμε τη δείκτρια συνάρτηση για τις περιπτώσεις που  $\alpha_s = \alpha$  και με  $N_t(\alpha)$  τις φορές που επιλέξαμε τη δράση  $\alpha$ , έτσι ώστε  $N_t(\alpha) = \sum_{s=1}^t \mathbb{1}[\alpha_s = \alpha]$ .

Σύμφωνα με τον  $\epsilon$ -Greedy, με μια μικρή πιθανότητα  $\epsilon$  επιτρέπουμε να πραγματοποιήσουμε τυχαία μια δράση. Όλες τις άλλες φορές, με πιθανότητα  $1 - \epsilon$  διαλέγουμε τη βέλτιστη δράση που ξέρουμε μέχρι εκείνη τη στιγμή:

$$\hat{\alpha}^* = \arg \max_{\alpha \in \mathcal{A}} \hat{Q}_t(\alpha). \quad (4.5)$$

### 4.2.3 Ο Αλγόριθμος Upper Confidence Bounds (UCB)

Η τυχαία εξερεύνηση μας δίνει την ευκαιρία να δοκιμάσουμε επιλογές για τις οποίες δε γνωρίζαμε πολλά. Ωστόσο, λόγω της τυχειότητας, είναι πιθανό να καταλήξουμε να δοκιμάζουμε μια δράση για την οποία είχαμε ήδη επιβεβαιώσει πως είναι κακή. Για να αποφύγουμε τέτοιες μη αποδοτικές εξερευνήσεις, η μία προσέγγιση θα ήταν να μειώνουμε το  $\epsilon$  με τον χρόνο και η άλλη θα ήταν να δείχναμε μια προτίμηση προς τις επιλογές που δεν έχουμε εξερευνήσει επαρκώς και δεν έχουμε ασφαλή εκτίμηση για τις αξίες τους. Με άλλα λόγια, ο αλγόριθμος προτιμάει τις δράσεις που έχουν περισσότερη πιθανότητα να είναι βέλτιστες.

Ο αλγόριθμος UCB μετρά αυτήν τη δυνατότητα ορίζοντας ένα άνω φράγμα εμπιστοσύνης (upper confidence bound) για την αξία της αμοιβής, έστω  $\hat{U}_t(\alpha)$ , έτσι ώστε η πραγματική αξία να βρίσκεται κάτω από αυτήν την τιμή με μεγάλη πιθανότητα:

$$Q(\alpha) \leq \hat{Q}_t(\alpha) + \hat{U}_t(\alpha). \quad (4.6)$$

Το άνω φράγμα είναι φθίνουσα συνάρτηση του  $N_t(\alpha)$ . Δηλαδή, όσο μεγαλώνει το  $N_t(\alpha)$  και άρα γινόμαστε πιο σίγουροι για την αξία της δράσης  $\alpha$ , το  $\hat{U}_t(\alpha)$  μικραίνει.

Στον αλγόριθμο UCB διαλέγουμε πάντα την πιο «άπληστη» δράση:

$$\alpha_t^{\text{UCB}} = \arg \max_{\alpha \in \mathcal{A}} [\hat{Q}_t(\alpha) + \hat{U}_t(\alpha)]. \quad (4.7)$$

Πλέον, το ερώτημα είναι πώς θα βρούμε το άνω φράγμα εμπιστοσύνης. Αυτό το ερώτημα απαντάται μέσω της ανισότητας του **Hoeffding**, η οποία μας λέει ότι για  $X_1, X_2, \dots, X_t$  ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο  $[0,1]$  και  $u > 0$ , έχουμε ότι:

$$P(E[X] > \bar{X}_t + u) \leq e^{-2tu^2}, \quad (4.8)$$

όπου  $\bar{X}_t = \frac{1}{t} \sum_{s=1}^t X_s$ .

Εφαρμόζοντάς την, έπεται ότι

$$P(Q(\alpha) > \hat{Q}_t(\alpha) + U_t(\alpha)) \leq e^{-2tU_t(\alpha)^2}. \quad (4.9)$$

Το  $e^{-2tU_t(\alpha)^2}$  ουσιαστικά θέλουμε να είναι μια πολύ μικρή πιθανότητα, έστω  $p$ . Δηλαδή,  $e^{-2tU_t(\alpha)^2} = p$  και τελικά:

$$U_t(\alpha) = \sqrt{\frac{-\log p}{2N_t(\alpha)}}. \quad (4.10)$$

Αξίζει να αναφέρουμε την ύπαρξη κάποιων ευριστικών παραλλαγών του UCB, όπως είναι ο UCB1 και ο Bayesian UCB. Στον πρώτο, η μόνη διαφορά είναι ότι η πιθανότητα  $p$  μικραίνει με τον χρόνο, έτσι ώστε να είμαστε πιο σίγουροι για τις εκτιμήσεις μας όσο μεγαλώνει ο αριθμός των επαναλήψεων. Πιο συγκεκριμένα, στον UCB1, ισχύει ότι  $p = t^{-4}$ . Αντίστοιχα, στον Bayesian UCB, μπορούμε να χρησιμοποιήσουμε κάποια πρότερη γνώση για τα άνω φράγματα. Για παράδειγμα αν πιστεύουμε πως οι κατανομές των αμοιβών είναι Gaussian και μας ενδιαφέρει ένα 95% διάστημα εμπιστοσύνης, θα μπορούσαμε να θέσουμε το  $\hat{U}_t(\alpha)$  να είναι ίσο με δύο φορές την τυπική απόκλιση.

### 4.3 Thompson Sampling

Η δειγματοληψία κατά Thompson (Thompson Sampling ή, από εδώ και στο εξής, απλώς "TS"), η οποία είναι γνωστή και ως *ύστερη δειγματοληψία και αντιστοιχία πιθανότητας*, αρχικά προτάθηκε από τον Thompson το 1933 ως μέθοδος επίλυσης ενός MAB 2 χεριών, με εφαρμογή στις κλινικές δοκιμές. Ωστόσο, δεν ήταν καθόλου δημοφιλής στην ακαδημαϊκή κοινότητα μέχρι πολύ πρόσφατα, όπου ο Scott το 2010 ([23]) και οι Chapelle & Li το 2011 ([24]) έδειξαν την πολύ ισχυρή εμπειρική απόδοση του αλγορίθμου. Από τότε, η βιβλιογραφία στο TS αυξήθηκε αισθητά, ενώ διάφορες προσαρμοσμένες εκδοχές του αλγορίθμου

χρησιμοποιούνται σε πολλούς τομείς, όπως η εμπορία, η βελτιστοποίηση ιστοσελίδων, το Monte Carlo tree search, η διαφήμιση στο internet κ.ά. Επίσης, ο TS χρησιμοποιείται από πολλές εταιρείες, όπως η Adobe, Amazon, Facebook, Google, LinkedIn, Microsoft, Netflix, Twitter. Ενδιαφέρον σχετικά με το Thompson Sampling παρουσιάζουν τα [25], [26], στα οποία παραπέμπουμε τον αναγνώστη για περαιτέρω ανάγνωση.

Ο TS βασίζεται σε μια απλή ιδέα που όμως δουλεύει εξαιρετικά όταν πρόκειται να λύσει ένα MAB.



Σχήμα 4.3: Υπάρχει ένα ολόκληρο (χιουμοριστικό) Θεώρημα που εξηγεί πώς εφευρέθηκε ο αλγόριθμος του Thompson όσο εκείνος έφαχνε πού να δώσει πάσα με την μπάλα. Το βρήκα έξυπνο και το έκλειψα.

Σε κάθε βήμα, θέλουμε να διαλέξουμε μία δράση  $\alpha$  σύμφωνα με την πιθανότητα ότι η  $\alpha$  είναι η βέλτιστη:

$$\pi(\alpha | \mathbf{h}_t) = P(Q(\alpha) > Q(\alpha'), \forall \alpha \neq \alpha' | \mathbf{h}_t) = E_{\mathcal{R} | \mathbf{h}_t} \left[ \mathbb{1} \left[ \alpha = \arg \max_{\alpha \in \mathcal{A}} Q(\alpha) \right] \right], \quad (4.11)$$

όπου  $\pi(\alpha | h_t)$  είναι η πιθανότητα να διαλέξουμε τη δράση  $\alpha$  δεδομένης της προϊστορίας  $h_t$ .

Για το Bernoulli Bandit που αναλύαμε στην υποενότητα 4.2, είναι φυσικό να θεωρήσουμε ότι το  $Q(\alpha)$  ακολουθεί κατανομή Beta, εφόσον το  $Q(\alpha)$  στην ουσία είναι η παράμετρος  $\theta$  της αντίστοιχης Bernoulli κατανομής και η Beta είναι συζυγής prior για την Bernoulli.

Ως πρώτο βήμα, αρχικοποιούμε τις παραμέτρους  $\alpha$  και  $\beta$  της εκάστοτε Beta κατανομής, με βάση την πρότερη πληροφορία που έχουμε για την απόδοση κάθε δράσης. Για παράδειγμα,

- $\alpha = 1$  και  $\beta = 1$ : περιμένουμε η πιθανότητα επιτυχίας της δράσης να είναι 50%, χωρίς να είμαστε καθόλου σίγουροι.
- $\alpha = 1000$  και  $\beta = 9000$ : πιστεύουμε πραγματικά πως η πιθανότητα επιτυχίας της δράσης θα είναι κοντά στο 10%.

Σε κάθε βήμα  $t$ , δειγματίζουμε μία αναμενόμενη αμοιβή  $\hat{Q}(\alpha)$  από την πρότερη κατανομή  $\text{Beta}(\alpha_i, \beta_i)$  για κάθε δράση  $i$ .

Στη συνέχεια, επιλέγουμε την καλύτερη δράση από το δείγμα μας, έτσι ώστε  $\alpha_t^{TS} = \arg \max_{\alpha \in \mathcal{A}} \hat{Q}(\alpha)$ .

Αφού διαλέξουμε την καλύτερη δράση, την εφαρμόζουμε («τραβάμε το χέρι στον κουλοχέρη»), παρατηρούμε την πραγματική αμοιβή που επιστρέφει ως αποτέλεσμα (0 ή 1) και ανανεώνουμε την πρότερη Beta κατανομή της εν λόγω δράσης, σύμφωνα με τον κανόνα:

$$\begin{aligned} \alpha_i &\leftarrow \alpha_i + r_t \mathbb{1}[\alpha_t^{TS} = \alpha_i] \\ \beta_i &\leftarrow \beta_i + (1 - r_t) \mathbb{1}[\alpha_t^{TS} = \alpha_i] \end{aligned} \quad (4.12)$$

---

**Algorithm 2** BernTS( $K, \alpha, \beta$ )

---

```
1: for  $t = 1, 2, \dots$  do
2:   #sample model:
3:   for  $k = 1, \dots, K$  do
4:     Sample  $\hat{\theta}_k \sim \text{beta}(\alpha_k, \beta_k)$ 
5:   end for
6:
7:   #select and apply action:
8:    $x_t \leftarrow \text{argmax}_k \hat{\theta}_k$ 
9:   Apply  $x_t$  and observe  $r_t$ 
10:
11:  #update distribution:
12:   $(\alpha_{x_t}, \beta_{x_t}) \leftarrow (\alpha_{x_t} + r_t, \beta_{x_t} + 1 - r_t)$ 
13: end for
```

---

Σχήμα 4.4: Ο ψευδοκώδικας για το Thompson Sampling σε προβλήματα Bernoulli Bandit.

## 4.4 Παράδειγμα Bernoulli Bandit

Έχοντας περιγράψει το πρόβλημα του Bernoulli MAB και του αλγορίθμου Thompson Sampling, θα παρουσιάσουμε ένα απλό παράδειγμα για να δείξουμε την αποτελεσματικότητα του TS.

Έστω ότι μια εταιρεία θέλει να προωθήσει το προϊόν της και έχει πέντε (5) διαφορετικές πιθανές online διαφημίσεις για αυτό. Η επιτυχία κάθε διαφήμισης ακολουθεί Bernoulli κατανομή και έστω ότι οι παράμετροι είναι  $w_1 = 0.2, w_2 = 0.4, w_3 = 0.5, w_4 = 0.3, w_5 = 0.7$ . Με άλλα λόγια, αν η διαφήμιση 2 εμφανιστεί στα μέσα κοινωνικής δικτύωσης ενός χρήστη, αυτός έχει 40% πιθανότητα να αγοράσει το προϊόν. Αυτές οι παράμετροι είναι άγνωστες στην εταιρεία και σκοπός της είναι να μεγιστοποιήσει τις πωλήσεις της χωρίς να χάσει αρκετά λεφτά σε άσκοπες διαφημίσεις μέχρι να βρει τη βέλτιστη.

Κατασκευάζουμε μια συνάρτηση στην R με όνομα `my_thompson_bernoullibandit`,

η οποία δέχεται ως ορίσματα:

- τον χρονικό ορίζοντα του πειράματος (horizon), δηλαδή το πόσες διαφημίσεις μπορούμε να θυσιάσουμε στο βωμό της εύρεσης της βέλτιστης.
- τον αριθμό των προσομοιώσεων (sims), για επιβεβαίωση της αποτελεσματικότητας του αλγορίθμου.
- τον αριθμό των χεριών του κουλοχέρη (arms).
- τις αντίστοιχες πιθανότητες επιτυχίας (weights).

Η συνάρτηση επιστρέφει μία λίστα με δύο στοιχεία:

- το διάνυσμα των μέσων αμοιβών για κάθε προσομοίωση (MeanRewards).
- έναν πίνακα με το πλήθος των φορών που επιλέχθηκε το χέρι  $k$  στην προσομοίωση  $j$  (Counter).

```
#thompson_sampling for multi-armed bernoulli bandit
#a function of time horizon, number of simulations,
#number of arms and the weights vector

my_thompson_bernoullibandit<-function(horizon,sims,
                                       arms,weights){

  #initializing

  counter <- matrix(rep(0,arms*sims),ncol=sims)
  reward<-matrix(rep(NA, horizon*sims), ncol=sims)
  mean_rew<-NA

  #running for each simulation
```

```

for (j in 1:sims){

  theta<-rep(NA,arms)
  alpha <- rep(1,arms)
  beta <- rep(1,arms)

  #running for each trial

  for (i in 1:horizon){

    #for each arm generate a random number from beta(0,1)
    #and choose the arm with the highest number

    for (k in 1:arms){

      theta[k]<-rbeta(1,alpha[k],beta[k])

    }

    k_max<-which.max(theta)

    #observe the true reward of this arm
    #by generating a binary number
    #from its bernoulli distribution
    #with parameter theta_k

    r <- rbinom(1,1,prob=weights[k_max])

    #update the parameters of the beta distributions
    #for each arm

    alpha[k_max] <- alpha[k_max]+r
    beta[k_max] <- beta[k_max]+ (1-r)

    #update the matrices of reward and counter
    reward[i,j]<-r
    counter[k_max,j] <- counter[k_max,j]+1
  }
}

```



```

    }

    #update the vector of the mean_rewards
    #for every simulation
    #by adding a column with the mean rewards
    #of simulation j
    mean_rew<-cbind(mean_rew,mean(reward[,j]))

}

mean_rew<-mean_rew[-1]

#return a list consisting of the mean_rewards vector
#and the arms counter for each simulation

thom_list<-list(mean_rew,counter)
names(thom_list)<-c("MeanRewards","Counter")
row.names(thom_list$Counter)<-paste0("Arm_", 1:arms)
colnames(thom_list$Counter)<-paste0("Sim_", 1:sims)

thom_list
}

```

Τρέχουμε τη συνάρτηση για να λύσει το πρόβλημά μας, με χρονικό ορίζοντα 200 επαναλήψεων, 100 προσομοιώσεις, 5 χέρια και διάνυσμα βαρών  $\mathbf{w} = (0.2, 0.4, 0.5, 0.3, 0.7)$ . Στη συνέχεια καλούμε το πρώτο στοιχείο της λίστας που επιστρέφεται:

```

bernoullibandit <- my_thompson_bernoullibandit(horizon=200,
                                              sims=100, arms=5,
                                              weights=c(0.2,0.4,0.5,0.3,0.7))

```

```
bernoullibandit$MeanRewards
```

```
## [1] 0.675 0.635 0.700 0.640 0.615 0.615 0.580 0.600 0.575 0.655 0.615 0.690
## [13] 0.585 0.655 0.590 0.685 0.615 0.645 0.610 0.630 0.580 0.595 0.705 0.630
## [25] 0.575 0.625 0.705 0.640 0.655 0.555 0.605 0.625 0.660 0.625 0.660 0.665
## [37] 0.690 0.575 0.590 0.560 0.710 0.600 0.590 0.710 0.610 0.670 0.620 0.645
## [49] 0.645 0.590 0.625 0.650 0.610 0.545 0.620 0.630 0.540 0.550 0.610 0.650
## [61] 0.685 0.485 0.605 0.655 0.570 0.670 0.580 0.590 0.685 0.645 0.585 0.695
## [73] 0.650 0.650 0.550 0.640 0.605 0.650 0.655 0.630 0.625 0.630 0.645 0.550
## [85] 0.530 0.620 0.605 0.635 0.500 0.620 0.655 0.580 0.585 0.630 0.610 0.620
## [97] 0.590 0.580 0.535 0.675
```

Σχήμα 4.5: Μέσες αμοιβές για κάθε προσομοίωση. Όλες οι αμοιβές είναι αρκετά κοντά στο 0.7 που είναι η βέλτιστη απόδοση ενώ μόνο μία είναι κάτω από το 0.5 που είναι η απόδοση της αμέσως επόμενης καλύτερης διαφήμισης.

Παρατηρούμε ότι ο αλγόριθμος σε 200 βήματα καταφέρνει και φτάνει μία μέση απόδοση κοντά στο 0.7, η οποία είναι και η βέλτιστη. Όπως θα δούμε παρακάτω, εάν του επιτρέπαμε μεγαλύτερο χρονικό ορίζοντα, αυτές οι τιμές θα ήταν ακόμα πιο κοντά στο 0.7. Λαμβάνοντας τη μέση τιμή των μέσων αμοιβών για όλες τις προσομοιώσεις, βλέπουμε ότι πράγματι ο αλγόριθμος καταλαβαίνει πως η καλύτερη διαφήμιση για το προϊόν είναι η διαφήμιση 5.

```
mean(bernoullibandit$MeanRewards)
```

```
## [1] 0.6205
```

Σχήμα 4.6: Ο μέσος των μέσων είναι 0.6205. Η τιμή είναι πολύ κοντά στο 0.7, γεγονός απολύτως θεμιτό και λογικό, δεδομένης της αποτελεσματικότητας του αλγορίθμου.

Στη συνέχεια, καλούμε το 2ο στοιχείο της λίστας για να δούμε τις επιλογές των διαφημίσεων που έκανε ο αλγόριθμος σε κάθε προσομοίωση.

```
bernoullibandit$Counter[,1:50]
```

```
##      Sim 1 Sim 2 Sim 3 Sim 4 Sim 5 Sim 6 Sim 7 Sim 8 Sim 9 Sim 10 Sim 11
## Arm 1     6     3     6     3    18    13    13     5     5     4     3
## Arm 2     3     7     3     6     4     3    11    12    14     6    10
## Arm 3    11    15    19    25    19    22    38    24    38    15    10
## Arm 4     8    10     3     6    11     4     4    14    10     4     7
## Arm 5   172   165   169   160   148   158   134   145   133   171   170
##      Sim 12 Sim 13 Sim 14 Sim 15 Sim 16 Sim 17 Sim 18 Sim 19 Sim 20 Sim 21
## Arm 1     2     7     8    12     9    11     4    16     9     4
## Arm 2     5    25     8     5     3    17     6     4    16    17
## Arm 3    21    26     8    64    12    15    24    35    20     3
## Arm 4    10     4     7     6     5    17     5     6     6    21
## Arm 5   162   138   169   113   171   140   161   139   149   155
##      Sim 22 Sim 23 Sim 24 Sim 25 Sim 26 Sim 27 Sim 28 Sim 29 Sim 30 Sim 31
## Arm 1     9     7     6     4     4     3     8     3    16     4
## Arm 2    14     5    20    16    15    11     7    12    23     7
## Arm 3    27    11    10    27    13     8     7    30    32    19
## Arm 4    13     5     8     3    16    13     8     9     4     8
## Arm 5   137   172   156   150   152   165   170   146   125   162
##      Sim 32 Sim 33 Sim 34 Sim 35 Sim 36 Sim 37 Sim 38 Sim 39 Sim 40 Sim 41
## Arm 1     6     6     9     5     3     4    10     4     7     3
## Arm 2    18     4    11    27    13     6    28    22    26    13
## Arm 3    56    10    17     3    28    14    31    20    58    14
## Arm 4     4    10    14     6     4     7     5     7     8     3
## Arm 5   116   170   149   159   152   169   126   147   101   167
##      Sim 42 Sim 43 Sim 44 Sim 45 Sim 46 Sim 47 Sim 48 Sim 49 Sim 50
## Arm 1     5     3     2     3     6     4     7     5     8
## Arm 2     3    22     3    11    16    10     8     4    16
## Arm 3    36    25     7    19     4    40    19    15    32
## Arm 4     8     4    14    11     7     7     9    10     7
## Arm 5   148   146   174   156   167   139   157   166   137
```

Σχήμα 4.7: Πίνακας με τις επιλογές του αλγορίθμου σε κάθε μία από τις 50 πρώτες προσομοιώσεις. Χρονικός ορίζοντας = 200.

```
bernoullibandit$Counter[,51:100]
```

##	Sim 51	Sim 52	Sim 53	Sim 54	Sim 55	Sim 56	Sim 57	Sim 58	Sim 59	Sim 60
## Arm 1	4	8	5	8	6	9	20	11	4	14
## Arm 2	13	10	20	15	6	6	48	25	6	14
## Arm 3	3	21	16	41	19	5	19	24	23	7
## Arm 4	9	11	6	16	13	10	9	10	7	4
## Arm 5	171	150	153	120	156	170	104	130	160	161
##	Sim 61	Sim 62	Sim 63	Sim 64	Sim 65	Sim 66	Sim 67	Sim 68	Sim 69	Sim 70
## Arm 1	3	12	4	4	7	3	9	4	3	5
## Arm 2	18	14	20	10	4	19	16	25	21	6
## Arm 3	5	19	27	14	48	7	30	16	3	7
## Arm 4	5	58	25	3	17	5	7	7	2	3
## Arm 5	169	97	124	169	124	166	138	148	171	179
##	Sim 71	Sim 72	Sim 73	Sim 74	Sim 75	Sim 76	Sim 77	Sim 78	Sim 79	Sim 80
## Arm 1	7	3	7	5	6	4	4	7	5	3
## Arm 2	12	3	25	7	26	18	8	3	4	9
## Arm 3	30	3	31	7	18	4	23	17	6	10
## Arm 4	6	3	3	4	5	8	13	6	5	8
## Arm 5	145	188	134	177	145	166	152	167	180	170
##	Sim 81	Sim 82	Sim 83	Sim 84	Sim 85	Sim 86	Sim 87	Sim 88	Sim 89	Sim 90
## Arm 1	4	7	5	9	6	4	3	4	5	5
## Arm 2	6	22	5	4	47	6	9	13	22	18
## Arm 3	8	26	13	52	51	20	3	15	60	14
## Arm 4	17	7	17	11	6	15	12	6	17	6
## Arm 5	165	138	160	124	90	155	173	162	96	157
##	Sim 91	Sim 92	Sim 93	Sim 94	Sim 95	Sim 96	Sim 97	Sim 98	Sim 99	Sim 100
## Arm 1	4	10	3	3	4	4	13	13	14	3
## Arm 2	7	28	30	3	7	6	9	14	30	20
## Arm 3	15	35	38	16	5	9	10	37	62	24
## Arm 4	9	5	14	10	4	10	9	11	19	7
## Arm 5	165	122	115	168	180	171	159	125	75	146

Σχήμα 4.8: Πίνακας με τις επιλογές του αλγορίθμου σε κάθε μία από τις 50 τελευταίες προσομοιώσεις. Χρονικός ορίζοντας = 200.

Παρατηρούμε ότι στις περισσότερες προσομοιώσεις ο αλγόριθμος αναγνωρίζει την ανωτερότητα της διαφήμισης 5 με μεγάλη ευκολία και τη διαλέγει σχεδόν αποκλειστικά. Αξίζει να σημειώσουμε ότι στις προσομοιώσεις 62, 89 και 99 έχει αργήσει λίγο περισσότερο να καταλήξει στο βέλτιστο χέρι και αυτό φαίνεται και στο Σχήμα 4.4, όπου η μέση αμοιβή είναι 0.485, 0.500 και 0.535, αντίστοιχα. Αυτό θα μπορούσε να εξαλειφθεί αυξάνοντας τον χρονικό ορίζοντα και επιτρέποντας στον αλγόριθμο να εξερευνήσει λίγο ακόμα. Ωστόσο, δεδομένου ότι τρεις από τις πέντε διαφημίσεις έχουν πιθανότητα επιτυχίας μικρότερη ή ίση του 0.4, καταλαβαίνει κανείς ότι το να αποκαλείς «αποτυχία» του αλγορίθμου μια απόδοση κοντά στο 0.5 είναι πολυτέλεια.

Επαναλαμβάνουμε το ίδιο πείραμα, για χρονικό ορίζοντα 1000 δοκιμών και καλούμε το διάνυσμα με τις μέσες αμοιβές:

```
bernoullibandit2 <- my_thompson_bernoullibandit(horizon=1000,  
sims=100, arms=5,  
weights=c(0.2,0.4,0.5,0.3,0.7))
```

```
bernoullibandit2$MeanRewards
```

```
## [1] 0.673 0.666 0.671 0.676 0.691 0.667 0.678 0.670 0.666 0.676 0.711 0.668  
## [13] 0.648 0.690 0.691 0.680 0.680 0.698 0.666 0.705 0.685 0.667 0.668 0.706  
## [25] 0.694 0.700 0.669 0.629 0.668 0.690 0.672 0.657 0.692 0.657 0.641 0.686  
## [37] 0.661 0.690 0.648 0.678 0.694 0.690 0.694 0.720 0.699 0.679 0.671 0.697  
## [49] 0.670 0.675 0.650 0.653 0.673 0.672 0.689 0.660 0.682 0.650 0.696 0.663  
## [61] 0.657 0.680 0.655 0.682 0.677 0.683 0.659 0.676 0.686 0.677 0.668 0.670  
## [73] 0.671 0.680 0.666 0.685 0.670 0.661 0.716 0.705 0.675 0.679 0.645 0.713  
## [85] 0.695 0.662 0.682 0.677 0.698 0.655 0.652 0.661 0.664 0.697 0.655 0.702  
## [97] 0.674 0.675 0.677 0.668
```

Σχήμα 4.9: Μέσες αμοιβές για κάθε προσομοίωση. Πλέον όλες οι αμοιβές είναι πάνω από 0.6 και πολύ κοντά στο 0.7.

Πράγματι, παρατηρούμε ότι οι μέσες αμοιβές σε κάθε προσομοίωση βελτιώθηκαν αισθητά, όπως περιμέναμε, εφόσον επιτρέψαμε στον αλγόριθμό μας να πειραματιστεί λίγο περισσότερο. Αυτό επιβεβαιώνεται και με τον μέσο των μέσων:

```
mean(bernoullibandit2$MeanRewards)
```

```
## [1] 0.67636
```

Σχήμα 4.10: Ο μέσος των μέσων είναι 0.67636 και πλέον φαίνεται πως ο αλγόριθμος πιάνει τη βέλτιστη απόδοση.

Τελειώνουμε το παράδειγμά μας με τον πίνακα επιλογών του αλγορίθμου σε κάθε προσομοίωση για την περίπτωση που ο χρονικός ορίζοντας είναι 1000 δοκιμές:

```
bernoullibandit2$Counter[,1:50]

##      Sim 1 Sim 2 Sim 3 Sim 4 Sim 5 Sim 6 Sim 7 Sim 8 Sim 9 Sim 10 Sim 11
## Arm 1    5    6   11   21   10   10   10    9   16    6    5
## Arm 2   17   17   13   25   11   13   21   14   28   18   10
## Arm 3   10  138    7   73    9   61   47   38   35   28   32
## Arm 4   23   10    9   24   11   12    7    9    8    9    7
## Arm 5  945  829  960  857  959  904  915  930  913  939  946
##      Sim 12 Sim 13 Sim 14 Sim 15 Sim 16 Sim 17 Sim 18 Sim 19 Sim 20 Sim 21
## Arm 1    7    9    8    8   16   14    4    4    8    5
## Arm 2   27   38    7    4   18   24   13   18    8   15
## Arm 3   77   57   33    9   14   43   10   33   23   21
## Arm 4   23   12    9   17    9   16   11    9    5    7
## Arm 5  866  884  943  962  943  903  962  936  956  952
##      Sim 22 Sim 23 Sim 24 Sim 25 Sim 26 Sim 27 Sim 28 Sim 29 Sim 30 Sim 31
## Arm 1   12    4    4    5    4    7    7    9    9   12
## Arm 2   19   23   25   20   14    8   20   50    4   23
## Arm 3   46   68    4   17    5   25   74   49    9   18
## Arm 4   22   15    6   10    4   18   12    8   27   23
## Arm 5  901  890  961  948  973  942  887  884  951  924
##      Sim 32 Sim 33 Sim 34 Sim 35 Sim 36 Sim 37 Sim 38 Sim 39 Sim 40 Sim 41
## Arm 1    8    6   14    6    7    7    5    7   14    9
## Arm 2   34   10   32   26   12   18   16   26   16   14
## Arm 3   46   47   83   22   87   17   26   91   20   34
## Arm 4   27   13    7   14    7   15    8   15   12    5
## Arm 5  885  924  864  932  887  943  945  861  938  938
##      Sim 42 Sim 43 Sim 44 Sim 45 Sim 46 Sim 47 Sim 48 Sim 49 Sim 50
## Arm 1    6    6    6    4   13   10   11    6   14
## Arm 2   17   16   11   12   36   20   20   41    8
## Arm 3   34   17   15   53    6   18   28   28   46
## Arm 4    9   31   32    4    5   31   10   18   12
## Arm 5  934  930  936  927  940  921  931  907  920
```

Σχήμα 4.11: Πίνακας με τις επιλογές του αλγορίθμου σε κάθε μία από τις 50 πρώτες προσομοιώσεις. Χρονικός ορίζοντας = 1000.

```
bernoullibandit2$Counter[,51:100]
```

```
##      Sim 51 Sim 52 Sim 53 Sim 54 Sim 55 Sim 56 Sim 57 Sim 58 Sim 59 Sim 60
## Arm 1    10    11     5    14     5    11     4     8    11     9
## Arm 2    49    20    40    16    23    20    20    56     5    21
## Arm 3    53    34    29    35    63    61    50    78    29    87
## Arm 4    13     5    11    15    10    14     9    11    11    11
## Arm 5   875   930   915   920   899   894   917   847   944   872
##      Sim 61 Sim 62 Sim 63 Sim 64 Sim 65 Sim 66 Sim 67 Sim 68 Sim 69 Sim 70
## Arm 1     8    12     7    15    12     5    23    11     7     5
## Arm 2    42     5    36    19    20    12     8    18    44    10
## Arm 3    36    24    52    14    14    57    62     4    16    18
## Arm 4    28     9    28     8    25    10    13     4    21    11
## Arm 5   886   950   877   944   929   916   894   963   912   956
##      Sim 71 Sim 72 Sim 73 Sim 74 Sim 75 Sim 76 Sim 77 Sim 78 Sim 79 Sim 80
## Arm 1     8    10     5     6     8    12     5     7     9    21
## Arm 2    24     8    24    16    11    16    11    12    15     4
## Arm 3    68    44    39    60    55    38    11    75    29    53
## Arm 4    12     8    20    16    15     7    26     7     9    10
## Arm 5   888   930   912   902   911   927   947   899   938   912
##      Sim 81 Sim 82 Sim 83 Sim 84 Sim 85 Sim 86 Sim 87 Sim 88 Sim 89 Sim 90
## Arm 1     7     4    10     6    13    12    11     4     5    10
## Arm 2    10    37    28    15    23    18     6    30     9     9
## Arm 3    12    36    18    35    16    17    58    43    45    40
## Arm 4     8     8    21     7     5     7    30    11     7    33
## Arm 5   963   915   923   937   943   946   895   912   934   908
##      Sim 91 Sim 92 Sim 93 Sim 94 Sim 95 Sim 96 Sim 97 Sim 98 Sim 99 Sim 100
## Arm 1    13    14    21    20     8     9     7    12     8    11
## Arm 2    23    10    27    21    38     8    26    24    30    17
## Arm 3   131    12    33    19    17    44     6    71    66    19
## Arm 4    12     6     9     8    15    19     9    10     9    16
## Arm 5   821   958   910   932   922   920   952   883   887   937
```

Σχήμα 4.12: Πίνακας με τις επιλογές του αλγορίθμου σε κάθε μία από τις 50 τελευταίες προσομοιώσεις. Χρονικός ορίζοντας = 1000.

Πλέον, είναι εμφανής η προτίμηση του αλγορίθμου στη διαφήμιση 5, αφού στις περισσότερες προσομοιώσεις διαλέγει τη βέλτιστη διαφήμιση με ποσοστό μεγαλύτερο του 90%.

# Ενότητα 5

## Εφαρμογές

### 5.1 Εύρεση μιας στοχαστικά βέλτιστης διαδρομής

Η πρώτη εφαρμογή, όπως μαρτυρά και ο τίτλος της, θα είναι η εύρεση της βέλτιστης διαδρομής από το σπίτι μου μέχρι το Εθνικό Μετσόβιο Πολυτεχνείο. Ουσιαστικά, το πρόβλημα μπορεί να αναχθεί σε ένα ΜΑΒ, αν το διατυπώσουμε ως εξής:

*Κάθε μέρα πηγαίνω στο Ε.Μ.Π. και θα ήθελα να χρησιμοποιώ τη διαδρομή που απαιτεί τη λιγότερη ώρα κατά μέσο όρο, χωρίς να γνωρίζω το χρόνο που χρειάζεται το εκάστοτε μονοπάτι. Πώς μπορώ να μάθω αποδοτικά και να ελαχιστοποιήσω -έπειτα από πολλές μέρες- το συνολικό χρόνο ταξιδιού;*

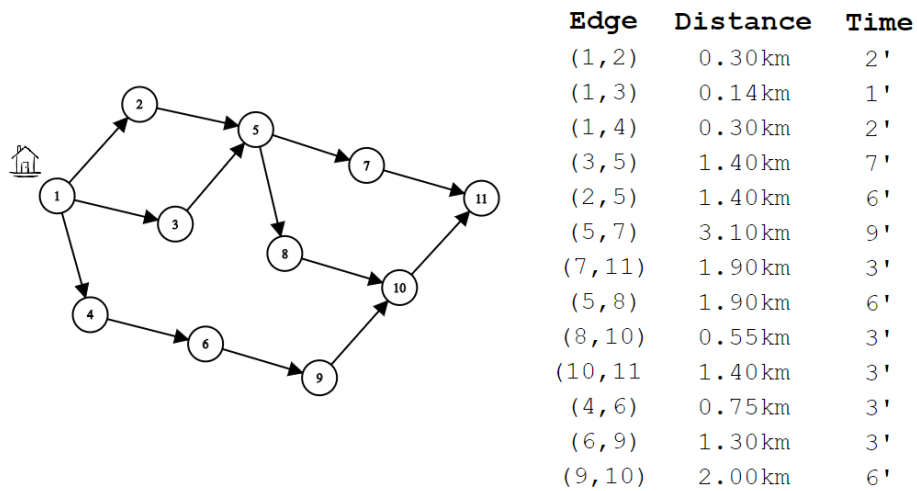
Θεωρούμε ως σημείο εκκίνησης την οδό Μασσαλίας (κόμβος 1), στην οποία σταθμεύω το αυτοκίνητο κάθε μέρα, ενώ σημείο τερματισμού (κόμβος 11) είναι η Σχολή των Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών (ΣΕΜ-ΦΕ). Οι κόμβοι αναλυτικά είναι οι εξής:

1. Μασσαλίας 12
2. Χαριλάου Τρικούπη & Σόλωνος
3. Ασκληπιού & Σόλωνος



4. Ακαδημίας & Ιπποκράτους
5. Λεωφόρος Αλεξάνδρας & Ασκληπιού
6. Βασιλίσσης Σοφίας & Ακαδημίας
7. Κατεχάκη & Κηφισίας
8. Νοσοκομείο Παίδων
9. Μιχαλακοπούλου (Hilton)
10. Κοκκινοπούλου & Αγήνορος
11. ΣΕΜΦΕ

Ακολουθεί ο κατευθυνόμενος γράφος του προβλήματος μαζί με τις αποστάσεις και τους μέσους χρόνους (από Google Maps) για την κάθε ακμή.



Σχήμα 5.1: Ο γράφος του προβλήματος με τα κόστη των επιμέρους ακμών.

Για να μοντελοποιήσουμε το πρόβλημα ([27], [28], [29], [30], [31], [32]), θεωρούμε γενικώς έναν γράφο  $G = (V, E)$ , όπου  $V = \{1, 2, \dots, N\}$  οι κορυφές (κόμβοι),  $E$  οι ακμές και  $\theta_e$  ο πραγματικός (και άγνωστος) μέσος χρόνος που αντιστοιχεί στην ακμή  $e \in E$ . Μία δράση ισοδυναμεί με μία αλληλουχία ακμών που οδηγεί από τον κόμβο 1 στον κόμβο  $N$ . Αφού εφαρμόσουμε τη δράση  $x_t$ , παρατηρούμε έναν πραγματικό χρόνο  $y_{t,e}$  για κάθε ακμή  $e \in x_t$ , το οποίο το παράγουμε τυχαία και ανεξάρτητα από μία κατανομή με μέση τιμή  $\theta_e$ . Στη συνέχεια, αθροίζουμε τους χρόνους κάθε ακμής,  $\sum_{e \in x_t} y_{t,e}$ , και το άθροισμα αυτό θα αποτελεί το συνολικό μας κόστος για τη διαδρομή που επιλέξαμε. Το κόστος μπορούμε να το φανταστούμε και ως αμοιβή  $r_t$ , για την οποία προφανώς θα ισχύει

$$r_t = - \sum_{e \in x_t} y_{t,e}. \quad (5.1)$$

Το πρόβλημα αυτό δεν ανήκει στην κατηγορία Bernoulli Bandit, καθώς η επιτυχία δεν προσμετράται με 0 και 1, αλλά με ένα κόστος, το οποίο θα ακολουθεί log-Gaussian κατανομή. Για αυτόν τον λόγο, υπάγεται στην κατηγορία Gaussian Bandit και συνεπώς, θεωρούμε μία πρότερη κατανομή για κάθε  $\theta_e$ , log-Gaussian, με παραμέτρους  $\mu_e$  και  $\sigma_e^2$ . Με άλλα λόγια,  $\ln(\theta_e) \sim N(\mu_e, \sigma_e^2)$ . Συνεπώς, έπεται ότι  $E[\theta_e] = e^{\mu_e + \sigma_e^2/2}$ . Επίσης, θεωρούμε για το  $y_{t,e}$  ότι είναι ανεξάρτητο για κάθε ακμή και ακολουθεί log-Gaussian κατανομή με παραμέτρους  $\ln(\theta_e) - \tilde{\sigma}^2/2$  και  $\tilde{\sigma}^2$ , ούτως ώστε  $E[y_{t,e} | \theta_e] = \theta_e$ . Οι ιδιότητες των συζυγών κατανομών μάς προσφέρουν έναν απλό κανόνα ανανέωσης των κατανομών των  $\theta_e$  αφού παρατηρήσουμε τα πραγματικά κόστη  $y_{t,e}$ :

$$(\mu_e, \sigma_e^2) \leftarrow \left( \frac{\frac{1}{\sigma_e^2} + \frac{1}{\tilde{\sigma}^2} \left( \ln(y_{t,e}) + \frac{\tilde{\sigma}^2}{2} \right)}{\frac{1}{\sigma_e^2} + \frac{1}{\tilde{\sigma}^2}}, \frac{1}{\frac{1}{\sigma_e^2} + \frac{1}{\tilde{\sigma}^2}} \right) \quad (5.2)$$

Στο παράδειγμα του Bernoulli Bandit, στην προηγούμενη υποενότητα, θεωρήσαμε ότι δεν είχαμε καμμία αρχική πληροφορία σχετικά με τις κατανομές των διαφημίσεων και ποιά από αυτές ήταν η βέλτιστη και για αυτόν τον λόγο ορίσαμε ως πρότερη κατανομή για κάθε διαφήμιση την  $\text{Beta}(1, 1)$ . Ωστόσο, στην εφαρμογή που αναλύουμε στην παρούσα υποενότητα, είναι λογικό να δείξουμε μια αρχική προτίμηση στις ακμές με τη μικρότερη απόσταση χιλιομέτρων. Κάτι τέτοιο μπορεί να επιτευχθεί, θέτοντας την αρχική παράμετρο για κάθε ακμή

$e \in E$  ως  $\mu_e = \ln(d_e) - \sigma_e^2/2$ . Τα αντίστοιχα αρχικά  $\sigma_e^2$  τα προσομοιώνουμε από Inverse-Gamma κατανομές, με παραμέτρους που επιτρέπουν θεμιτή και λογική αβεβαιότητα, ούτως ώστε να συμβάλουμε στην επίτευξη μιας ομαλούς εξερεύνησης. Στην παρούσα εφαρμογή, οι παράμετροι της Inverse-Gamma διαλέχθηκαν εμπειρικά, με δοκιμή και απόρριψη.

Όπως και στο Bernoulli Bandit, τα βήματα είναι απλά και ξεκάθαρα. Αρχικά, προσομοιώνουμε τα  $\hat{\theta}_e$  για κάθε ακμή  $e \in E$  από τις αντίστοιχες log-Gaussian κατανομές, με  $\mu_e = \ln(d_e) - \sigma_e^2/2$ .

Στη συνέχεια, το πρόβλημά μας θα έχει αναχθεί σε ένα ντετερμινιστικό πρόβλημα εύρεσης βέλτιστου μονοπατιού. Συνεπώς, εφαρμόζουμε τον αλγόριθμο του Dijkstra για να βρούμε το βέλτιστο μονοπάτι και κατ' επέκταση, ποιές ακμές έχουν συμμετάσχει σε αυτό.

Εφόσον δεν είναι ρεαλιστικά εφικτό να πάρουμε το αυτοκίνητο και να χρονομετρήσουμε τη διαδρομή, αντ' αυτού προσομοιώνουμε χρόνους  $y_{t,e}$  για κάθε ακμή από log-Gaussian κατανομές με μέση τιμή το χρόνο που μας δίνει το Google Maps. Το  $\tilde{\sigma}^2$  το προσομοιώνουμε από μία Inverse-Gamma με παραμέτρους τέτοιες ώστε να επιτρέπουμε απρόσμενη, αλλά ρεαλιστική κίνηση στο δρόμο. Οι τιμές των παραμέτρων της Inverse-Gamma διαλέχθηκαν εμπειρικά, με δοκιμή και απόρριψη.

Τελικά, ανανεώνουμε τις παραμέτρους  $\mu_e$  και  $\sigma_e^2$  για κάθε ακμή  $e \in E$ , σύμφωνα με τον κανόνα (5.2) και επαναλαμβάνουμε τη διαδικασία για κάθε μέρα.

---

**Algorithm 4** Thompson( $\mathcal{X}, p, q, r$ )

---

```
1: for  $t = 1, 2, \dots$  do
2:   #sample model:
3:   Sample  $\hat{\theta} \sim p$ 
4:
5:   #select and apply action:
6:    $x_t \leftarrow \operatorname{argmax}_{x \in \mathcal{X}} \mathbb{E}_{q_{\hat{\theta}}} [r(y_t) | x_t = x]$ 
7:   Apply  $x_t$  and observe  $y_t$ 
8:
9:   #update distribution:
10:   $p \leftarrow \mathbb{P}_{p,q}(\theta \in \cdot | x_t, y_t)$ 
11: end for
```

---

Σχήμα 5.2: Ο ψευδοκώδικας για το Thompson Sampling σε προβλήματα Gaussian Bandit.

Έχοντας περιγράψει το πρόβλημα και τον αλγόριθμο που το επιλύει, κατασκευάζουμε τη συνάρτηση *my\_thompson\_gaussian* που δέχεται ως ορίσματα:

- τον χρονικό ορίζοντα του πειράματος (horizon)
- το πλαίσιο δεδομένων (df) τεσσάρων στηλών, του οποίου η πρώτη στήλη περιέχει τους αρχικούς κόμβους των ακμών, η δεύτερη στήλη περιέχει τους τελικούς κόμβους των ακμών, η τρίτη στήλη περιέχει τις αποστάσεις (σε km) των ακμών (από το Google Maps) και η τέταρτη στήλη περιέχει τους μέσους χρόνους (σε λεπτά) των ακμών (επίσης από το Google Maps). Για παράδειγμα αν μια γραμμή του πλαισίου δεδομένων είναι η (1, 3, 0.14, 1), αυτό συνεπάγεται ότι η ακμή (1, 3) έχει απόσταση 0.14km και μέσο χρόνο να τη διασχίσεις, 1 λεπτό.

Η συνάρτηση επιστρέφει μία λίστα με δύο στοιχεία:

- τη λίστα ChosenPath, με στοιχεία όσα και οι μέρες του χρονικού ορίζοντα. Κάθε στοιχείο της λίστας είναι ένας πίνακας ο οποίος περιέχει

το μονοπάτι που επιλέχθηκε για εκείνη τη μέρα του χρονικού ορίζοντα (στήλη 1), μαζί με το χρονικό κόστος κάθε ακμής που συμμετείχε στο επιλεγμένο μονοπάτι (στήλη 2) αλλά και του αθροιστικού κόστους (στήλη 3).

- τον πίνακα Counter, που έχει οχτώ γραμμές και στήλες ίσες με το πλήθος των διαφορετικών μονοπατιών που επιλέχθηκαν κατά το πείραμα. Κάθε γραμμή αποτελεί το 1/8 των συνολικών ημερών και το στοιχείο (i, j) του πίνακα μας δείχνει πόσες φορές επιλέχθηκε το μονοπάτι j στο i όγδοο του πειράματος.

Για τη λειτουργία της συνάρτησής μας, χρησιμοποιούμε τα εξής πακέτα:

- *invgamma*, για να προσομοιάσουμε τιμές από την Inverse-Gamma για τα  $\sigma_e^2$  και  $\tilde{\sigma}^2$ , με χρήση της συνάρτησης `rinvgamma`.
- *cppRouting*. Το πακέτο αυτό είναι πολύ βασικό, καθώς περιέχει τη συνάρτηση `get_path_pair` που λύνει το ντετερμινιστικό πρόβλημα εύρεσης βέλτιστου μονοπατιού, με χρήση του αλγορίθμου του Dijkstra. Ωστόσο, για να γίνει αυτό, χρησιμοποιούμε και τη συνάρτηση `makegraph` του ίδιου πακέτου, για να ορίσουμε τον γράφο στη μορφή που απαιτείται ώστε να μπορέσει να εισέλθει ως όρισμα στην `get_path_pair`. Σημειώνουμε ότι η συνάρτηση `makegraph` δέχεται ως όρισμα ένα πλαίσιο δεδομένων, ακριβώς όπως το ορίζουμε ως όρισμα στη συνάρτησή μας (`my_thompson_gaussian`), μόνο που αντί για τέσσερις στήλες χρησιμοποιεί μόνο τις τρεις. Πλέον, η τρίτη στήλη είναι τα κόστη της κάθε ακμής (είτε σε απόσταση χιλιομέτρων, είτε σε χρόνο διάσχισης). Τέλος, από τα στοιχεία που επιστρέφει η συνάρτηση `makegraph`, χρησιμοποιούμε και το `nbnode` που είναι το πλήθος των κορυφών του γράφου.

Πλέον, είμαστε σε θέση να παρουσιάσουμε τη συνάρτησή μας, με επαρκή σχόλια σε κάθε σημείο του αλγορίθμου, ώστε να γίνουν όσο το δυνατόν πιο κατανοητά τα βήματα και οι χειρισμοί που απαιτήθηκαν για την ομαλή διεξαγωγή του.

```

require(invgamma)
require(cppRouting)
my_thompson_gaussian<-function(horizon,df){

  #PARAMETER INITIALIZATION AND ADJUSTMENTS

  #CREATING A GRAPH USING THE DATAFRAME OF VERTICES,
  #EDGES AND COSTS (DISTANCES - IN KM)
  df_temp<-df[,-4]
  graph<-makegraph(df_temp)

  #GETTING THE NUMBER OF VERTICES IN THE GRAPH
  total_vertices<-graph$nbnode

  #SAMPLING VARIANCES FOR EACH EDGE
  #FROM INVERSE GAMMA DISTRIBUTIONS
  S<-matrix(rinvgamma(total_vertices*total_vertices,4,3),
            ncol=total_vertices)

  #CREATING A SQUARE MATRIX
  #OF THE DISTANCES BETWEEN VERTICES
  D<-matrix(rep(0,total_vertices*total_vertices),
            ncol=total_vertices)

  for(i in 1:nrow(df)){

    D[df[i,1],df[i,2]]<-df[i,3]
  }

  #CREATING A SQUARE MATRIX
  #OF THE TRAVELING TIMES BETWEEN VERTICES
  P<-matrix(rep(0,total_vertices*total_vertices),
            ncol=total_vertices)

```

```

for(i in 1:nrow(df)){

  P[df[i,1],df[i,2]]<-df[i,4]
}

#SETTING EDGES' PRIOR DISTRIBUTION'S MEAN
mu <- log(D) - S/2

#INITIALIZING THETA
theta <- matrix(rep(0,ncol(D)*nrow(D)), ncol=ncol(D))

#INITIALIZING THE LIST CONSISTING OF THE DIFFERENT PATHS
#FOR EACH DAY
chosen_paths<-vector(mode="list",length=horizon)

#INITIALIZING A COMBINED LIST OF CHOSEN_PATHS,
#EDGE COSTS AND CUMULATIVE COSTS
combined<-vector(mode="list",length = horizon)

#STARTING THE FOR LOOP FOR EACH DAY

for( l in (1:horizon)){

  #SAMPLING S_TILDE FROM AN INVERSE GAMMA
  S_tilde<-rinvgamma(1,17,8)

  #SAMPLING THETA FOR EACH EDGE
  #FROM LOGNORMAL DISTRIBUTIONS
  #WITH MEAN = MU AND VARIANCE = S

  for (i in 1:nrow(D)){

```

```

for (j in 2:ncol(D)){
  if (D[i,j] != 0){
    theta[i,j]<- rlnorm(1,mu[i,j],S[i,j])
  }
}

}

#DEFINING A NEW DATA FRAME OF VERTICES,
#EDGES AND COSTS (THETA)
df_new<-df[,c(-3,-4)]
df_new<-cbind(df_new,0)

for ( i in 1:nrow(df_new)){

  df_new[i,3]<-theta[df_new[i,1],df_new[i,2]]
}

#CREATING A NEW GRAPH USING THE NEW DATAFRAME
graph_new<-makegraph(df_new)

#GET THE MINIMUM PATH USING DIJKSTRA'S ALGORITHM
opt_path<-get_path_pair(graph_new,1,total_vertices,
                        algorithm="Dijkstra")

#GET THE CORRESPONDING COST
#(NOT USED IN THE ALGORITHM)
path_cost<-get_distance_pair(graph_new,1,total_vertices,
                              algorithm="Dijkstra")

#THE FUNCTION GET_PATH_PAIR RETURNS THE OPTIMUM PATH
#AS A ONE-ELEMENT LIST, FROM LAST TO FIRST

```



```

#ADJUSTING THE OPTIMUM PATH IN ORDER TO GET IT
#AS A VECTOR, FROM FIRST TO LAST

opt_path<-opt_path[[1]]

op_inv<-as.numeric(opt_path)

length_op<-length(op_inv)

op<-NA

for (i in 0:(length_op-1)){

  op[i+1]<-op_inv[length_op-i]

}

#INITIALIZING THE EDGE_COST VECTOR
edge_cost<-NA

#SAMPLING A COST FOR EACH EDGE OF THE OPTIMUM PATH

for ( i in 1:(length_op-1)){

  V_start<-op[i]
  V_end<-op[i+1]

  edge_cost[i]<-rlnorm(1,meanlog = (log(P[V_start,V_end])-
                                (S_tilde/2)),
                    sdlog = S_tilde)

#UPDATING THE PARAMETERS OF THE LOGNORMAL DISTRIBUTIONS
#OF THESE EDGE'S THETA
#WHICH WERE INCLUDED IN THE OPTIMUM PATH

a1<-(1/(S[V_start,V_end]))*(mu[V_start,V_end])
a2<-(1/(S_tilde))*(log(edge_cost[i])+(S_tilde/2))

```

```

    b<-(1/S[V_start,V_end])+(1/S_tilde)

    mu[V_start,V_end]<-(a1+a2)/b
    S[V_start,V_end] <- 1/b

}

#ADDING ZERO AS THE FIRST VALUE OF THE VECTOR
#(FOR N VERTICES IN THE OPTIMUM PATH, WE HAVE N-1 EDGES)
edge_cost<-c(0,edge_cost)

#INITIALIZING AND BUILDING THE CUMULATIVE_COST VECTOR
cumulative_cost<-rep(0,(length(op)))

for (i in 2:(length(op))) {

    cumulative_cost[i]<-cumulative_cost[i-1]+edge_cost[i]

}

#SETTING THE OPTIMUM PATH AS THE CHOSEN PATH OF DAY 1
chosen_paths[[1]]<-op

#SETTING THE 1 ELEMENT OF COMBINED LIST
#AS A MATRIX OF THREE COLUMNS:
#OPTIMUM PATH, EDGE COST (MINUTES)
#AND CUMULATIVE COST (MINUTES)
combined[[1]]<-cbind(op,edge_cost,(cumulative_cost))
colnames(combined[[1]])<-c("Vertices",
                           "Cost_of_Edge_(minutes)",
                           "Cumulative_Path_Cost_(minutes)")

}

#THE FOR LOOP FOR EACH DAY IS FINISHED

```

```

#DEFINING THE UNIQUE PATHS THAT WERE CHOSEN IN THESE 1 DAYS
unique_paths<-unique(chosen_paths)

#GETTING THE TOTAL LENGTH OF THESE UNIQUE PATHS
number_of_paths<-length(unique_paths)

#SETTING THE NAME OF EACH PATH TO BE THE PATH ITSELF
names_of_paths<-rep(1,length(unique_paths))

for ( i in 1:number_of_paths){

  lngth<-length(unique_paths[[i]])
  for ( j in 1:(lngth-1)){

    names_of_paths[i]<-paste(names_of_paths[i],
                           unique_paths[[i]][j+1])

  }

}

names(unique_paths)<-names_of_paths

#DEFINING EIGHT EQUAL TIME PERIODS
q1<-0.125*horizon
q2<-0.250*horizon
q3<-0.375*horizon
q4<-0.5*horizon
q5<-0.625*horizon
q6<-0.75*horizon
q7<-0.875*horizon
q8<-horizon

#INITIALIZING THE CORRESPONDING COUNTER FOR EACH PERIOD
#AS A VECTOR OF LENGTH = NUMBER_OF_PATHS
counter_q1<-rep(0,number_of_paths)

```

```

counter_q2<-rep(0,number_of_paths)
counter_q3<-rep(0,number_of_paths)
counter_q4<-rep(0,number_of_paths)
counter_q5<-rep(0,number_of_paths)
counter_q6<-rep(0,number_of_paths)
counter_q7<-rep(0,number_of_paths)
counter_q8<-rep(0,number_of_paths)

#FOR LOOP FOR EACH TIME PERIOD
#ADD 1 TO THE ith VALUE OF COUNTER_Q1
#FOR EVERY TIME THAT PATH i WAS CHOSEN
#IN THE FIRST Q1 DAYS
#ADD 1 TO THE ith VALUE OF COUNTER_Q2
#FOR EVERY TIME THAT PATH i WAS CHOSEN
#BETWEEN THE DAYS Q1+1 AND Q2
#ETC.

for ( i in 1:number_of_paths){

  for (j in 1:q1){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q1[i]<-counter_q1[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

  for (j in (q1+1):q2){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q2[i]<-counter_q2[i]+1
    }
  }
}

```

```

}

for ( i in 1:number_of_paths){

  for (j in (q2+1):q3){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q3[i]<-counter_q3[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

  for (j in (q3+1):q4){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q4[i]<-counter_q4[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

  for (j in (q4+1):q5){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q5[i]<-counter_q5[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

```

```

for (j in (q5+1):q6){

  if( isTRUE(all.equal(chosen_paths[[j]],
                      unique_paths[[i]]))==TRUE){
    counter_q6[i]<-counter_q6[i]+1
  }
}

for ( i in 1:number_of_paths){

  for (j in (q6+1):q7){

    if( isTRUE(all.equal(chosen_paths[[j]],
                      unique_paths[[i]]))==TRUE){
      counter_q7[i]<-counter_q7[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

  for (j in (q7+1):horizon){

    if( isTRUE(all.equal(chosen_paths[[j]],
                      unique_paths[[i]]))==TRUE){
      counter_q8[i]<-counter_q8[i]+1
    }
  }
}

#COMBINING ALL THE COUNTERS TO PRESENT THEM TOGETHER
counter_final<-rbind(counter_q1,counter_q2,
                    counter_q3,counter_q4,
                    counter_q5,counter_q6,

```

```

        counter_q7, counter_q8)
counter_final<-as.matrix(counter_final)

colnames(counter_final)<-names_of_paths
rownames(counter_final)<-c("1/8_of_Total_Horizon",
                           "2/8_of_Total_Horizon",
                           "3/8_of_Total_Horizon",
                           "4/8_of_Total_Horizon",
                           "5/8_of_Total_Horizon",
                           "6/8_of_Total_Horizon",
                           "7/8_of_Total_Horizon",
                           "8/8_of_Total_Horizon")

names(combined)<-paste0("Day_", 1:horizon)

#DEFINING THE LIST TO BE RETURNED
results<-list(ChosenPath=combined,Counter=counter_final)

return(results)

}

```

Αφού κατασκευάσαμε τη συνάρτηση `my_thompson_gaussian`, ορίζουμε το πλαίσιο δεδομένων `df_gaussianbandit`, σύμφωνα με τον γράφο στο Σχήμα 5.1 και τρέχουμε τη συνάρτησή μας για 1000 ημέρες και το εν λόγω πλαίσιο δεδομένων:

```

df_gaussianbandit<-cbind(c(1,1,1,3,2,5,7,5,8,10,4,6,9),
                        c(2,3,4,5,5,7,11,8,10,11,6,9,10),
                        c(0.3,0.14,0.3,1.4,1.4,3.1,1.9,
                          1.9,0.55,1.4,0.75,1.3,2),
                        c(2,1,2,7,6,9,3,6,3,3,3,3,6))

```

```
gaussthom<-my_thompson_gaussian(horizon=1000,  
                                df=df_gaussianbandit)
```

Πριν προβούμε στα αποτελέσματα της συνάρτησης, αξίζει να αναφέρουμε πως το ντετερμινιστικά βέλτιστο μονοπάτι με βάση την απόσταση είναι το "1 3 5 8 10 11". Ωστόσο, το ντετερμινιστικά βέλτιστο μονοπάτι με βάση το χρόνο διάσχισης είναι το "1 4 6 9 10 11". Αυτό επιβεβαιώνεται με τη βοήθεια του πακέτου `cppRouting`:

```
graph_distance<-makegraph(df_gaussianbandit[,-4])  
graph_time<-makegraph(df_gaussianbandit[,-3])  
best_path_distance<-get_path_pair(graph_distance,  
                                  1,11,algorithm="Dijkstra")  
best_path_time<-get_path_pair(graph_time,  
                              1,11,algorithm = "Dijkstra")
```

```
best_path_distance
```

```
## $'1_11'  
## [1] "11" "10" "8" "5" "3" "1"
```

```
best_path_time
```

```
## $'1_11'  
## [1] "11" "10" "9" "6" "4" "1"
```

Σχήμα 5.3: Ντετερμινιστικά βέλτιστο μονοπάτι ως προς την απόσταση: "1 3 5 8 10 11". Ντετερμινιστικά βέλτιστο μονοπάτι ως προς το χρόνο διάσχισης: "1 4 6 9 10 11".

Έπειτα, λαμβάνουμε το `Counter`, το στοιχείο της λίστας που επιστρέφει η συ-



νάρτησή μας με τους μετρητές για κάθε μονοπάτι και κάθε χρονική περίοδο.

```
gaussthom$Counter
```

```
##           1 3 5 8 10 11 1 2 5 7 11 1 4 6 9 10 11 1 2 5 8 10 11
## 1/8 of Total Horizon           3           7           108           5
## 2/8 of Total Horizon           0           0           125           0
## 3/8 of Total Horizon           0           0           125           0
## 4/8 of Total Horizon           0           0           125           0
## 5/8 of Total Horizon           0           0           125           0
## 6/8 of Total Horizon           0           0           125           0
## 7/8 of Total Horizon           0           0           125           0
## 8/8 of Total Horizon           0           0           125           0
##           1 3 5 7 11
## 1/8 of Total Horizon           2
## 2/8 of Total Horizon           0
## 3/8 of Total Horizon           0
## 4/8 of Total Horizon           0
## 5/8 of Total Horizon           0
## 6/8 of Total Horizon           0
## 7/8 of Total Horizon           0
## 8/8 of Total Horizon           0
```

Σχήμα 5.4: Πίνακας με το πλήθος των φορών που επιλέχθηκε κάθε μονοπάτι σε κάθε όγδοο του χρονικού ορίζοντα.

Παρατηρούμε ότι ο αλγόριθμος καταλήγει με ευκολία στην επιλογή του αγαπημένου του μονοπατιού, το οποίο, όπως φαίνεται και στο Σχήμα 5.3, είναι το "1 4 6 9 10 11". Ήδη από την πρώτη χρονική περίοδο, και αφού πειραματίστηκε ελαφρώς (μόνο 17 φορές δοκίμασε κάποιο άλλο), άρχισε να διαλέγει το βέλτιστο χρονικά μονοπάτι με συνέπεια.

Τελειώνουμε, παρουσιάζοντας τα 4 πρώτα και τα 4 τελευταία στοιχεία της λίστας ChosenPath.

```
head(gaussthom$ChosenPath,4)
```

```
## $'Day 1'  
##      Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,]         1         0.000000         0.000000  
## [2,]         3         0.8259853         0.8259853  
## [3,]         5         6.3995891         7.2255744  
## [4,]         8         2.8187237        10.0442982  
## [5,]        10         0.8947695        10.9390676  
## [6,]        11         1.2859914        12.2250591  
##  
## $'Day 2'  
##      Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,]         1         0.000000         0.000000  
## [2,]         2         0.493165         0.493165  
## [3,]         5         3.424901         3.918066  
## [4,]         7        21.826841        25.744907  
## [5,]        11         2.956178        28.701085  
##  
## $'Day 3'  
##      Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,]         1         0.000000         0.000000  
## [2,]         4         1.719395         1.719395  
## [3,]         6         3.553533         5.272928  
## [4,]         9         1.104393         6.377321  
## [5,]        10        15.535438        21.912759  
## [6,]        11         3.059773        24.972532  
##  
## $'Day 4'  
##      Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,]         1         0.000000         0.000000  
## [2,]         3         0.7577943         0.7577943  
## [3,]         5         6.6092746         7.3670689  
## [4,]         8         7.6242786        14.9913475  
## [5,]        10         4.5796707        19.5710182  
## [6,]        11         1.3381568        20.9091750
```

Σχήμα 5.5: Τέσσερις πρώτες ημέρες του πειράματος. Στην πρώτη στήλη κάθε ημέρας βρίσκονται οι κορυφές του μονοπατιού που επιλέχθηκε, στη δεύτερη στήλη βρίσκονται οι επί μέρους χρόνοι κάθε ακμής, ενώ στην τρίτη στήλη βρίσκονται οι αθροιστικοί χρόνοι.

Παρατηρούμε ότι ο αλγόριθμος ξεκινάει με την επιλογή του μονοπατιού με τη μικρότερη αθροιστική απόσταση, "1 3 5 8 10 11". Το γεγονός αυτό είναι απλώς λογικό, αν θυμηθούμε τις μέσες τιμές των πρότερων κατανομών που ορίσαμε για τις ακμές μας, οι οποίες είναι συναρτήσεις των αποστάσεών μας. Συνεπώς, κάτι τέτοιο το περιμέναμε και είναι θεμιτό. Επίσης, βλέπουμε ότι το βέλτιστο χρονικά μονοπάτι "1 4 6 9 10 11" έχει επιλεγεί την ημέρα 3, αλλά έχει πολύ χειρότερη απόδοση (24.97 λεπτά) σε σχέση με το "1 3 5 8 10 11" (12.22 λεπτά), το οποίο επιλέγεται ξανά την ημέρα 4 και πάλι με καλύτερη απόδοση (20.90 λεπτά) από αυτή του "1 4 6 9 10 11". Είναι επόμενο, λοιπόν, να γεννηθεί η εξής απορία: *Πώς εξηγείται το γεγονός ότι παρόλο που δοκιμάζει ένα μονοπάτι και αυτό έχει καλύτερη απόδοση από αυτή του "1 4 6 9 10 11", τελικά καταλήγει στο χρονικά βέλτιστο με τόση ευκολία;*

Η απάντηση χωρίζεται σε δύο μέρη. Πρώτον, χρειάζεται να κάνουμε την απλή σχέση ότι κάθε μονοπάτι αποτελείται από ακμές και ως εκ τούτου, ο χρόνος κάθε μονοπατιού είναι το άθροισμα των χρόνων των επί μέρους ακμών. Παρατηρούμε ότι στον κόμβο 11 φτάνουμε είτε από τον 10, είτε από τον 7. Επειδή, όμως, για να βρεθούμε στον 7 πρέπει να διασχίσουμε την ακμή (5,7) που είναι χρονικά κοστοβόρα, ο αλγόριθμος αποφασίζει πως θα πάει στο 11 από τον κόμβο 10. Εκεί, θα βρεθεί είτε από τον 9, είτε από τον 8. Από τις πρώτες 4 ημέρες, του είναι ξεκάθαρο πως το να βρεθεί στον 9 είναι πολύ πιο συμφέρον χρονικά από το να βρεθεί στον 8. Οπότε η αρχική ερώτηση ανάγεται στην εξής: γιατί επιλέγει το "1 4 6 9 10 11" εφόσον η ακμή (9,10) είναι τόσο πιο «ακριβή» από την (8,10); Σε αυτό το σημείο επικαλούμαστε τους μέσους χρόνους των ακμών από το Σχήμα 5.1 και παρατηρούμε ότι ο μέσος χρόνος της ακμής (8,10) είναι 3 λεπτά, που δε διαφέρει και τόσο από το χρόνο που έκανε στο πείραμά μας τις μέρες 1 και 4 (~1 λεπτό και ~4.5 λεπτά, αντίστοιχα). Συγκρίνοντας, όμως, το μέσο χρόνο της ακμής (9,10) που είναι 6 λεπτά με τον χρόνο που έκανε να τη διασχίσει ο αλγόριθμός μας τη μέρα 3, που είναι ~15 λεπτά, έρχεται και δένει το δεύτερο επιχείρημα του συλλογισμού. Το επιχείρημα, αυτό, μας λέει πως σύμφωνα με τον κανόνα ανανέωσης της σχέσης (5.2), ο αλγόριθμος δεν «τιμωρεί» κακές αποδόσεις για τις οποίες ευθύνεται κυρίως μία μεγάλη διασπορά  $\sigma^2$ . Προφανώς, λοιπόν, βρεθήκαμε σε αυτήν την περίπτωση, όπου ο αλγόριθμός μας ικανοποιήθηκε από την πορεία μέχρι και τον κόμβο 9 και, από εκεί και πέρα, δεν «τιμώρησε» την κακή απόδοση μέχρι τον κόμβο 10, λόγω μεγάλης διασποράς.

Ο παραπάνω συλλογισμός επιβεβαιώνεται και από τις τελευταίες 4 ημέρες του

πειράματος:

```
tail(gaussthom$ChosenPath,4)
```

```
## $'Day 997'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 4 2.3536720 2.353672  
## [3,] 6 2.8814777 5.235150  
## [4,] 9 5.9303055 11.165455  
## [5,] 10 8.1380340 19.303489  
## [6,] 11 0.6963264 19.999816  
##  
## $'Day 998'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 4 1.260655 1.260655  
## [3,] 6 4.588886 5.849540  
## [4,] 9 1.544358 7.393898  
## [5,] 10 3.594166 10.988064  
## [6,] 11 2.850577 13.838640  
##  
## $'Day 999'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 4 1.473960 1.473960  
## [3,] 6 3.090123 4.564084  
## [4,] 9 3.864649 8.428733  
## [5,] 10 4.852254 13.280987  
## [6,] 11 1.664985 14.945972  
##  
## $'Day 1000'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 4 1.193392 1.193392  
## [3,] 6 2.132885 3.326277  
## [4,] 9 1.501477 4.827754  
## [5,] 10 7.204992 12.032746  
## [6,] 11 2.019975 14.052721
```

Σχήμα 5.6: Τέσσερις τελευταίες ημέρες του πειράματος.

Πράγματι, παρατηρούμε ότι και τις τέσσερις τελευταίες ημέρες, που το "1 4 6 9 10 11" επιλέχθηκε ως βέλτιστο, ο χρόνος διάσχισης της ακμής (9,10) δεν ξεπέρασε τα 9 λεπτά.

Κλείνοντας την παρούσα εφαρμογή, σημειώνουμε ότι επαναλάβουμε το πείραμα αρκετές φορές και θα μπορούσαμε να έχουμε παραθέσει στην εργασία κάποιες από τις επαναλήψεις. Ωστόσο, εφόσον δεν αλλάζουν τα αποτελέσματα και η ανάλυση αυτών, κρίνουμε φρόνιμο να μη συμπεριλάβουμε καμμία άλλη προσομοίωση του πειράματος και να προχωρήσουμε άμεσα στην επόμενη εφαρμογή.

## 5.2 Εύρεση λύσης σε ένα στοχαστικό Πρόβλημα του Πλανόδιου Πωλητή (Traveling Salesman Problem)

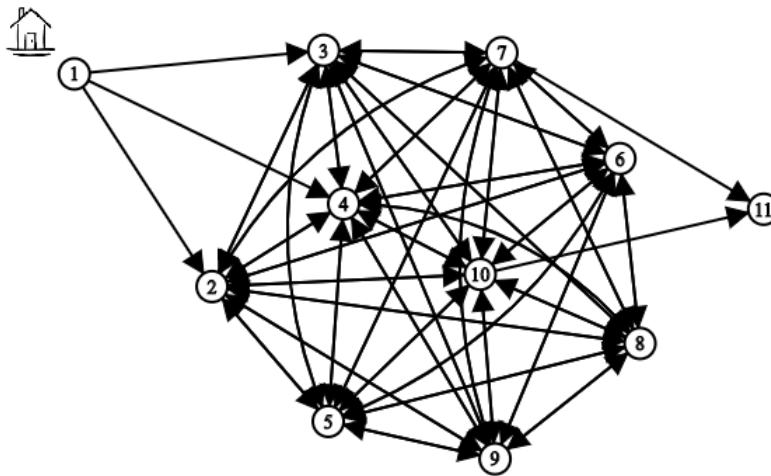
Στην τελευταία υποενότητα και εφαρμογή της παρούσας εργασίας, θα μετατρέψουμε το πρόβλημα εύρεσης βέλτιστου μονοπατιού της υποενότητας 5.1 σε πρόβλημα πλανόδιου πωλητή. Δηλαδή, σκοπός μας πλέον δε θα είναι να βρούμε το βέλτιστο χρονικά μονοπάτι από τον κόμβο 1 στον κόμβο 11, αλλά να βρούμε τη βέλτιστη χρονικά σειρά με την οποία θα επισκεφθούμε όλους τους ενδιαμέσους κόμβους, ξεκινώντας από την ίδια αφετηρία και καταλήγοντας στο Ε.Μ.Π.

Για να προσδώσουμε επιπλέον ενδιαφέρον στο πρόβλημά μας, θεωρούμε ότι ο πρωταγωνιστής της ιστορίας ξεκινάει από την αφετηρία, όπου έχει παρκαρσμένο το αυτοκίνητό του, και θέλει να καταλήξει στη ΣΕΜΦΕ αφού πρώτα κάνει κάποιες δουλειές στο κέντρο της Αθήνας. Πιο συγκεκριμένα, θέλει να διευθετήσει τις εξής υποχρεώσεις:

1. Να ξεκινήσει από τη **Μασσαλίας 12**.
2. Να αγοράσει ένα ζευγάρι ακουστικά από το Athens Pro Audio στην **Ασκληπιού 26**.
3. Να αγοράσει μία κιθάρα, από το Armaos Music Store στη **Χαριλάου Τρικούπη 24**.
4. Να αγοράσει φάρμακα από το φαρμακείο του Κοτσερώνη, στην **Ακαδημίας 39**.
5. Να κάνει εξετάσεις στη Βιοϊατρική στη **Λεωφόρο Αλεξάνδρας 142**.
6. Να διευθετήσει εκκρεμότητες με την πρεσβεία του Περού, στη **Βασιλίσσης Σοφίας 19**.
7. Να περάσει από το σταθμό του **Μετρό στην Κατεχάκη**, να παραλάβει τον φίλο του που θα έρθει μαζί του στη σχολή.
8. Να βοηθήσει την αδερφή του που εμβολιάζει τις κόρες της στο νοσοκομείο Παίδων, **Φωκίδος 63**.

9. Να πάρει κάτι να φάει από την καντίνα στη **Μιχαλακοπούλου 38**.
10. Να αγοράσει ένα μπουκάλι κρασί από το 24ωρο mini market στην **Κοκκινοπούλου 11**, με σκοπό να γιορτάσει αργότερα.
11. Να φτάσει στη ΣΕΜΦΕ για να παρουσιάσει την εργασία του.

Ο κατευθυνόμενος και όχι συμμετρικός γράφος του προβλήματος, θα έχει τη μορφή ενός πλέγματος. Ξεκινάει από τον κόμβο 1, πάει σε έναν εκ των 2, 3 ή 4 και καταλήγει στον κόμβο 11 από τον 7 ή τον 10. Οι κόμβοι 2 έως 10 αποτελούν έναν πλήρη γράφο. Το πλέγμα είναι αρκετά σύνθετο στην όψη και δεν προσδίδει κάποια εύπεπτη πληροφορία, αλλά μοιάζει κάπως έτσι:



Σχήμα 5.7: Ο κατευθυνόμενος γράφος του Προβλήματος του Πλανόδιου Πωλητή.

Στην πραγματικότητα, ο γράφος θα είναι πλήρης και δηλαδή θα μπορούμε από οποιαδήποτε κορυφή να βρεθούμε σε όποια άλλη θέλουμε. Ωστόσο, για χάρη του προβλήματος, θεωρούμε ότι πάντα ξεκινάμε από την κορυφή 1 και πάντα καταλήγουμε στην 11. Για να το πετύχουμε αυτό, δημιουργούμε έναν πλήρη γράφο, θέτοντας όμως κάποιες αποστάσεις να είναι ακραία μεγάλες, ώστε να αποθαρρύνουμε τον αλγόριθμο από το να τις χρησιμοποιήσει. Για παράδειγμα,

θα θέσουμε την ακμή (1,5) να έχει απόσταση  $d_{1,5} = e^{20} \simeq 485165195.41$  και παράλληλα εξίσου μεγάλο μέσο χρόνο διάσχισης  $\theta_{1,5} = e^{20} \simeq 485165195.41$ . Επιπλέον, στη διαγώνιο του πίνακα αποστάσεων, αλλά και του πίνακα μέσων χρόνων, θέτουμε τα στοιχεία να είναι και αυτά ίσα με  $e^{20}$ , χωρίς να επηρεάζουν κάπως τον αλγόριθμο, αφού ούτως ή άλλως στο πρόβλημά μας θα πρέπει να περάσουμε από όλους τους κόμβους ακριβώς μία φορά.

Έχοντας πει όλα αυτά, ο πίνακας μέσων χρόνων διάσχισης των ακμών σε λεπτά (από Google Maps) ορίζεται ως εξής (τον χωρίζουμε σε δύο μέρη, στήλες 1-6 και στήλες 7-11, για περισσότερη ευκολία στην ανάγνωση):

```
df_tsp_times<-rbind(c(exp(20),2,6,4,exp(20),exp(20),
                    exp(20),exp(20),exp(20),
                    exp(20),exp(20)),
                    c(exp(20),exp(20),6,6,6,10,
                    15,13,15,16,exp(20)),
                    c(exp(20),3,exp(20),4,7,7,
                    16,13,11,17,exp(20)),
                    c(exp(20),5,7,exp(20),11,2,
                    17,14,13,18,exp(20)),
                    c(exp(20),12,14,12,exp(20),8,
                    9,7,12,10,exp(20)),
                    c(exp(20),8,10,9,12,exp(20),
                    10,9,5,12,exp(20)),
                    c(exp(20),20,19,16,13,11,
                    exp(20),5,12,7,13),
                    c(exp(20),22,19,16,14,11,
                    5,exp(20),10,3,exp(20)),
                    c(exp(20),17,17,13,14,8,
                    7,5,exp(20),6,exp(20)),
                    c(exp(20),21,23,18,17,14,
                    6,3,12,exp(20),7),
                    c(exp(20),exp(20),exp(20),
                    exp(20),exp(20),exp(20),
                    exp(20),exp(20),exp(20),
                    exp(20),exp(20)))

vertices_names<-c("Massalias","Asklipiou","Trikoupi",
```



```

      "Akadimias", "Alexandras",
      "Vas. Sofias", "Katexaki", "Goudi",
      "Mixalakopoulou", "Kokkinopoulou", "SEMFE")
colnames(df_tsp_times) <- vertices_names
rownames(df_tsp_times) <- vertices_names

```

```
knitr::kable(df_tsp_times[,1:6])
```

	Massalias	Asklipiou	Trikoupi	Akadimias	Alexandras	Vas. Sofias
Massalias	485165195	2	6	4	485165195	485165195
Asklipiou	485165195	485165195	6	6	6	10
Trikoupi	485165195	3	485165195	4	7	7
Akadimias	485165195	5	7	485165195	11	2
Alexandras	485165195	12	14	12	485165195	8
Vas. Sofias	485165195	8	10	9	12	485165195
Katexaki	485165195	20	19	16	13	11
Goudi	485165195	22	19	16	14	11
Mixalakopoulou	485165195	17	17	13	14	8
Kokkinopoulou	485165195	21	23	18	17	14
SEMFE	485165195	485165195	485165195	485165195	485165195	485165195

Σχήμα 5.8: Ο τετραγωνικός πίνακας με τους μέσους χρόνους διάσχισης των ακμών (λεπτά). Στήλες 1-6.

```
knitr::kable(df_tsp_times[,7:11])
```

	Katexaki	Goudi	Mixalakovoulou	Kokkinopoulou	SEMFE
Massalias	485165195	485165195	485165195	485165195	485165195
Asklipiou	15	13	15	16	485165195
Trikoupi	16	13	11	17	485165195
Akadimias	17	14	13	18	485165195
Alexandras	9	7	12	10	485165195
Vas. Sofias	10	9	5	12	485165195
Katexaki	485165195	5	12	7	13
Goudi	5	485165195	10	3	485165195
Mixalakovoulou	7	5	485165195	6	485165195
Kokkinopoulou	6	3	12	485165195	7
SEMFE	485165195	485165195	485165195	485165195	485165195

Σχήμα 5.9: Ο τετραγωνικός πίνακας με τους μέσους χρόνους διάσχισης των ακμών (λεπτά). Στήλες 7-11.

Αντίστοιχα, ορίζουμε τον τετραγωνικό πίνακα αποστάσεων των ακμών σε χιλιόμετρα (από Google Maps) (τον χωρίζουμε σε δύο μέρη, στήλες 1-6 και στήλες 7-11, για περισσότερη ευκολία στην ανάγνωση).

```
df_tsp_distances<-rbind(c(exp(20),0.35,0.60,
                          0.60,exp(20),exp(20),
                          exp(20),exp(20),exp(20),
                          exp(20),exp(20)),
                        c(exp(20),exp(20),0.85,
                          0.85,1.40,1.60,
                          4.40,3.20,2.60,
                          4.00,exp(20)),
                        c(exp(20),0.40,exp(20),
                          0.70,1.70,1.40,
                          4.60,3.40,2.70,
                          4.10,exp(20)),
                        c(exp(20),0.75,0.90,
                          exp(20),2.30,0.70,
```

```
5.00,3.70,2.00,  
4.50,exp(20)),  
c(exp(20),2.30,2.40,  
2.50,exp(20),2.80,  
2.80,1.70,2.50,  
2.40,exp(20)),  
c(exp(20),1.10,1.40,  
1.40,2.70,exp(20)),  
4.00,3.10,1.30,  
3.20,exp(20)),  
c(exp(20),5.00,5.20,  
4.60,3.00,3.80,  
exp(20),1.60,3.40,  
2.10,4.40),  
c(exp(20),4.10,4.40,  
3.70,2.20,3.00,  
1.50,exp(20),2.50,  
0.60,exp(20)),  
c(exp(20),3.30,3.00,  
2.30,4.10,2.10,  
3.00,1.80,exp(20),  
2.30,exp(20)),  
c(exp(20),4.00,4.30,  
3.60,3.00,2.80,  
2.60,0.60,2.70,  
exp(20),2.50),  
c(exp(20),exp(20),exp(20),  
exp(20),exp(20),exp(20),  
exp(20),exp(20),exp(20),  
exp(20),exp(20))
```

```
colnames(df_tsp_distances)<-vertices_names  
rownames(df_tsp_distances)<-vertices_names
```

```
knitr::kable(df_tsp_distances[,1:6])
```

	Massalias	Asklipiou	Trikoupi	Akadimias	Alexandras	Vas. Sofias
Massalias	485165195	0.35	0.60	0.60	485165195.4	485165195.4
Asklipiou	485165195	485165195.41	0.85	0.85	1.4	1.6
Trikoupi	485165195	0.40	485165195.41	0.70	1.7	1.4
Akadimias	485165195	0.75	0.90	485165195.41	2.3	0.7
Alexandras	485165195	2.30	2.40	2.50	485165195.4	2.8
Vas. Sofias	485165195	1.10	1.40	1.40	2.7	485165195.4
Katexaki	485165195	5.00	5.20	4.60	3.0	3.8
Goudi	485165195	4.10	4.40	3.70	2.2	3.0
Mixalakopoulou	485165195	3.30	3.00	2.30	4.1	2.1
Kokkinopoulou	485165195	4.00	4.30	3.60	3.0	2.8
SEMFE	485165195	485165195.41	485165195.41	485165195.41	485165195.4	485165195.4

Σχήμα 5.10: Ο τετραγωνικός πίνακας με τις αποστάσεις των ακμών (χιλιόμετρα). Στήλες 1-6.

```
knitr::kable(df_tsp_distances[,7:11])
```

	Katexaki	Goudi	Mixalakopoulou	Kokkinopoulou	SEMFE
Massalias	485165195.4	485165195.4	485165195.4	485165195.4	485165195.4
Asklipiou	4.4	3.2	2.6	4.0	485165195.4
Trikoupi	4.6	3.4	2.7	4.1	485165195.4
Akadimias	5.0	3.7	2.0	4.5	485165195.4
Alexandras	2.8	1.7	2.5	2.4	485165195.4
Vas. Sofias	4.0	3.1	1.3	3.2	485165195.4
Katexaki	485165195.4	1.6	3.4	2.1	4.4
Goudi	1.5	485165195.4	2.5	0.6	485165195.4
Mixalakopoulou	3.0	1.8	485165195.4	2.3	485165195.4
Kokkinopoulou	2.6	0.6	2.7	485165195.4	2.5
SEMFE	485165195.4	485165195.4	485165195.4	485165195.4	485165195.4

Σχήμα 5.11: Ο τετραγωνικός πίνακας με τις αποστάσεις των ακμών (χιλιόμετρα). Στήλες 7-11.

Η λογική του αλγορίθμου που θα επιλύει αυτό το πρόβλημα είναι η ίδια με

αυτήν του αλγορίθμου της εφαρμογής 5.1, καθώς αμφότεροι χρησιμοποιούν την Thompson Sampling. Η διαφορά έγκειται στο πώς γίνεται η επιλογή του βέλτιστου, αφού παραγάγουμε τα  $\hat{\theta}_e$  για κάθε ακμή. Στο πρόβλημα της εφαρμογής 5.1, φάχναμε το βέλτιστο μονοπάτι από τον κόμβο 1 στον κόμβο 11 και συνεπώς, χρησιμοποιούσαμε τον αλγόριθμο του Dijkstra για να μας υποδείξει ποιο μονοπάτι είναι το βέλτιστο. Στην παρούσα εφαρμογή, από τη στιγμή που ολοκληρωθεί το στάδιο της παραγωγής των  $\hat{\theta}_e$ , ερχόμαστε αντιμέτωποι με ένα ντετερμινιστικό πρόβλημα πλανόδιου πωλητή. Δηλαδή, το πρόβλημα ανάγεται στην ερώτηση «πώς θα επισκεφθούμε όλους τους κόμβους από μία φορά με χρονικά βέλτιστο τρόπο;».

Το Πρόβλημα του Πλανόδιου Πωλητή υπάγεται στην κατηγορία των NP-hard προβλημάτων και ως εκ τούτου, η επίλυση του απαιτεί κάποιον ευριστικό αλγόριθμο (heuristics). Ένας εύκολος, εύχρηστος, διαδεδομένος και πολύ αποτελεσματικός αλγόριθμος είναι αυτός του Simulated Annealing (SA, από εδώ και στο εξής), ο οποίος λύνει το Traveling Salesman Problem βρίσκοντας μία λύση «πολύ κοντά» στη βέλτιστη.

Το όνομα και η έμπνευση του SA προέρχεται από την ανόπτηση στη μεταλλουργία, μία τεχνική που περιλαμβάνει θέρμανση και ελεγχόμενη ψύξη ενός υλικού, προκειμένου να αυξηθεί το μέγεθος των κρυστάλλων του και να μειωθούν τα ελαττώματά τους. Αμφότερα τα μεγέθη αυτά είναι χαρακτηριστικά του υλικού που εξαρτώνται από τη θερμοδυναμική ελεύθερη ενέργεια. Με τη θέρμανση και την ψύξη του υλικού, επηρεάζεται και η θερμοκρασία, αλλά και η ελεύθερη ενέργεια. Ο Simulated Annealing μπορεί να χρησιμοποιηθεί για την προσέγγιση του ολικού ελαχίστου σε συναρτήσεις με πολλές μεταβλητές. Για περισσότερες πληροφορίες σχετικά με τον SA και τη στοχαστική προσομοίωση, παραπέμπουμε τον αναγνώστη στα [33], [34], [35].

Η ιδέα του αλγορίθμου είναι απλή:

1. Διάλεξε μία αρχική τυχαία κατάσταση  $x_{\text{start}}$  και θέσε  $x = x_{\text{start}}$ .
2. Θέσε μία αρχική θερμοκρασία  $T_1$ .
3. Διάλεξε μία τυχαία γειτονική κατάσταση  $x'$ .
4. Αν  $f$  η συνάρτηση κόστους και  $f(x') < f(x)$ , τότε θέσε  $x = x'$ .

5. Διαφορετικά, θέσε  $x = x'$  με πιθανότητα

$$\exp\left\{\frac{f(x) - f(x')}{T}\right\}. \quad (5.3)$$

6. Μείωσε το  $T$  βάσει κάποιου προγράμματος ψύξης.

7. Επανάλαβε τα βήματα 3-6 μέχρι το  $T$  να γίνει ίσο με  $T_f$ .

Το πρόγραμμα ψύξης ουσιαστικά είναι η συνάρτηση με την οποία φθίνει η θερμοκρασία με κάθε βήμα του αλγορίθμου. Στη συνάρτηση SA που θα παρουσιάσουμε παρακάτω, έχουμε επιλέξει το γραμμικό πρόγραμμα ψύξης, για το οποίο ισχύει:

$$T_i = \frac{T_1 - T_f}{M - 1}(M - i) + T_f, \quad (5.4)$$

όπου  $M$  είναι ο συνολικός αριθμός των επαναλήψεων του αλγορίθμου.

Για αρχή, λοιπόν, κατασκευάζουμε τη συνάρτηση του Simulated Annealing, η οποία λύνει το Traveling Salesman Problem και δέχεται ως ορίσματα:

- έναν τετραγωνικό πίνακα (df) διάστασης  $n = \text{πλήθος των κόμβων}$ , όπου το στοιχείο  $(i, j)$  του πίνακα αντιστοιχεί στο κόστος της ακμής  $(i, j)$  (σε χρόνο ή απόσταση).
- τον αριθμό των επαναλήψεων ( $M$ ).

Η συνάρτηση θα επιστρέφει μία λίστα δύο στοιχείων:

- το Table1, το οποίο είναι το διάνυσμα με τις κορυφές του γράφου, στη σειρά με την οποία επιλέχθηκαν ώστε να είναι βέλτιστο το μονοπάτι.
- το Table2, το οποίο είναι το συνολικό κόστος του μονοπατιού που επιλέχθηκε.

Παραθέτουμε τον κώδικα της συνάρτησης *my\_sa* στην R, που επιλύει το πρόβλημα του πλανόδιου πωλητή με χρήση SA, ξεκινώντας από τον κόμβο 1 και καταλήγοντας στον 11.

```
my_sa<-function(df, M){

#GETTING THE TOTAL NUMBER OF VERTICES
  cities<-ncol(df)

#DEFINING THE COST FUNCTION FOR A PATH
#IN THE TRAVELING SALESMAN PROBLEM
  cost<-function(x){
    edge<-NA
    edge[1]<-0
    for(i in 1:(length(x)-1)){

      edge[i+1]<-df[x[i],x[i+1]]
    }
    sum(edge)
  }

#INITIALIZING AND RANDOMIZING THE STARTING PATH
  x_start<-NA

#ALWAYS STARTING AT VERTEX 1
  x_start[1]<-1

#ALWAYS ENDING AT THE LAST VERTEX
  x_start[cities]<-cities

#RANDOMIZING THE REST OF THE VERTICES
  sampler<-sample(2:(cities-1))

#COMBINING VERTEX 1
#WITH THE RANDOM VERTICES
#AND THE LAST VERTEX
```

```

x_start<-c(x_start[1],sampler,x_start[cities])

#INITIALIZING THE TEMPERATURE
T_start<-1
T_final<-0.1
T_mid <- T_start

#STARTING THE LOOP
#REPEAT WHILE T_mid > T_final
j=2
while (T_mid > T_final){

  x_new <- x_start

  #CHOOSING RANDOMLY N/2 INTERMEDIATE VERTICES
  #AND RANDOMLY CHANGING THEIR ORDER
  #TO TRY IT AS A NEW CONFIGURATION
  #N = TOTAL_NUMBER_OF_VERTICES
  changes<-sample(x_new[-c(1,cities)],floor(cities/2))
  changes_new <- sample(changes)

  #GENERATE A RANDOM NUMBER IN (0,1)
  rand_prob<-runif(1)

  #DEFINING X_ASSIST TO ASSIST
  #IN CHANGING THE ORDER OF X_NEW
  x_assist<-x_new
  for (i in 1:length(changes)){

    x_new[x_assist==changes[i]]<-changes_new[i]

  }

  #CHECKING THE VALIDITY
  #OF CONDITION NUMBER 1
  #COST(X_NEW) < COST(X_OLD)
  if (cost(x_new)<= cost(x_start)){

    x_start<-x_new
  }
}

```



```

}

#ELSE, CHECKING THE VALIDITY
#OF CONDITION NUMBER 2
#PROBABILITY > RUNIF(1)
else if (exp((cost(x_start)-cost(x_new))/T_mid)
>= rand_prob){

  x_start <- x_new
}

#COUNTER ++
j<-j+1

#DECREASING TEMPERATURE
T_mid <- ((T_start-T_final)/(M-1))*(M-j))+T_final

}

#THE WHILE LOOP IS OVER

#GETTING THE COST OF THE OPTIMUM PATH
path_cost<-cost(x_start)

#DEFINING THE LIST
#WITH THE OPTIMUM PATH VECTOR
#AND THE COST OF THE OPTIMUM PATH
results<-list(Table1=x_start,Table2=path_cost)

#RETURNING THE RESULTS
return(results)
}

```

Δοκιμάζουμε την αποτελεσματικότητα της συνάρτησής μας, για τον πίνακα  $df\_tsp\_times$  και  $M = 1000$ . Επαναλαμβάνουμε τη διαδικασία 50 φορές για να έχουμε μια εικόνα για το πού περίπου κυμαίνεται ο βέλτιστος χρόνος.

```

sa_test_paths<-rep(0,11)
sa_test_times<-NA

for(i in 1:50){

  sa_trial<-my_sa(df_tsp_times,1000)
  path<-sa_trial$Table1
  time<-sa_trial$Table2
  sa_test_times[i]<-time
  sa_test_paths<-rbind(sa_test_paths,path)
}

sa_test_paths<-sa_test_paths[-1,]

final_sa_test_df<-cbind(sa_test_paths,sa_test_times)
colnames(final_sa_test_df)<-c(paste0("Vertex_",1:11),
                             "Time")
rownames(final_sa_test_df)<-paste0("Path_",1:50)

View(final_sa_test_df)

```

	Vertex 1	Vertex 2	Vertex 3	Vertex 4	Vertex 5	Vertex 6	Vertex 7	Vertex 8	Vertex 9	Vertex 10	Vertex 11	Time
Path 1	1	4	6	3	2	5	9	7	8	10	11	59
Path 2	1	3	2	5	7	4	6	9	8	10	11	62
Path 3	1	2	3	4	6	9	7	5	8	10	11	56
Path 4	1	4	6	9	7	3	2	5	8	10	11	63
Path 5	1	4	6	9	8	7	3	2	5	10	11	66
Path 6	1	4	3	2	5	7	6	9	8	10	11	60
Path 7	1	2	5	3	4	6	9	7	8	10	11	55
Path 8	1	4	6	9	8	7	3	2	5	10	11	66
Path 9	1	4	6	3	2	5	9	7	8	10	11	59
Path 10	1	2	5	3	4	6	9	8	7	10	11	57
Path 11	1	4	6	3	2	5	9	7	8	10	11	59
Path 12	1	2	3	4	6	9	7	5	8	10	11	56
Path 13	1	2	5	3	4	6	9	7	8	10	11	55
Path 14	1	4	3	2	5	7	6	9	8	10	11	60
Path 15	1	4	3	2	5	6	9	7	8	10	11	55
Path 16	1	2	4	6	9	3	5	7	8	10	11	63
Path 17	1	2	3	4	6	9	7	5	8	10	11	56
Path 18	1	4	3	2	5	6	9	7	8	10	11	55
Path 19	1	4	3	2	5	6	9	7	8	10	11	55
Path 20	1	4	3	2	5	6	9	7	8	10	11	55
Path 21	1	3	4	2	5	6	9	7	8	10	11	56
Path 22	1	3	4	2	5	6	9	7	8	10	11	56
Path 23	1	2	5	3	4	6	9	7	8	10	11	55
Path 24	1	3	2	5	4	6	9	7	8	10	11	56
Path 25	1	4	3	2	5	6	9	7	8	10	11	55

Σχήμα 5.12: Οι 25 πρώτες δοκιμές της συνάρτησης *my\_sa* για τον πίνακα *df\_tsp\_times*.

	Vertex 1	Vertex 2	Vertex 3	Vertex 4	Vertex 5	Vertex 6	Vertex 7	Vertex 8	Vertex 9	Vertex 10	Vertex 11	Time
Path 26	1	2	5	3	4	6	9	7	8	10	11	55
Path 27	1	4	6	3	2	5	7	9	8	10	11	61
Path 28	1	4	3	2	5	6	9	7	8	10	11	55
Path 29	1	3	2	5	8	7	4	6	9	10	11	63
Path 30	1	4	6	9	3	2	5	8	7	10	11	63
Path 31	1	2	3	5	4	6	9	7	8	10	11	56
Path 32	1	2	5	8	3	4	6	9	7	10	11	66
Path 33	1	4	3	2	5	8	7	6	9	10	11	61
Path 34	1	2	3	5	8	7	4	6	9	10	11	63
Path 35	1	2	3	4	6	9	7	5	8	10	11	56
Path 36	1	3	4	6	2	5	9	7	8	10	11	60
Path 37	1	4	3	2	5	6	9	7	8	10	11	55
Path 38	1	2	3	4	6	5	9	7	8	10	11	60
Path 39	1	2	5	3	4	6	9	7	8	10	11	55
Path 40	1	2	3	4	6	9	7	5	8	10	11	56
Path 41	1	4	3	2	5	6	9	7	8	10	11	55
Path 42	1	4	3	2	5	6	9	7	8	10	11	55
Path 43	1	4	6	3	2	5	9	7	8	10	11	59
Path 44	1	3	2	5	4	6	9	7	8	10	11	56
Path 45	1	2	3	4	6	9	5	7	8	10	11	57
Path 46	1	2	5	3	4	6	9	7	8	10	11	55
Path 47	1	2	5	3	4	6	9	7	8	10	11	55
Path 48	1	4	3	2	5	6	9	7	8	10	11	55
Path 49	1	3	2	4	6	9	7	5	8	10	11	59
Path 50	1	4	3	2	5	6	9	7	8	10	11	55

Σχήμα 5.13: Οι 25 τελευταίες δοκιμές της συνάρτησης *my\_sa* για τον πίνακα *df\_tsp\_times*.

```
mean(sa_test_times)
```

```
## [1] 57.92
```

Σχήμα 5.14: Ο «μέσος» βέλτιστος χρόνος είναι 57.92.

Παρατηρούμε πως ο «καλύτερος» βέλτιστος χρόνος που εμφανίζεται στο πε-  
 ιράμά μας είναι τα 55 λεπτά και επιτυγχάνεται από τις διαδρομές "1 2 5 3 4 6

9 7 8 10 11" και "1 4 3 2 5 6 9 7 8 10 11". Ο αμέσος επόμενος καλύτερος χρόνος είναι τα 56 λεπτά, από τα μονοπάτια "1 2 3 5 4 6 9 7 8 10 11", "1 2 3 4 6 9 7 5 8 10 11" και "1 3 2 5 4 6 9 7 8 10 11". Από την άλλη, ο «χειρότερος» βέλτιστος χρόνος είναι τα 66 λεπτά, από τα μονοπάτια "1 4 6 9 8 7 3 2 5 10 11" και "1 2 5 8 3 4 6 9 7 10 11". Ο «μέσος» καλύτερος χρόνος είναι 57.92. Εννοείται ότι ενδέχεται να υπάρχουν και άλλα μονοπάτια στα οποία να επιτυγχάνονται ίδιοι χρόνοι και απλά να μην έτυχε να εμφανιστούν στο πείραμα των 50 επαναλήψεων, καθώς το συνολικό πλήθος των μονοπατιών είναι της τάξης του  $n!$ . Τελικά, συμπεραίνουμε ότι οποιοσδήποτε χρόνος κοντά στη 1 ώρα είναι ένας καλός χρόνος για το πρόβλημά μας.

Είμαστε πλέον σε θέση να κατασκευάσουμε την τελική συνάρτηση που θα λέγεται *my\_thompson\_tsp* και θα εφαρμόζει τον αλγόριθμο του TS, ώστε με τη χρήση της συνάρτησης *my\_sa* να λύνει το πρόβλημα του πλανόδιου πωλητή όπως το έχουμε ορίσει. Η συνάρτηση θα δέχεται ως ορίσματα:

- το χρονικό ορίζοντα του πειράματος (horizon).
- τον τετραγωνικό πίνακα  $P$ , διάστασης  $n = \text{πλήθος κόμβων που θέλουμε να επισκεφθούμε}$ , με το στοιχείο  $(i, j)$  του πίνακα να αντιστοιχεί στο κόστος της ακμής  $(i, j)$  σε χρόνο.
- τον τετραγωνικό πίνακα  $D$ , διάστασης  $n = \text{πλήθος κόμβων που θέλουμε να επισκεφθούμε}$ , με το στοιχείο  $(i, j)$  του πίνακα να αντιστοιχεί στο κόστος της ακμής  $(i, j)$  σε απόσταση.
- τον αριθμό των επαναλήψεων ( $M$ ) για τη συνάρτηση *my\_sa* που θα καλείται σε κάθε βήμα του αλγορίθμου.

Η συνάρτηση θα επιστρέφει μία λίστα τριών στοιχείων:

- τη λίστα ChosenPath, με στοιχεία όσες και οι μέρες του χρονικού ορίζοντα. Κάθε στοιχείο της λίστας είναι ένας πίνακας ο οποίος περιέχει το μονοπάτι που επιλέχθηκε για εκείνη τη μέρα του χρονικού ορίζοντα (στήλη 1), μαζί με το χρονικό κόστος κάθε ακμής που συμμετείχε στο επιλεγμένο μονοπάτι (στήλη 2), αλλά και του αθροιστικού κόστους (στήλη 3).

- τον πίνακα Counter, που έχει τέσσερις γραμμές και στήλες ίσες με το πλήθος των διαφορετικών μονοπατιών που επιλέχθηκαν κατά το πείραμα. Κάθε γραμμή αποτελεί το 1/4 των συνολικών ημερών και το στοιχείο (i, j) του πίνακα μας δείχνει πόσες φορές επιλέχθηκε το μονοπάτι j στο i τέταρτο του πειράματος.
- το AlgoTime, που είναι ο χρόνος που διήρκτησε η υλοποίηση του πειράματος στον ηλεκτρονικό υπολογιστή.

Για τη λειτουργία της συνάρτησής μας, χρησιμοποιούμε το πακέτο *invgamma*, όπως και στην εφαρμογή 5.1, για να παράγουμε τυχαίες τιμές για τα  $\sigma_e^2$  και  $\tilde{\sigma}^2$ , με χρήση της συνάρτησης *rinvgamma*.

Πλέον, είμαστε σε θέση να παραθέσουμε τον κώδικα για τη *my\_thompson\_tsp*.

```
my_thompson_tsp<-function(horizon, P, D, M){

  #GETTING THE TIME
  #AT WHICH THE ALGORITHM STARTS RUNNING
  start_time<-Sys.time()

  #GETTING THE NUMBER OF VERTICES
  total_vertices<-ncol(P)

  #SAMPLING VARIANCES FOR EACH EDGE
  #FROM INVERSE GAMMA DISTRIBUTIONS
  S<-matrix(rinvgamma(total_vertices*total_vertices,4,3),
            ncol=total_vertices)

  #SETTING EDGES' PRIOR DISTRIBUTION'S MEAN
  mu <- log(D) - S/2
```

```

#INITIALIZING THETA
theta <- matrix(rep(0,ncol(D)*nrow(D)), ncol=ncol(D))

#INITIALIZING THE LIST CONSISTING OF THE DIFFERENT PATHS
#FOR EACH DAY
chosen_paths<-vector(mode="list",length=horizon)

#INITIALIZING A COMBINED LIST OF CHOSEN_PATHS,
#EDGE COSTS AND CUMULATIVE COSTS
combined<-vector(mode="list",length = horizon)

#STARTING THE FOR LOOP FOR EACH DAY

for( l in (1:horizon)){

  #SAMPLING S_TILDE FROM AN INVERSE GAMMA
  S_tilde<-rinvgamma(1,17,8)

  #SAMPLING THETA FOR EACH EDGE
  #FROM LOGNORMAL DISTRIBUTIONS
  #WITH MEAN = MU AND VARIANCE = S

  for (i in 1:nrow(D)){

    for (j in 2:ncol(D)){

      theta[i,j]<- rlnorm(1,mu[i,j],S[i,j])
    }
  }
}

```

```

#RUNNING SIMULATED ANNEALING
#TO SOLVE THE TSP PROBLEM FOR THIS THETA MATRIX

sa_tsp<- my_sa(df=theta, M)

#GETTING THE OPTIMUM PATH AS A VECTOR

op<- sa_tsp$Table1

length_op<-length(op)

#INITIALIZING THE EDGE_COST VECTOR

edge_cost<-NA

#SAMPLING A COST FOR EACH EDGE OF THE OPTIMUM PATH

for ( i in 1:(length_op-1)){

  V_start<-op[i]
  V_end<-op[i+1]

  edge_cost[i]<-rlnorm(1,meanlog =
                      (log(P[V_start,V_end])-
                      (S_tilde/2)), sdlog = S_tilde)

#UPDATING THE PARAMETERS
#OF THE LOGNORMAL DISTRIBUTIONS
#OF THESE EDGE'S THETA
#WHICH WERE INCLUDED IN THE OPTIMUM PATH

a1<-(1/(S[V_start,V_end]))*(mu[V_start,V_end])
a2<-(1/(S_tilde))*(log(edge_cost[i])+(S_tilde/2))

```



```

    b<-(1/S[V_start,V_end])+(1/S_tilde)

    mu[V_start,V_end]<-(a1+a2)/b
    S[V_start,V_end] <- 1/b
  }

  #ADDING ZERO AS THE FIRST VALUE OF THE VECTOR
  #(FOR N VERTICES IN THE OPTIMUM PATH, WE HAVE N-1 EDGES)
  edge_cost<-c(0,edge_cost)

  #INITIALIZING AND BUILDING THE CUMULATIVE_COST VECTOR

  cumulative_cost<-rep(0,(length(op)))

  for (i in 2:(length(op))) {

    cumulative_cost[i]<-cumulative_cost[i-1]+edge_cost[i]
  }

  #SETTING THE OPTIMUM PATH AS THE CHOSEN PATH OF DAY 1
  chosen_paths[[1]]<-op

  #SETTING THE 1 ELEMENT OF COMBINED LIST
  #AS A MATRIX OF THREE COLUMNS:
  #OPTIMUM PATH, EDGE COST (MINUTES)
  #AND CUMULATIVE COST (MINUTES)
  combined[[1]]<-cbind(op,edge_cost,(cumulative_cost))
  colnames(combined[[1]])<-c("Vertices",
                             "Cost_of_Edge_(minutes)",
                             "Cumulative_Path_Cost_(minutes)")
}

```

```

#THE FOR LOOP FOR EACH DAY IS FINISHED

#DEFINING THE UNIQUE PATHS THAT WERE CHOSEN IN THESE 1 DAYS
unique_paths<-unique(chosen_paths)

#GETTING THE TOTAL LENGTH OF THESE UNIQUE PATHS
number_of_paths<-length(unique_paths)

#SETTING THE NAME OF EACH PATH TO BE THE PATH ITSELF
names_of_paths<-rep(1,length(unique_paths))

for ( i in 1:number_of_paths){

  lngth<-length(unique_paths[[i]])
  for ( j in 1:(lngth-1)){

    names_of_paths[i]<-paste(names_of_paths[i],
                           unique_paths[[i]][j+1])

  }

}

names(unique_paths)<-names_of_paths

#DEFINING EIGHT EQUAL TIME PERIODS
q1<-0.25*horizon
q2<-0.5*horizon
q3<-0.75*horizon
q4<-horizon

#INITIALIZING THE CORRESPONDING COUNTER FOR EACH PERIOD,
#AS A VECTOR OF LENGTH = NUMBER_OF_PATHS
counter_q1<-rep(0,number_of_paths)
counter_q2<-rep(0,number_of_paths)
counter_q3<-rep(0,number_of_paths)

```

```

counter_q4<-rep(0,number_of_paths)

#FOR LOOP FOR EACH TIME PERIOD
#ADD 1 TO THE ith VALUE OF COUNTER_Q1
#FOR EVERY TIME THAT PATH i WAS CHOSEN
#IN THE FIRST Q1 DAYS
#ADD 1 TO THE ith VALUE OF COUNTER_Q2
#FOR EVERY TIME THAT PATH i WAS CHOSEN
#BETWEEN THE DAYS Q1+1 AND Q2
#ETC.

for ( i in 1:number_of_paths){

  for (j in 1:q1){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q1[i]<-counter_q1[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

  for (j in (q1+1):q2){

    if( isTRUE(all.equal(chosen_paths[[j]],
                        unique_paths[[i]]))==TRUE){
      counter_q2[i]<-counter_q2[i]+1
    }
  }
}

for ( i in 1:number_of_paths){

  for (j in (q2+1):q3){

```

```

        if( isTRUE(all.equal(chosen_paths[[j]],
                            unique_paths[[i]]))==TRUE){
            counter_q3[i]<-counter_q3[i]+1
        }
    }
}

for ( i in 1:number_of_paths){

    for (j in (q3+1):q4){

        if( isTRUE(all.equal(chosen_paths[[j]],
                            unique_paths[[i]]))==TRUE){
            counter_q4[i]<-counter_q4[i]+1
        }
    }
}

#COMBINING ALL THE COUNTERS TO PRESENT THEM TOGETHER
counter_final<-rbind(counter_q1,counter_q2,
                    counter_q3,counter_q4)
counter_final<-as.matrix(counter_final)

colnames(counter_final)<-names_of_paths
rownames(counter_final)<-c("1/4_of_Total_Horizon",
                          "2/4_of_Total_Horizon",
                          "3/4_of_Total_Horizon",
                          "4/4_of_Total_Horizon")

names(combined)<-paste0("Day_", 1:horizon)

#GETTING THE TIME
#AT WHICH THE ALGORITHM STOPS RUNNING

```

```

end_time<-Sys.time()

#GETTING THE TOTAL TIME
#OF THE ALGORITHM
total_time<-end_time - start_time

#DEFINING THE LIST TO BE RETURNED
results<-list(ChosenPath=combined,Counter=counter_final,
              AlgoTime=total_time)

return(results)

}

```

Καλούμε τη συνάρτησή μας, για έναν χρονικό ορίζοντα 10000 δοκιμών (horizon=1000), πίνακα μέσων χρόνων  $P = df\_tsp\_times$ , πίνακα αποστάσεων  $D = df\_tsp\_distances$  και αριθμό επαναλήψεων για το SA του κάθε βήματος  $M = 1000$ . Ως πρώτη επίσημη δοκιμή, εκχωρούμε τα αποτελέσματα σε μια λίστα με όνομα `test_official_1`.

```

test_official_1<-my_thompson_tsp(horizon=10000,
                                P = df_tsp_times,
                                D = df_tsp_distances,
                                M = 1000)

```

Για αρχή, περισσότερο ενδιαφέρον παρουσιάζει ο πίνακας Counter, που όμως έχει διαστάσεις  $4 \times 198$  (198 ήταν τα μοναδικά μονοπάτια που επιλέχθηκαν τουλάχιστον 1 φορά στο πείραμα των 10000 δοκιμών) και δεν είναι πρακτικό να τον παραθέσουμε ως έχει. Ενδεικτικά, δείχνουμε τις 16 πρώτες στήλες χωρίς να προβούμε σε άμεσες αναλύσεις, απλώς και μόνο για να δείξουμε πώς μοιάζει ο πίνακας Counter.

```
tsp_counters[,1:16]
```

```
##          1 3 2 10 5 4 6 9 8 7 11 1 2 5 10 4 8 3 6 9 7 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          0          0
## 4/4 of Total Horizon          0          0
##          1 4 9 2 3 5 6 7 8 10 11 1 2 7 9 5 3 4 6 8 10 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          0          0
## 4/4 of Total Horizon          0          0
##          1 3 4 7 2 8 5 9 6 10 11 1 2 3 6 5 7 8 9 4 10 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          0          0
## 4/4 of Total Horizon          0          0
##          1 4 2 6 9 3 7 5 8 10 11 1 2 6 3 9 8 4 5 7 10 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          0          0
## 4/4 of Total Horizon          0          0
##          1 4 3 5 10 8 6 2 9 7 11 1 2 9 7 4 6 3 5 8 10 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          0          0
## 4/4 of Total Horizon          0          0
##          1 3 5 2 4 6 9 8 7 10 11 1 4 6 7 9 8 2 5 3 10 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          2          0
## 4/4 of Total Horizon          0          0
##          1 2 5 4 3 7 6 8 9 10 11 1 2 4 3 5 6 9 7 8 10 11
## 1/4 of Total Horizon          1          58
## 2/4 of Total Horizon          0          23
## 3/4 of Total Horizon          0          18
## 4/4 of Total Horizon          0          30
##          1 4 9 7 5 6 2 3 8 10 11 1 3 2 6 4 5 9 7 8 10 11
## 1/4 of Total Horizon          1          1
## 2/4 of Total Horizon          0          0
## 3/4 of Total Horizon          0          0
## 4/4 of Total Horizon          0          0
```

Σχήμα 5.15: Οι 16 πρώτες στήλες του πίνακα Counter.

Αντ' αυτού, θα χειριστούμε τα δεδομένα μας για να παρουσιάσουμε τις τρεις πιο δημοφιλείς διαδρομές για τον αλγόριθμό μας, ξεκινώντας από την πρώτη.

```
tsp_counters<-test_official_1$Counter
counter_sums<-colSums(tsp_counters)

tsp_best<-max(counter_sums)
which_tsp_best<-which(counter_sums==tsp_best)
```

```
tsp_best
```

```
## [1] 2580
```

```
which_tsp_best
```

```
## 1 4 3 2 5 6 9 7 8 10 11
##                               38
```

```
tsp_counters[,which_tsp_best]
```

```
## 1/4 of Total Horizon 2/4 of Total Horizon 3/4 of Total Horizon
##                   680                   677                   626
## 4/4 of Total Horizon
##                   597
```

Σχήμα 5.16: Δημοφιλέστερη επιλογή αλγορίθμου. Το `tsp_best` μας δείχνει πόσες φορές από τις 10000 επιλέχθηκε το εν λόγω μονοπάτι, το `which_tsp_best` μας δείχνει ποιο μονοπάτι είναι αυτό ("1 4 3 2 5 6 9 7 8 10 11") και σε ποιά στήλη του πίνακα `Counter` βρίσκεται (38). Τέλος, το `tsp_counters[,which_tsp_best]` μας δείχνει πόσες φορές επιλέχθηκε το δημοφιλέστερο μονοπάτι σε κάθε τέταρτο του χρονικού ορίζοντα.

Ανακαλούμε τα σχόλια που κάναμε κάτω από τα σχήματα 5.12 και 5.13, ανα-

λύοντας τα ντετερμινιστικά βέλτιστα χρονικά μονοπάτια και παρατηρούμε ότι το δημοφιλέστερο μονοπάτι, "1 4 3 2 5 6 9 7 8 10 11", είναι ένα από τα δύο μονοπάτια που απαιτούσαν τον ελάχιστο χρόνο των 55 λεπτών. Συνεπώς, φαίνεται ότι η συνάρτηση *my\_thompson\_tsp* πέτυχε στην αναγνώριση του βέλτιστου.

Στην ίδια λογική, πάμε να δούμε το δεύτερο και το τρίτο δημοφιλέστερο μονοπάτι του αλγορίθμου.

```
tsp_second_best<-max(counter_sums[-which_tsp_best])
which_tsp_second_best<-which(counter_sums==tsp_second_best)

tsp_third_best<-max(counter_sums[-c(which_tsp_best,
                                     which_tsp_second_best)])
which_tsp_third_best<-which(counter_sums==tsp_third_best)
```



```
tsp_second_best
```

```
## [1] 1713
```

```
which_tsp_second_best
```

```
## 1 2 5 3 4 6 9 7 8 10 11  
##                                40
```

```
tsp_counters[,which_tsp_second_best]
```

```
## 1/4 of Total Horizon 2/4 of Total Horizon 3/4 of Total Horizon  
##                    279                    485                    452  
## 4/4 of Total Horizon  
##                    497
```

Σχήμα 5.17: Δεύτερη δημοφιλέστερη επιλογή το "1 2 5 3 4 6 9 7 8 10 11", με 1713 εμφανίσεις, στη θέση 40 του πίνακα Counter.

```
tsp_third_best
```

```
## [1] 1211
```

```
which_tsp_third_best
```

```
## 1 2 3 4 6 9 7 5 8 10 11  
##                               25
```

```
tsp_counters[,which_tsp_third_best]
```

```
## 1/4 of Total Horizon 2/4 of Total Horizon 3/4 of Total Horizon  
##                   203                   235                   375  
## 4/4 of Total Horizon  
##                   398
```

Σχήμα 5.18: Τρίτη δημοφιλέστερη επιλογή το "1 2 3 4 6 9 7 5 8 10 11", με 1211 εμφανίσεις, στη θέση 25 του πίνακα Counter

Ανακαλούμε ξανά τα σχόλια κάτω από τα σχήματα 5.12 και 5.13 που σχετίζονται με τα ντετερμινιστικά χρονικά βέλτιστα μονοπάτια. Παρατηρούμε, λοιπόν, ότι η δεύτερη δημοφιλέστερη επιλογή του αλγορίθμου είναι το δεύτερο μονοπάτι ελάχιστου χρόνου (55 λεπτών), ενώ η τρίτη επιλογή είναι το αμέσως επόμενο καλύτερο, με μέσο χρόνο 56 λεπτών. Επίσης, παρατηρούμε ότι ήδη από τις πρώτες 2500 επαναλήψεις, ο αλγόριθμος έχει κατασταλλάξει στις προτιμήσεις του, καθώς η ανθροιστική επιλογή των τριών αγαπημένων του μονοπατιών είναι ίση με 1162. Με άλλα λόγια, 3 από τα 198 μονοπάτια τα διαλέγει σχεδόν το 50% των φορών. Συνεπώς, συνυπολογίζοντας πως αυτά τα 3 μονοπάτια είναι και βέλτιστα, η απόδοση του αλγορίθμου κρίνεται ικανοποιητικότερη.

Στη συνέχεια, παραθέτουμε τις επιλογές του αλγορίθμου τις 3 πρώτες μέρες του πειράματος και τις 3 τελευταίες.

```
head(test_official_1$ChosenPath,3)
```

```
## $'Day 1'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 3 5.187056 5.187056  
## [3,] 2 3.972896 9.159951  
## [4,] 10 16.474558 25.634509  
## [5,] 5 25.936341 51.570851  
## [6,] 4 9.505243 61.076094  
## [7,] 6 1.839735 62.915829  
## [8,] 9 3.353392 66.269221  
## [9,] 8 3.595833 69.865055  
## [10,] 7 4.508433 74.373488  
## [11,] 11 4.992185 79.365672  
##  
## $'Day 2'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 2 1.595949 1.595949  
## [3,] 5 3.770996 5.366945  
## [4,] 10 4.434014 9.800959  
## [5,] 4 27.851540 37.652499  
## [6,] 8 6.002672 43.655172  
## [7,] 3 5.067067 48.722238  
## [8,] 6 5.934776 54.657014  
## [9,] 9 4.453247 59.110262  
## [10,] 7 4.299756 63.410018  
## [11,] 11 22.654008 86.064026  
##  
## $'Day 3'  
## Vertices Cost Of Edge (minutes) Cumulative Path Cost (minutes)  
## [1,] 1 0.000000 0.000000  
## [2,] 4 3.644489 3.644489  
## [3,] 9 11.023599 14.668089  
## [4,] 2 14.767005 29.435093  
## [5,] 3 4.849952 34.285046  
## [6,] 5 5.041218 39.326264  
## [7,] 6 5.401470 44.727734  
## [8,] 7 6.512719 51.240453  
## [9,] 8 8.392483 59.632936  
## [10,] 10 3.871777 63.504713  
## [11,] 11 3.002747 66.507461
```

Παρατηρούμε πως τα μονοπάτια των πρώτων τριών ημερών είναι ενδεχομένως λίγο απρόσμενα, καθώς η επιλογή δεν έχει γίνει με βάση τις πιο κοντινές αποστάσεις, που θα ήταν το άμεσα λογικό. Ωστόσο, αυτό το γεγονός δε μας επηρεάζει αρνητικά, γιατί υποδηλώνει μία τάση του αλγορίθμου για άμεση τυχαία εξερεύνηση. Η πιο λογική εξήγηση είναι ότι τα  $\sigma_e^2$  που παράχθηκαν ήταν τέτοια ώστε να έχουμε μεγάλες αποκλίσεις στα πρώτα  $\hat{\theta}_k$ .

```
tail(test_official_1$ChosenPath,3)
```

```
## $'Day 9998'
```

##	Vertices	Cost Of Edge (minutes)	Cumulative Path Cost (minutes)
##	[1,]	1	0.0000000
##	[2,]	2	1.5712998
##	[3,]	3	5.9230079
##	[4,]	4	3.5158126
##	[5,]	6	0.7878267
##	[6,]	9	3.0056182
##	[7,]	7	4.4438573
##	[8,]	5	14.1427150
##	[9,]	8	6.2258129
##	[10,]	10	3.5391440
##	[11,]	11	4.7499859

```
##
```

```
## $'Day 9999'
```

##	Vertices	Cost Of Edge (minutes)	Cumulative Path Cost (minutes)
##	[1,]	1	0.0000000
##	[2,]	2	1.113846
##	[3,]	3	2.473521
##	[4,]	5	7.266805
##	[5,]	4	17.290065
##	[6,]	6	1.038930
##	[7,]	9	2.460294
##	[8,]	7	4.772088
##	[9,]	8	1.258265
##	[10,]	10	1.620935
##	[11,]	11	6.471309

```
##
```

```
## $'Day 10000'
```

##	Vertices	Cost Of Edge (minutes)	Cumulative Path Cost (minutes)
##	[1,]	1	0.0000000
##	[2,]	2	0.6699180
##	[3,]	5	3.7154434
##	[4,]	3	5.4871413
##	[5,]	4	4.4430533
##	[6,]	6	0.6580054
##	[7,]	9	6.0570185
##	[8,]	7	4.6704707
##	[9,]	8	1.1921770
##	[10,]	10	3.0422416
##	[11,]	11	4.3447365

Παρατηρούμε ότι τη μέρα 10000 έχει γίνει η επιλογή του βέλτιστου και δεύτερου δημοφιλέστερου μονοπατιού. Επίσης, τη μέρα 9998 επιλέχθηκε το τρίτο δημοφιλέστερο μονοπάτι μέσου χρόνου 56 λεπτών. Ωστόσο, τη μέρα 9999 έχουμε μία έκπληξη, από το μονοπάτι "1 2 3 5 4 6 9 7 8 10 11". Ανατρέχοντας, όμως, στα σχήματα 5.12 και 5.13, μπορούμε εύκολα να δούμε ότι δεν πρόκειται και για τόσο μεγάλη έκπληξη, καθώς είναι κι αυτό ένα μονοπάτι με μέσο χρόνο 56 λεπτών. Συνεπώς, στον πείραμά μας θα μπορούσε, για παράδειγμα, να είναι η τέταρτη ή η πέμπτη δημοφιλέστερη επιλογή, ενώ σε μια επανάληψη του ίδιου πειράματος θα μπορούσε κάλλιστα να είναι η τρίτη δημοφιλέστερη επιλογή.

Τέλος, παραθέτουμε το τρίτο στοιχείο από τη λίστα που επιστρέφει η συνάρτηση `my_thompson_tsp`, το οποίο είναι το `AlgoTime`, δηλαδή ο χρόνος που πέρασε μέχρι να υλοποιηθεί το πείραμα στον ηλεκτρονικό υπολογιστή.

```
test_official_1$AlgoTime
```

```
## Time difference of 9.045731 mins
```

Σχήμα 5.19: Ο χρόνος που χρειάστηκε για να υλοποιηθεί το πείραμα ήταν  $\simeq 9$  λεπτά.

Βλέπουμε, λοιπόν, ότι σε λιγότερο από 10 λεπτά καταφέραμε και λύσαμε ένα αρκετά σύνθετο πρόβλημα, γεγονός που προσδίδει επιπλέον αξία στον αλγόριθμο.

Προτού κλείσουμε αυτήν την υποενότητα, θα μπορούσαμε να κάνουμε τον συνήγορο του διαβόλου και να ισχυριστούμε ότι ο χρονικός ορίζοντας των 10000 ημερών είναι πολυτέλεια, δεδομένου ότι στην πραγματικότητα θα έπρεπε κάποιος να δοκιμάζει αυτά τα μονοπάτια για 10000 μέρες, γεγονός αδύνατο.

Επαναλαμβάνουμε, λοιπόν, το πείραμά μας για μόνο 100 ημέρες.

```
test_official_2<-my_thompson_tsp(horizon=100,  
                                P = df_tsp_times,
```

```
D=df_tsp_distances,  
M=1000)
```

Και ζητάμε το δημοφιλέστερο μονοπάτι.

```
tsp_counters2<-test_official_2$Counter  
counter_sums2<-colSums(tsp_counters2)  
  
tsp_best2<-max(counter_sums2)  
which_tsp_best2<-which(counter_sums2==tsp_best2)
```

```
tsp_best2
```

```
## [1] 19
```

```
which_tsp_best2
```

```
## 1 4 3 2 5 6 9 7 8 10 11  
##                               35
```

```
tsp_counters2[,which_tsp_best2]
```

```
## 1/4 of Total Horizon 2/4 of Total Horizon 3/4 of Total Horizon  
##                               0                               3                               7  
## 4/4 of Total Horizon  
##                               9
```

Σχήμα 5.20: Το αγαπημένο μονοπάτι είναι το "1 4 3 2 5 6 9 7 8 10 11", μέσου χρόνου 55 λεπτών.

Παρατηρούμε ότι παρόλο που δεν επιτρέψαμε στον αλγόριθμο επαρκή χώρο και χρόνο για να εξερευνήσει καταλλήλως, εκείνος διαλέγει το 20% των φορών

το ντετερμινιστικά χρονικά βέλιστο μονοπάτι και επιτυγχάνει ξανά. Αυτό επιβεβαιώνεται και παρακάτω, από τη δεύτερη θέση στην οποία ισοβαθμούν δύο μονοπάτια, καλού μέσου χρόνου, που όμως επιλέχθηκαν μόνο τέσσερις φορές έκαστο. Το ίδιο ισχύει και για την τρίτη θέση, στην οποία ισοβαθμούν τέσσερα μονοπάτια που επιλέχθηκαν τρεις φορές το καθένα. Με τόσες λίγες φορές επιλογής, όμως, παύει να αποτελεί «διάκριση» για τα μονοπάτια τα οποία εξυπηρετούν καθαρά τη διαδικασία της εξερεύνησης.

```
tsp_second_best2<-max(counter_sums2[-which_tsp_best2])
which_tsp_second_best2<-which(counter_sums2==tsp_second_best2)

tsp_third_best2<-max(counter_sums2[-c(which_tsp_best2,
                                       which_tsp_second_best2)])
which_tsp_third_best2<-which(counter_sums2==tsp_third_best2)
```



```
tsp_second_best2
```

```
## [1] 4
```

```
which_tsp_second_best2
```

```
## 1 2 3 4 6 9 5 7 8 10 11 1 2 3 5 4 6 9 7 8 10 11  
##                               38                               51
```

```
tsp_counters2[,which_tsp_second_best2]
```

```
##           1 2 3 4 6 9 5 7 8 10 11 1 2 3 5 4 6 9 7 8 10 11  
## 1/4 of Total Horizon           0           0  
## 2/4 of Total Horizon           1           0  
## 3/4 of Total Horizon           1           1  
## 4/4 of Total Horizon           2           3
```

Σχήμα 5.21: Στη δεύτερη θέση ισοβαθμούν τα "1 2 3 4 6 9 5 7 8 10 11" και "1 2 3 5 4 6 9 7 8 10 11" με μόνο 4 εμφανίσεις.

```
tsp_third_best2
```

```
## [1] 3
```

```
which_tsp_third_best2
```

```
## 1 4 3 2 5 6 9 8 7 10 11 1 3 2 5 4 6 9 7 8 10 11 1 4 3 2 5 8 7 6 9 10 11
##                               20                               26                               30
## 1 2 3 4 6 9 7 5 8 10 11
##                               40
```

```
tsp_counters2[,which_tsp_third_best2]
```

```
##           1 4 3 2 5 6 9 8 7 10 11 1 3 2 5 4 6 9 7 8 10 11
## 1/4 of Total Horizon           1           0
## 2/4 of Total Horizon           0           2
## 3/4 of Total Horizon           1           0
## 4/4 of Total Horizon           1           1
##           1 4 3 2 5 8 7 6 9 10 11 1 2 3 4 6 9 7 5 8 10 11
## 1/4 of Total Horizon           0           0
## 2/4 of Total Horizon           2           1
## 3/4 of Total Horizon           1           1
## 4/4 of Total Horizon           0           1
```

Σχήμα 5.22: Στην τρίτη θέση ισοβαθμούν τα "1 4 3 2 5 6 9 8 7 10 11", "1 3 2 5 4 6 9 7 8 10 11", "1 4 3 2 5 8 7 6 9 10 11" και "1 2 3 4 6 9 7 5 8 10 11" με μόνο 3 εμφανίσεις.

Με μεγάλη μας χαρά, λοιπόν, ασφαλώς συμπεραίνουμε ότι ο αλγόριθμός μας επιτυγχάνει στην επίλυση ενός Traveling Salesman Problem στο πλαίσιο του Multi-Armed Bandit με χρήση του Thompson Sampling και του Simulated Annealing, ακόμα και σε ένα διάστημα 100 μόνο ημερών.

# Βιβλιογραφία

- [1] Bertsekas, Dimitri P. and Tsitsiklis and John N, 2002. *Introduction to probability*. Athena, Scientific, Belmont.
- [2] Brémaud, Pierre, 1999. *Markov chains: Gibbs fields, Monte Carlo simulation and queues*. Texts in applied mathematics. Springer, New York, NY.
- [3] Varadhan, S. R. S., 2001. *Probability theory*. Courant Lecture Notes in Mathematics. American Mathematical Society, Providence, RI.
- [4] Χελιώτης, Δ., 2015. *Εισαγωγή στον στοχαστικό λογισμό*. [ηλεκτρ. βιβλ.] Αθήνα:Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.  
<http://hdl.handle.net/11419/4143>
- [5] Χελιώτης, Δ., 2015. *Ένα δεύτερο μάθημα στις πιθανότητες*. [ηλεκτρ. βιβλ.] Αθήνα:Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.  
<http://hdl.handle.net/11419/2825>
- [6] Doyle, Peter G, 1985. *Random Walks and Electrical Networks*. Carus Mathematical Monographs. John Wiley & sons, New York, NY.
- [7] Häggström, Olle, 2002. *Finite Markov chains and algorithmic applications*. London Mathematical Society student texts. Cambridge University Press, Cambridge, UK.
- [8] Krauth, Werner, 2006. *Statistical Mechanics: Algorithms and Computations*. Oxford Master Series in Physics. Oxford University Press, Oxford, UK.
- [9] Levin, David Asher, Peres, Yuval and Wilmer, Elizabeth Lee, 2009. *Markov chains and mixing times*. American Mathematical Society, Providence, RI.

- [10] Norris, James R, 1998. *Markov chains*. Cambridge series in statistical and probabilistic mathematics, Cambridge, UK. Cambridge University Press.
- [11] Williams, David, 1991. *Probability With Martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, UK.
- [12] Λουλάκης Μιχάλης, 2015. *Στοχαστικές Διαδικασίες*. [ηλεκτρ. βιβλ.] Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.  
<http://hdl.handle.net/11419/6003>
- [13] Γιαννακόπουλος Αθανάσιος, 2003. *Στοχαστική Ανάλυση και Εφαρμογές στη Χρηματοοικονομική, Τόμος I: Εισαγωγή Στην Στοχαστική Ανάλυση* [ηλεκτρ. βιβλ.]  
[https://mathbooksgroups.files.wordpress.com/2011/08/stox\\_diadikasies\\_ii.pdf](https://mathbooksgroups.files.wordpress.com/2011/08/stox_diadikasies_ii.pdf)
- [14] Goran Peskir and Albert Shiryaev, 2006. *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics, ETH Zürich.
- [15] Albert Shiryaev, 1978. *Optimal Stopping Rules*. Springer.
- [16] Y. S. Chow and H. Robbins, 1963. *On optimal stopping rules*, [*Z. Wahrscheinlichkeitstheorie und verw. Gebiete* 2, 33-49].
- [17] Thomas S. Ferguson. *Optimal Stopping and Applications*. [electr. text] Mathematics Department, UCLA.  
<https://www.math.ucla.edu/~tom/Stopping/Contents.html>
- [18] D. A. Berry, 1972. *A Bernoulli two-armed bandit*, [*Ann. Math. Statist.* 43, 871-897].
- [19] D. A. Berry and B. Fristedt, 1985. *Bandit Problems; Sequential Allocation of Experiments*. Chapman and Hall, New York.
- [20] M. N. Katehakis and A. F. Vienott, Jr., 1987. *The multi-armed bandit problem: Decomposition and computation*, [*Math. of Op. Res.* 12, 262-268].
- [21] P. P. Varaiya, J. C. Walrand and C. Buyukkoc, 1985. *Extensions of the multi-armed bandit problem. The discounted case*. [*IEEE Trans. Autom. Control AC-30*, 426-439].

- [22] J. C. Gittins and D. M. Jones, 1979. *A Dynamic Allocation Index for the Discounted Multiarmed Bandit Problem*, [*Biometrika* Vol. 66, No. 3, pp. 561-565].
- [23] Scott, S.L., 2010. *A modern Bayesian look at the multi-armed bandit*. [*Applied Stochastic Models in Business and Industry*, 26, 639-658].
- [24] Chapelle, O. and Li, L., 2011. *An Empirical Evaluation of Thompson Sampling*. [*NIPS*].
- [25] Thompson, W. R., 1935. *On the theory of apportionment*. [*American Journal of Mathematics*. 57(2): 450-456].
- [26] Thompson, W., 1933. *On the likelihood that one unknown probability exceeds another in view of the evidence of two samples*. [*Biometrika* 25(3/4): 285-294].
- [27] Daniel J. Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband and Zheng Wen, 2018. *A Tutorial on Thompson Sampling*. [*Foundations and Trends® in Machine Learning: Vol. 11: No. 1, pp 1-96*].  
<http://dx.doi.org/10.1561/22000000070>
- [28] Russo, D. and B. Van Roy, 2014a. *Learning to optimize via information-directed sampling*. [*Advances in Neural Information Processing Systems* 27. 1583-1591].
- [29] Russo, D. and B. Van Roy, 2014b. *Learning to optimize via posterior sampling*. [*Mathematics of Operations Research*. 39(4): 1221-1243].
- [30] Russo, D. and B. Van Roy, 2016. *An Information-Theoretic analysis of Thompson sampling*. [*Journal of Machine Learning Research*. 17(68): 1-30].
- [31] Russo, D., 2016. *Simple bayesian algorithms for best arm identification*. [*Conference on Learning Theory*. 1417-1418].
- [32] Russo, D. and B. Van Roy, 2018a. *Learning to optimize via information-directed sampling*. [*Operations Research*. 66(1): 230-252.]
- [33] Strens, Malcolm J. A., 2000. *A Bayesian Framework for Reinforcement Learning*. [*ICML*].

- [34] Ripley, B.D., 1987. *Stochastic Simulation*. John Wiley & Sons, New Jersey.
- [35] Efron, B. and Hastie, T., 2016. *Computer Age Statistical Inference Algorithms, Evidence, and Data Science*. Cambridge University Press, New York.

