Εισαγωγή στα Martingales

Αντώνης Οικονόμου - Μάιος 2012

1. Εισαγωγη

Τα martingales αποτελούν μια πολύ σημαντική κατηγορία στοχαστικών διαδικασιών, αφενός διότι έχουν θεωρητικό ενδιαφέρον και αφετέρου επειδή εμφανίζονται σε πολλά πεδία εφαρμογών (μαθηματική χρηματοοικονομία, μαθηματική επιδημιολογία κλπ.). Σκοπός αυτών των σημειώσεων είναι να παρέχουν μια εισαγωγή στα martingales, αναπτύσσοντας τη θεωρία τους και ταυτόχρονα παρουσιάζοντας μια επιλογή από τις εφαρμογές τους. Η αυστηρή ανάπτυξη της θεωρίας των martingales απαιτεί τη χρησιμοποίηση εννοιών από τη θεμελίωση της Θεωρίας Πιθανοτήτων μέσω της Θεωρίας Μέτρου. Επομένως, για την καλύτερη κατανόηση των σημειώσεων είναι επιθυμητό ο αναγνώστης να έχει γνώση αυτής της θεμελίωσης. Παρόλα αυτά, στα πρώτα κεφάλαια γίνεται μια σύντομη επισκόπηση των αποτελεσμάτων από τη θεωρητική θεμελίωση της Θεωρίας Πιθανοτήτων που χρειαζόμαστε για την ανάπτυξη της θεωρίας των martingales. Με τον τρόπο αυτό, οι σημειώσεις πιθανά να γίνουν προσιτές και για όσους έχουν τις γνώσεις που διδάσκονται σε ένα προπτυχιακό μάθημα Θεωρίας Πιθανοτήτων, χωρίς την αυστηρή θεμελίωση μέσω της Θεωρίας Μέτρου. Ασφαλώς, στην περίπτωση αυτή, η κατανόηση των διάφορων εννοιών και οι αποδείξεις απαιτούν πρόσθετη προσπάθεια.

Οι σημειώσεις αναπτύχθηκαν για την κάλυψη των διδακτικών αναγκών μέρους του μαθήματος 'Στοχαστικές Ανελίξεις ' που διδάσκεται στο Μεταπτυχιακό Πρόγραμμα Σπουδών στη Στατιστική και στην Επιχειρησιακή Έρευνα του Μαθηματικού Τμήματος του Πανεπιστημίου Αθηνών. Δεν διεκδικούν πρωτοτυπία, αλλά απλά αποτελούν μια εισαγωγή που προκύπτει από τη σύνθεση των διαφόρων πηγών που αναφέρονται στη βιβλιογραφία. Είναι σίγουρα γεμάτες λάθη και παραλείψεις μια και αποτελούν μια πρώτη πρόχειρη έκδοση που δημοσιεύτηκε στην ηλεκτρονική τάξη του μαθήματος κάτω από ασφυκτικά χρονικά περιθώρια. Θα ήμουν ευγνώμων στον οποιονδήποτε μου υποδείξει τέτοια λάθη και παραλείψεις. Οποιοδήποτε σχόλιο ή πρόταση για βελτίωση είναι ιδιαίτερα ευπρόσδεκτο.

2. Επισκοπηση της Θεωριας Πιθανοτητων

2.1. **Χώροι πιθανότητας - Τυχαίες μεταβλητές - Μέση τιμή.** Η Θεωρία Πιθανοτήτων ασχολείται με την ποσοτική μελέτη διαδικασιών με αβέβαιο αποτέλεσμα. Τέτοιες διαδικασίες αναφέρονται ως 'πειράματα τύχης ' και σκοπός της Θεωρίας Πιθανοτήτων είναι να μετρήσει τις πιθανότητες πραγματοποίησης των διάφορων αποτελεσμάτων - ενδεχομένων που μπορεί να συμβούν (ό,τι κι αν σημαίνει διαισθητικά αυτό). Οι θεμελιώδεις έννοιες για μια αυστηρή θεμελίωση της Θεωρίας Πιθανοτήτων είναι ο δειγματικός χώρος, η οικογένεια των ενδεχομένων, το μέτρο πιθανότητας και η τυχαία μεταβλητή.

Ορισμός 1. *Έστω Ω ένα συνολο και $\mathcal{F}$ μια οικογένεια υποσυνόλων του Ω με τις ιδιότητες*

(1) $\emptyset \in \mathcal{F}$ *(η οικογένεια περιλαμβάνει το κενό σύνολο),*
(2) *Αν $A \in \mathcal{F}$ τότε και $A^c \in \mathcal{F}$ (η οικογένεια είναι κλειστή ως προς τα συμπληρώματα),*
(3) *Αν $A_1, A_2, \ldots \in \mathcal{F}$, τότε $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ (η οικογένεια είναι κλειστή ως προς τις αριθμήσιμες ενώσεις).*

Το σύνολο Ω αναφέρεται ως δειγματικός χώρος και η $\mathcal{F}$ λέγεται σ-άλγεβρα ή σ-πεδίο ενδεχομένων του Ω.

Μπορούμε να σκεφτόμαστε ότι ο χώρος Ω είναι το σύνολο των δυνατών αποτελεσμάτων ενός πειράματος τύχης, ενώ η οικογένεια $\mathcal{F}$ περιλαμβάνει τα 'ενδεχόμενα' του πειράματος τύχης. Ένα ενδεχόμενο είναι στην ουσία μια ιδιότητα που μπορούμε να αποφανθούμε αν ισχύει ή όχι μετά την πραγματοποίηση του πειράματος. Σε πολλά πειράματα τύχης ξεκινάμε από κάποια οικογένεια συνόλων $\mathcal{C}$ που επιθυμούμε να περιλαμβάνει κάποια συγκεκριμένα σύνολα-ενδεχόμενα. Η οικογένεια αυτή, όμως, μπορεί όμως να μην είναι σ-άλγεβρα. Στην περίπτωση αυτή, για να μπορέσουμε να προχωρήσουμε, εργαζόμαστε με την ελάχιστη σ-άλγεβρα που περιέχει τα σύνολα της $\mathcal{C}$ (η τομή σ-αλγεβρών ειναι επίσης σ-άλγεβρα και επομένως έχει νόημα να μιλάμε για την ελάχιστη σ-άλγεβρα που περιέχει όλα τα στοιχεία της $\mathcal{C}$). Αυτή η σ-άλγεβρα συμβολίζεται με $\sigma(\mathcal{C})$. Σε πειράματα τύχης που ο ο Ω είναι ο $\mathbb{R}^{\ltimes}$ (ή κάποιο υποσύνολό του) χρειάζεται να θεωρήσουμε την ελάχιστη σ-άλγεβρα που περιέχει όλα τα ανοικτά υποσύνολα του $\mathbb{R}^{\ltimes}$. Αυτή η σ-άλγεβρα γράφεται ως $\mathcal{B}(\mathbb{R}^{\ltimes})$ και αναφέρεται ως η Borel σ-άλγεβρα του $\mathbb{R}^{\ltimes}$. Για $n = 1$ γράφουμε απλά $\mathcal{B}(\mathbb{R})$. Ένα χρήσιμο αποτέλεσμα είναι το ότι

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(-\infty, x]\}_{x\in\mathbb{Q}}).$$

Ορισμός 2. *Έστω Ω ένα συνολο, $\mathcal{F}$ μια σ-άλγεβρα επί του Ω και $P : \mathcal{F} \to [0, 1]$ μια συνολοσυνάρτηση που ικανοποιεί τις ιδιότητες*

(1) $P(\Omega) = 1$ *(κανονικοποιημένη),*
(2) *Για κάθε αριθμήσιμη ακολουθία $A_1, A_2, \ldots$ ξένων ανά δύο συνόλων της $\mathcal{F}$ ισχύει $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ (σ-προσθετικότητα).*

Η συνολοσυνάρτηση P λέγεται μέτρο πιθανότητας επί του Ω και η τριάδα $(\Omega, \mathcal{F}, P)$ λέγεται χώρος πιθανότητας.

Μπορούμε να σκεφτόμαστε ότι η συνολοσυνάρτηση P αποδίδει πιθανότητα στα διάφορα ενδεχόμενα του πειράματος τύχης, δηλαδή μετρά την πιθανότητα να έχουν συμβεί μετά την εκτέλεση του πειράματος. Στο βέβαιο ενδεχόμενο Ω αποδίδεται πιθανότητα 1. Επίσης απαιτείται ότι αν έχουμε μια ακολουθία από αμοιβαία αποκλειόμενες ιδιότητες του πειράματος τύχης, τότε η πιθανότητα το αποτέλεσμα του πειράματος να έχει μια από αυτές ισούται με το άθροισμα των πιθανοτήτων τους.

Ορισμένες χρήσιμες και ευρέως χρησιμοποιούμενες ιδιότητες της συνολοσυνάρτησης της πιθανότητας συνοψίζονται στην ακόλουθη πρόταση.

Πρόταση 1. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας. Τότε ισχύουν τα ακόλουθα:*

(1) *Αν A και B ενδεχόμενα της $\mathcal{F}$ με $A \subseteq B$, τότε $P(A) \leq P(B)$ και $P(B \backslash A) = P(B) - P(A)$.*

(2) *Για κάθε ακολουθία ενδεχομένων $(A_n)_{n\geq 1}$ της $\mathcal{F}$, ισχύει ότι*

$$P(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n).$$

(3) *Για κάθε αύξουσα ακολουθία ενδεχομένων $(A_n)_{n\geq 1}$ της $\mathcal{F}$ (δηλαδή $A_n \subseteq A_{n+1}$, $n \geq 1$), ισχύει ότι*

$$P(\cup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

(4) *Για κάθε φθίνουσα ακολουθία ενδεχομένων $(A_n)_{n\geq 1}$ της $\mathcal{F}$ (δηλαδή $A_n \supseteq A_{n+1}$, $n \geq 1$), ισχύει ότι*

$$P(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

Ορισμός 3. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $X : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$ μια συνάρτηση επί του Ω με την ιδιότητα $X^{-1}(B) \in \mathcal{F}$, για κάθε σύνολο $B \in \mathcal{B}(\mathbb{R})$. Η X λέγεται τυχαία μεταβλητή επί του $(\Omega, \mathcal{F}, P)$ ή $\mathcal{F}$-μετρήσιμη (στη γλώσσα της Θεωρίας Μέτρου).*

Η συνάρτηση $P_X : \mathcal{B}(\mathbb{R}^n) \to [0,1]$ με

$$P_X(B) = P(X^{-1}(B)),\ B \in \mathcal{B}(\mathbb{R}^n)$$

αναφέρεται ως κατανομή της X ή επαγόμενο μέτρο της X.

Η συνάρτηση $F_X : \mathbb{R} \to [0,1]$ με $F_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\})$ αναφέρεται ως συνάρτηση κατανομής της X.

Λόγω του ότι $\mathcal{B}(\mathbb{R}) = \sigma(\{(-\infty, x]\}_{x\in\mathbb{Q}})$, αρκεί η γνώση της συνάρτησης κατανομής στους ρητούς για να προσδιοριστεί η κατανομή της X.

Μπορούμε να σκεφτόμαστε ότι μια τυχαία μεταβλητή είναι ένα ποσοτικό χαρακτηριστικό του πειράματος. Έτσι για κάθε δυνατό αποτέλεσμα ω το χαρακτηριστικό παίρνει μια τιμή, την $X(\omega)$. Η απαίτηση $X^{-1}(B) \in \mathcal{F}$, για κάθε σύνολο $B \in \mathcal{B}(\mathbb{R})$ σημαίνει ότι για κάθε 'καλό' υποσύνολο B του $\mathbb{R}$ ('καλό'=Borel), μπορούμε να αποδόσουμε πιθανότητα στο ενδεχόμενο η τιμή του χαρακτηριστικού του πειράματος να ανήκει στο B.

Δοθείσης μιας συνάρτησης $X : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$, το σύνολο

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$$

αποδεικνύεται ότι είναι μια σ-άλγεβρα επί του Ω, που αναφέρεται ως παραγόμενη σ-άλγεβρα από τη X. Μια συνάρτηση $X : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$ είναι τυχαία μεταβλητή στον χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$ αν $\sigma(X) \subseteq \mathcal{F}$. Με άλλα λόγια, η $\sigma(X)$ είναι η μικρότερη σ-άλγεβρα που μπορεί να οριστεί στον Ω ώστε η X να είναι τυχαία μεταβλητή.

Γενικότερα, δοθείσης μιας οικογένειας συναρτήσεων $\{X_i\}_{i\in I}$ επί του Ω με τιμές στο $\mathbb{R} \cup \{-\infty, \infty\}$, ως παραγόμενη σ-άλγεβρα της $\{X_i\}_{i\in I}$ ορίζεται η

$$\sigma(\{X_i\}_{i\in I}) = \sigma(\cup_{i\in I}\sigma(X_i)),$$

που είναι η ελάχιστη σ-άλγεβρα που κάνει όλες τις συναρτήσεις της $\{X_i\}_{i\in I}$ τυχαίες μεταβλητές. Ένα χρήσιμο αποτέλεσμα είναι το ακόλουθο:

Πρόταση 2. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $X, X_1, X_2, \ldots, X_n$ τυχαίες μεταβλητές επί του $(\Omega, \mathcal{F}, P)$. Η X είναι $\sigma(X_1, X_2, \ldots, X_n)$-μετρήσιμη, αν και μόνο αν υπάρχει $f : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{B}(\mathbb{R}^n)$-μετρήσιμη ώστε $X = f(X_1, X_2, \ldots, X_n)$.*

Η οικογένεια των τυχαίων μεταβλητών που είναι ορισμένες σε ένα χώρο πιθανότητας είναι κλειστή ως προς όλες τις συνήθεις αλγεβρικές πράξεις. Δηλαδή, αν έχουμε X, Y τυχαίες μεταβλητές και $a \in \mathbb{R}$, τότε όλες οι συνήθεις συναρτήσεις των X, Y (άθροισμα, διαφορά, γινόμενο, πηλίκο, maximum ή minimum) καθώς και το βαθμωτό γινόμενο aX είναι επίσης τυχαίες μεταβλητές. Επιπλέον, δοθείσης μιας ακολουθίας τυχαίων μεταβλητών $(X_n)_{n\geq 0}$, τα inf, sup, $\liminf$ και $\limsup$ είναι επίσης τυχαίες μεταβλητές (και συνεπώς και το κατά σημείο όριο της ακολουθίας όταν υπάρχει).

Στα στοιχειώδη μαθήματα Θεωρίας Πιθανοτήτων η προσοχή εστιάζεται σε δυο είδη τυχαίων μεταβλητών, τις διακριτές και τις απόλυτα συνεχείς τυχαίες μεταβλητές. Μια τυχαία μεταβλητή X είναι διακριτή αν με πιθανότητα 1 παίρνει τιμές σε ένα αριθμήσιμο σύνολο A. Στην περίπτωση αυτή η συνάρτηση κατανομής της προσδιορίζεται πλήρως από τη συνάρτηση $f_X(x)$, $x \in A$ με $f_X(x) = P(X = x)$, $x \in A$. Η $f_X(x)$ αναφέρεται ως συνάρτηση πιθανότητας της X. Μια τυχαία μεταβλητή X λέγεται απόλυτα συνεχής αν η συνάρτηση κατανομής της δίνεται ως $F_X(x) = \int_0^x f_X(x)dx$, $x \in \mathbb{R}$. Η $f_X(x)$ αναφέρεται ως συνάρτηση πυκνότητας πιθανότητας της X.

Η τυχαία μεταβλητή είναι ένα πολύπλοκο αντικείμενο υπό την έννοια ότι για να τη γνωρίζουμε είναι αναγκαίο να γνωρίζουμε τις τιμές της σε όλο το σύνολο Ω. Στις εφαρμογές είναι σημαντικό να έχουμε μια πληροφορία όσο το δυνατό πιο συνοπτική για το χαρακτηριστικό ενός πειράματος τύχης, χρησιμοποιώντας έναν μόνο αριθμό. Την ανάγκη αυτή εξυπηρετεί η μέση ή αναμενόμενη τιμή μιας τυχαίας μεταβλητής.

Η μέση τιμή μιας τυχαίας μεταβλητής X είναι διαισθητικά ο καταλληλότερος αριθμός για περιγράψουμε τη 'θέση' γύρω από την οποία κινούνται οι τιμές της X σε όλο τον Ω, αν αυτοπεριοριστούμε στο να χρησιμοποιήσουμε έναν μόνο αριθμό. Μαθηματικά, η μέση τιμή μιας τυχαίας μεταβλητής είναι το ολοκλήρωμα Lebesgue της τυχαίας μεταβλητής ως προς το μέτρο πιθανότητας. Οι υπολογισμοί που χρησιμοποιούμε στις πιθανότητες δεν απαιτούν να ανατρέχουμε στον ορισμό του ολοκληρώματος Lebesgue. Σχεδόν πάντα, αρκεί να χρησιμοποιούμε τις ιδιότητες της μη-αρνητικότητας και της γραμμικότητας της μέσης τιμής.

Ο ορισμός της μέσης τιμής, χτίζεται σταδιακά ξεκινώντας από 0-1 τυχαίες μεταβλητές (δείκτριες τυχαίες μεταβλητές), συνεχίζοντας σε απλές, περνώντας σε μη αρνητικές και καταλήγοντας σε αυθαίρετες τυχαίες μεταβλητές. Πιο συγκεκριμένα έχουμε τον ακόλουθο ορισμό.

Ορισμός 4. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $X : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$ μια τυχαία μεταβλητή. Η X λέγεται δείκτρια του ενδεχομένου $A \in \mathcal{F}$, συμβολικά $X = 1_A$, αν $X(\omega) = 1$, για $\omega \in A$ και $X(\omega) = 0$, για $\omega \notin A$. Στην περίπτωση αυτή ορίζουμε τη μέση τιμή της X ως*

$$E[1_A] = P(A).$$

Η X λέγεται απλή αν γράφεται ως πεπερασμένος γραμμικός συνδυασμός από δείκτριες συναρτήσεις ενδεχομένων, δηλαδή αν

$$X = \sum_{i=1}^{n} c_i 1_{A_i}.$$

Στην περίπτωση αυτή ορίζουμε τη μέση τιμή της X ως

$$E[X] = \sum_{i=1}^{n} c_i P(A_i).$$

Η X λέγεται μη-αρνητική αν $X(\omega) \geq 0$ για κάθε $\omega \in \Omega$. Στην περίπτωση αυτή ορίζουμε τη μέση τιμή της X ως

$$E[X] = \sup\{E[Y] : Y \leq X, Y \text{ είναι απλή}\}.$$

Αν X είναι μια οποιαδήποτε τυχαία μεταβλητή ορίζουμε τις μη-αρνητικές τυχαίες μεταβλητές

$$X^+ = X1_{\{X \geq 0\}},\ X^- = -X1_{\{X \leq 0\}},$$

για τις οποίες έχουν οριστεί οι αντίστοιχες μέσες τιμές $E[X^+]$ και $E[X^-]$. Στην περίπτωση αυτή ορίζουμε τη μέση τιμή της X ως

$$E[X] = E[X^+] - E[X^-],$$

με την προϋπόθεση ότι δεν ισχύει $E[X^+] = E[X^-] = \infty$.

Για την απόδειξη των ιδιοτήτων της μέσης τιμής καθώς και άλλων αποτελεσμάτων στα οποία εμπλέκεται η έννοια της μέσης τιμής, αποδεικνύουμε πρώτα ότι το αποτέλεσμα ισχύει για δείκτριες τυχαίες μεταβλητές, μετά για απλές, μετά για μη-αρνητικές και τέλος για γενικές τυχαίες μεταβλητές. Για το τρίτο βήμα τέτοιων αποδείξεων (για να περάσουμε δηλαδή από τις απλές στις μη-αρνητικές) είναι χρήσιμο το ακόλουθο αποτέλεσμα.

Πρόταση 3. *Έστω $X : \Omega \to [0, \infty]$ τυχαία μεταβλητή. Τότε υπάρχει μια αύξουσα ακολουθία μη αρνητικών, απλών τυχαίων μεταβλητών $(X_n)_{n\geq 0}$ ώστε $X = \lim_{n\to\infty} X_n$.*

Οι δυο βασικές ιδιότητες της μέσης τιμής που χρησιμοποιούμε στους υπολογισμούς πιθανοτήτων είναι οι ιδιότητες της μη-αρνητικότητας και της γραμμικότητας που παρουσιάζονται στην παρακάτω πρόταση.

Πρόταση 4. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας, X, Y τυχαίες μεταβλητές και $a, b \in \mathbb{R}$. Τότε*

(1) *Αν $P(X \geq 0) = 1$, τότε $E[X] \geq 0$.*
(2) *Αν $P(X \geq 0) = 1$ και $E[X] = 0$, τότε $P(X = 0) = 1$.*
(3) *$E[aX + bY] = aE[X] + b[Y]$.*

Η απόδειξη της πρότασης αυτής δεν είναι δύσκολη, αλλά δεν είναι και προφανής. Όπως και ο ορισμός της μέσης τιμής, πρέπει να γίνει σταδιακά προχωρώντας από δείκτριες σε απλές, μετά σε μη-αρνητικές και τελικά σε αυθαίρετες τυχαίες μεταβλητές.

Για τις διακριτές και τις απόλυτα συνεχείς τυχαίες μεταβλητές που εμφανίζονται στα στοιχειώδη μαθήματα Θεωρίας Πιθανοτήτων, ο γενικός ορισμός που δόθηκε παραπάνω επεκτείνει τον ορισμό της μέσης τιμής που έχει δοθεί για αυτές. Συγκεκριμένα μπορούμε με το γενικότερο ορισμό να αποδείξουμε ότι αν X είναι διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας $f_X(x)$, τότε η μέση της τιμή δίνεται από τη σχέση $E[X] = \sum_x x f_X(x)$. Ανάλογα, αν X είναι απόλυτα συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f_X(x)$, τότε η μέση της τιμή δίνεται από τη σχέση $E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$.

Η μέση τιμή είναι ένα 'μέτρο θέσης ' με την έννοια ότι δίνει 'έναν αριθμό γύρω από τον οποίο παίρνει τιμές ' η τυχαία μεταβλητή. Για το 'πόσο πυκνά ή αραιά είναι κατανεμημένες οι τιμές ' μιας τυχαίας μεταβλητής χρειαζόμαστε ένα 'μέτρο διακύμανσης '. Το ρόλο αυτό παίζει η διασπορά μιας τυχαίας μεταβλητής που ορίζουμε αμέσως.

Ορισμός 5. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $X : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$ μια τυχαία μεταβλητή με $E[|X|] < \infty$. Η διασπορά της X ορίζεται να είναι ο αριθμός*

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

Εύκολα αποδεικνύεται ο εναλλακτικός τύπος υπολογισμού της διασποράς

$$Var[X] = E[X^2] - E[X]^2.$$

2.2. **Ανεξαρτησία.** Η θεμελιώδης έννοια που διαφοροποιεί τη Θεωρία Πιθανοτήτων από τη Θεωρία Μέτρου και δίνει άλλη δυναμική στην ανάπτυξή της είναι η έννοια της ανεξαρτησίας. Έχουμε τους ακόλουθους ορισμούς που θεμελιώνουν την έννοια της ανεξαρτησίας για ενδεχόμενα, σ-άλγεβρες και τυχαίες μεταβλητές.

Ορισμός 6. *Έστω χώρος πιθανότητας $(\Omega, \mathcal{F}, P)$ και $\{A_i\}_{i\in I} \subseteq \mathcal{F}$ μια οικογένεια ενδεχομένων.*

(1) *Τα $\{A_i\}_{i\in I}$ λέγονται ανεξάρτητα, αν για κάθε $k \geq 2$ και $i_1, i_2, \ldots, i_k$ διαφορετικά ανά δύο στοιχεία του I έχουμε*

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}).$$

(2) *Τα $\{A_i\}_{i\in I}$ λέγονται ανεξάρτητα ανά δύο, αν για κάθε i_1, i_2 διαφορετικά στοιχεία του I έχουμε*

$$P(A_{i_1} \cap A_{i_2}) = P(A_{i_1})P(A_{i_2}).$$

Ορισμός 7. *Έστω χώρος πιθανότητας $(\Omega, \mathcal{F}, P)$ και $\{\mathcal{G}_i\}_{i\in I}$ μια οικογένεια σ-αλγεβρών με $\mathcal{G}_i \subseteq \mathcal{F}$, $i \in I$. Οι $\{\mathcal{G}_i\}_{i\in I}$ λέγονται ανεξάρτητες αν για κάθε επιλογή $\{A_i\}_{i\in I}$ συνόλων με $A_i \in \mathcal{G}_i$, $i \in I$, τα $\{A_i\}_{i\in I}$ είναι ανεξάρτητα.*

Ορισμός 8. *Έστω χώρος πιθανότητας $(\Omega, \mathcal{F}, P)$ και $\{X_i\}_{i\in I}$ μια οικογένεια τυχαίων μεταβλητών. Οι τυχαίες μεταβλητές $\{X_i\}_{i\in I}$ λέγονται ανεξάρτητες αν οι παραγόμενες σ-άλγεβρες $\{\sigma(X_i)\}_{i\in I}$ είναι ανεξάρτητες.*

Στο πλαίσιο των ενδεχομένων και των πιθανοτήτων, η ανεξαρτησία λέει 'πολλαπλασίασε' για να βρεις την πιθανότητα. Στο πλαίσιο των τυχαίων μεταβλητών και των μέσων τιμών, η ανεξαρτησία λέει 'πολλαπλασίασε' για να βρεις τη μέση τιμή. Πιο συγκεκριμένα, έχουμε την ακόλουθη πρόταση.

Πρόταση 5. *Έστω X, Y τυχαίες μεταβλητές με $X, Y \geq 0$ ή $E[|X|], E[|Y|] < \infty$. Τότε*

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

Η παραπάνω πρόταση αποδεικνύεται εύκολα, ακολουθώντας τη συνήθη διαδικασία, δηλαδή ξεκινώντας από δείκτριες τυχαίες μεταβλητές και περνώντας σε απλές, μη-αρνητικές και τελικά σε γενικές τυχαίες μεταβλητές.

2.3. **Βασικές ανισότητες στη Θεωρία Πιθανοτήτων.** Στη Θεωρία Πιθανοτήτων ένα σημαντικό εργαλείο είναι διάφορες ανισότητες που χρησιμοποιούνται συχνά τόσο για την απόδειξη θεωρητικών αποτελεσμάτων όσο και στις εφαρμογές. Στις παρακάτω προτάσεις παρουσιάζουμε τις βασικότερες από αυτές.

Πρόταση 6. *Ανισότητα Markov: Αν $c > 0$ και X τυχαία μεταβλητή με $X \geq 0$ και $E[X] < \infty$, τότε*

$$P(X \geq c) \leq \frac{E[X]}{c}.$$

Πρόταση 7. *Ανισότητα Chebyshev: Αν $c > 0$ και X τυχαία μεταβλητή με $E[X^2] < \infty$, τότε*

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

Πρόταση 8. *Ανισότητα Jensen: Έστω X τυχαία μεταβλητη, $I \subseteq \mathbb{R}$ διάστημα και $\phi : I \to \mathbb{R}$ κυρτή συνάρτηση με $P(X \in I) = 1$, $E[|X|] < \infty$ και $E[|\phi(X)|] < \infty$. Τότε*

$$\phi(E[X]) \leq E[\phi(X)].$$

Πρόταση 9. *Ανισότητα Hölder: Έστω X, Y τυχαίες μεταβλητές με $E[|X|^p] < \infty$ και $E[|Y|^q] < \infty$, όπου $p > 1$ και $\frac{1}{p} + \frac{1}{q} = 1$. Τότε*

$$|E[XY]| \leq E[|XY|] \leq E[|X|^p]^{1/p} E[|Y|^q]^{1/q}.$$

Πρόταση 10. *Ανισότητα Lyapounov: Έστω X τυχαία μεταβλητή και $r > s \geq 1$. Τότε*

$$E[|X|^r]^{1/r} \geq E[|X|^s]^{1/s}.$$

Πρόταση 11. *Ανισότητα Minkowski: Έστω X, Y τυχαίες μεταβλητές και $p \geq 1$ ώστε $E[|X|^p] < \infty$ και $E[|Y|^p] < \infty$. Τότε*

$$E[|X + Y|^p]^{1/p} \leq E[|X|^p]^{1/p} + E[|Y|^p]^{1/p}.$$

2.4. **Σύγκλιση με πιθανότητα 1 και μέση τιμή.** Όπως είδαμε οι τυχαίες μεταβλητές είναι συναρτήσεις με πεδίο ορισμού το δειγματικό χώρο. Επομένως μπορούμε να μεταφέρουμε έννοιες σύγκλισης από την μαθηματική ανάλυση στην οποία ορίζονται ποικίλες μορφές σύγκλισης ακολουθιών συναρτήσεων. Το ανάλογο της κατά σημείο σύγκλισης συναρτήσεων για τυχαίες μεταβλητές είναι η σύγκλιση με πιθανότητα 1. Έχουμε τον ακόλουθο ορισμό.

Ορισμός 9. *Μια ακολουθία τυχαίων μεταβλητών $(X_n)_{n \geq 0}$ συγκλίνει με πιθανότητα 1 ή σχεδόν βέβαια ή σχεδόν παντού (almost surely - a.s) στην τυχαία μεταβλητή X αν*

$$P(\{\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)\}) = 1.$$

Γράφουμε τότε

$$\lim_{n\to\infty} X_n = X \, a.s. \text{ ή } \lim_{n\to\infty} X_n = X \text{ με πιθανότητα 1 ή } X_n \stackrel{a.s.}{\to} X.$$

Η σύγκλιση με πιθανότητα 1 αποδεικνύεται συχνά με επιχειρήματα που βασίζονται στη θεώρηση μιας τυπικής πραγματοποίησης της ακολουθίας των τυχαίων μεταβλητών που μελετώνται. Κατόπιν, έχοντας τη σύγκλιση της ακολουθίας των τυχαίων μεταβλητών με πιθανότητα 1, τίθεται το ζήτημα κατά πόσον η σύγκλιση 'περνάει' και στις αντίστοιχες μέσες τιμές. Δηλαδή, πότε μπορούμε να συνάγουμε τη σχέση $\lim_{n\to\infty} E[X_n] = E[X]$ από τη σχέση $\lim_{n\to\infty} X_n = X$ a.s.;

Υπάρχουν τρία βασικά αποτελέσματα τα οποία απαντάνε στο ζήτημα αυτό και τα οποία χρησιμοποιούνται πολύ συχνά τόσο για την απόδειξη θεωρητικών αποτελεσμάτων όσο και στις εφαρμογές. Καθένα από αυτά επιτρέπει να εναλλάξουμε τους τελεστές της μέσης τιμής και του ορίου ακολουθίας κάτω από συγκεκριμένες συνθήκες.

Πρόταση 12. *Θεώρημα Μονότονης Σύγκλισης: Αν η $(X_n)_{n\geq 0}$ είναι αύξουσα ακολουθία μη-αρνητικών τυχαίων μεταβλητών (ή γενικότερα φραγμένων κάτω από σταθερά), τότε αυτή συγκλίνει με πιθανότητα 1 και*

$$\lim_{n\to\infty} E[X_n] = E[\lim_{n\to\infty} X_n].$$

Τα όρια μπορεί να παίρνουν και το ∞ ως δυνατή τιμή.

Πρόταση 13. *Λήμμα Fatou: Αν η $(X_n)_{n\geq 0}$ είναι ακολουθία μη-αρνητικών τυχαίων μεταβλητών (ή γενικότερα φραγμένων κάτω από σταθερά), τότε*

$$E[\liminf_{n\to\infty} X_n] \leq \liminf_{n\to\infty} E[X_n].$$

Τα όρια μπορεί να παίρνουν και το ∞ ως δυνατή τιμή.

Πρόταση 14. *Θεώρημα Κυριαρχημένης Σύγκλισης: Αν $X_n \stackrel{a.s.}{\to} X$ και υπάρχει τυχαία μεταβλητή Y με $E[Y] < \infty$ και $|X_n| \leq Y$ για κάθε $n = 0, 1, 2, \ldots$, τότε $E[|X|] < \infty$ και*

$$\lim_{n\to\infty} E[X_n] = E[\lim_{n\to\infty} X_n] = E[X].$$

Η σειρά με την οποία δίνονται τα παραπάνω αποτελέσματα αντιστοιχεί στη σειρά με την οποία παρουσιάζονται συνήθως στην ανάπτυξη της Θεωρίας Μέτρου στα περισσότερα σχετικά συγγράμματα. Από πλευράς συχνότητας χρήσης το πιο σημαντικό αποτέλεσμα φαίνεται να είναι το θεώρημα κυριαρχημένης σύγκλισης. Πράγματι, οι συνθήκες για την εφαρμογή του παρουσιάζονται συχνότερα στην πράξη από τη μονοτονία της ακολουθίας που απαιτεί το θεώρημα μονότονης σύγκλισης. Επιπλέον, το θεώρημα κυριαρχημένης σύγκλισης δίνει ισότητα αντί για την ανισότητα που δίνει το λήμμα Fatou. Από την άλλη μεριά, το λήμμα Fatou είναι πολύ χρήσιμο όταν θέλουμε να δείξουμε ότι το όριο μιας ακολουθίας τυχαίων μεταβλητών έχει πεπερασμένη μέση τιμή και έχουμε πληροφορίες για τις (περιθώριες) κατανομές των όρων της ακολουθίας, αλλά όχι για τις από κοινού κατανομές τους. Π.χ. αν γνωρίζουμε ότι $X_n \geq 0$ για κάθε n, $X_n \overset{a.s.}{\to} X$ και $E[X_n] \leq 1$ για κάθε n, το λήμμα Fatou μας δίνει ότι η $E[X]$ υπάρχει και επιπλέον $E[X] \leq 1$. Για να αιτιολογήσουμε το ίδιο αποτέλεσμα, χρησιμοποιώντας το θεώρημα κυριαρχημένης σύγκλισης, θα χρειαζόταν να θεωρήσουμε την τυχαία μεταβλητή $\sup_n X_n$ και να δείξουμε ότι έχει πεπερασμένη μέση τιμή. Αλλά για το σκοπό αυτό θα χρειαζόμαστε τις από κοινού κατανομές των όρων της $\{X_n\}$ που μπορεί να μην είναι διαθέσιμες.

Άμεση συνέπεια των θεωρημάτων μονότονης και κυριαρχημένης σύγκλισης είναι το θεώρημα Beppo Levi το οποίο μας δίνει συνθήκες ώστε να εναλλάσουμε τον τελεστή της μέσης τιμής και του άπειρου αθροίσματος.

Πρόταση 15. *Θεώρημα Beppo Levi:*

(1) *Αν η $(X_n)_{n\geq 0}$ είναι ακολουθία μη-αρνητικών τυχαίων μεταβλητών, τότε*

$$E[\sum_{n=0}^{\infty} X_n] = \sum_{n=0}^{\infty} E[X_n].$$

(2) *Αν η $(X_n)_{n\geq 0}$ είναι ακολουθία τυχαίων μεταβλητών με $\sum_{n=0}^{\infty} E[|X_n|] < \infty$ τότε η σειρά $\sum_{n=0}^{\infty} X_n$ συγκλίνει σε τυχαία μεταβλητή με τιμές στο $\mathbb{R}$ με πιθανότητα 1 και επιπλέον*

$$E[\sum_{n=0}^{\infty} X_n] = \sum_{n=0}^{\infty} E[X_n].$$

2.5. **Άλλα είδη σύγκλισης τυχαίων μεταβλητών και μεταξύ τους σχέσεις.** Όπως είδαμε, το ανάλογο της σύγκλισης κατά σημείο ακολουθιών συναρτήσεων στη Θεωρία Πιθανοτήτων είναι η σύγκλιση με πιθανότητα 1 ακολουθίας τυχαίων μεταβλητών. Δυό άλλα είδη σύγκλισης ακολουθιών συναρτήσεων είναι η σύγκλιση κατά νόρμα και η σύγκλιση κατά μέτρο. Τα αντίστοιχά τους στη Θεωρία Πιθανοτήτων είναι η σύγκλιση κατά p-ροπή και η σύγκλιση κατά πιθανότητα. Αρχίζουμε υπενθυμίζοντας την έννοια της νόρμας και της σύγκλισης κατά νόρμα.

Ορισμός 10. *Έστω V ένας διανυσματικός χώρος συναρτήσεων με τιμές στο $\mathbb{R}$. Θεωρούμε μια συνάρτηση $\|\cdot\|: V \to \mathbb{R}$ με*

(1) *$\| f \| \geq 0$, για κάθε $f \in V$,*
(2) *$\| f \| = 0$, αν και μόνο αν $f = 0$,*
(3) *$\| af \| = |a| \, \| f \|$, για κάθε $f \in V$ και $a \in \mathbb{R}$,*
(4) *$\| f + g \| \leq \| f \| + \| g \|$, για κάθε $f, g \in V$.*

Τότε η $\|\cdot\|$ λέγεται νόρμα στον V.

Έστω $(f_n)_{n\geq 0}$ ακολουθία συναρτήσεων στον V και f συνάρτηση του V. Η $(f_n)_{n\geq 0}$ συγκλίνει στην f ως προς τη νόρμα $\| \cdot \|$ (συμβολικά $f_n \stackrel{\|\cdot\|}{\to} f$) αν $\lim_{n\to\infty} \| f_n - f \| = 0$.

Στη Θεωρία Πιθανοτήτων, δοθέντος ενός χώρου πιθανότητας $(\Omega, \mathcal{F}, P)$, μπορούμε για κάθε $p \geq 1$ να ορίσουμε το σημαντικό (όπως θα δούμε παρακάτω) διανυσματικό χώρο

$$\mathcal{L}^p(\Omega, \mathcal{F}, P) = \{X \text{ τυχαία μεταβλητή με } E[|X|^p] < \infty\},$$

που περιλαμβάνει όλες τις τυχαίες μεταβλητές επί του Ω με πεπερασμένη ροπή p-τάξης (ο όρος ροπή p-τάξης αναφέρεται στην $E[X^p]$).

Στον $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ ορίζουμε τη συνάρτηση $\| \cdot \|_p$ από τη σχέση

$$\| X \|_p = E[|X|^p]^{1/p}.$$

Είναι εύκολο να δούμε ότι η συνάρτηση $\| \cdot \|_p$ ικανοποιεί τις ιδιότητες (1), (3) και (4) του ορισμού της νόρμας αλλά όχι και τη (2), αφού $\| X \|_p = 0$ δεν συνεπάγεται ότι $X = 0$, αλλά μόνο ότι $P(X = 0) = 1$. Για το λόγο αυτό ορίζουμε στον $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ τη σχέση ισοδυναμίας σύμφωνα με την οποία δυο τυχαίες μεταβλητές θεωρούνται ισοδύναμες αν είναι ίσες με πιθανότητα 1. Στο χώρο πηλίκο που προκύπτει και τον οποίο συμβολίζουμε με $L^p(\Omega, \mathcal{F}, P)$, η $\| \cdot \|_p$ είναι νόρμα. Επιπλέον αποδεικνύεται ότι ο νορμαρισμένος χώρος $(L^p(\Omega, \mathcal{F}, P), \| \cdot \|_p)$ είναι πλήρης (δηλαδή είναι χώρος Banach).

Μια άλλη σύγκλιση που ορίζεται για ακολουθίες πραγματικών συναρτήσεων είναι η σύγκλιση κατά μέτρο. Στη Θεωρία Πιθανοτήτων το αντίστοιχό της είναι η σύγκλιση κατά πιθανότητα. Επιπλέον στη Θεωρία Πιθανοτήτων υπάρχει και η σύγκλιση κατά κατανομή ή νόμο που όπως θα δούμε αναφέρεται σε μια ακόμα πιο ασθενή σύγκλιση που διατυπώνεται με βάση τις κατανομές των αντίστοιχων τυχαίων μεταβλητών, χωρίς να χρειάζεται αυτές να είναι ορισμένες στον ίδιο χώρο πιθανότητας. Για να διασαφηνιστούν οι έννοιες των διαφορετικών συγκλίσεων, παρουσιάζουμε και τα τέσσερα είδη συγκλίσεων τυχαίων μεταβλητών (τα τρία καινούργια και τη σύγκλιση με πιθανότητα 1 που έχουμε ήδη δει) στον ίδιο ορισμό.

Ορισμός 11. *Έστω $X, X_0, X_1, X_2, \ldots$ τυχαίες μεταβλητές στο χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$. Τότε:*

(1) *Η $(X_n)_{n\geq 0}$ συγκλίνει στη X με πιθανότητα 1 ή σχεδόν βέβαια ή σχεδόν παντού (almost surely - a.s.), αν*

$$P(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1.$$

Συμβολικά γράφουμε $X_n \stackrel{a.s.}{\to} X$.

(2) *Η $(X_n)_{n\geq 0}$ συγκλίνει στη X στον L^p ή κατά p-ροπή, αν*

$$E[|X_n^p|] < \infty, n = 0, 1, 2, \dots \text{ και } \lim_{n\to\infty} E[|X_n - X|^p] = 0.$$

Συμβολικά γράφουμε $X_n \xrightarrow{L^p} X$.

(3) *Η $(X_n)_{n\geq 0}$ συγκλίνει στη X κατά πιθανότητα, αν για κάθε $\epsilon > 0$ ισχύει ότι*

$$\lim_{n\to\infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}) = 0.$$

Συμβολικά γράφουμε $X_n \xrightarrow{p} X$.

(4) *Η $(X_n)_{n\geq 0}$ συγκλίνει στη X κατά κατανομή ή νόμο, αν για κάθε σημείο συνέχειας της συνάρτησης κατανομής F_X της X ισχύει*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

Συμβολικά γράφουμε $X_n \xrightarrow{d} X$.

Για $p = 1, 2$ η σύγκλιση στον L^p αναφέρεται αντίστοιχα ως σύγκλιση κατά μέσο (in mean) και ως σύγκλιση κατά μέσο τετραγωνικό σφάλμα (in mean square). Συμβολικά γράφουμε $X_n \xrightarrow{i.m.} X$ και $X_n \xrightarrow{m.s} X$. Είναι σημαντικό να συνειδητοποιήσει κανείς ότι μόνο οι τρεις πρώτες συγκλίσεις αναφέρονται πραγματικά στις τυχαίες μεταβλητές ως πραγματικές συναρτήσεις επί του δειγματικού χώρου. Η κατά κατανομή σύγκλιση αναφέρεται στην πραγματικότητα στις συναρτήσεις κατανομών και όχι στις ίδιες τις τυχαίες μεταβλητές. Με την έννοια αυτή δεν λαμβάνει υπόψη το δειγματικό χώρο Ω και συνεπώς ούτε την εξάρτηση των τυχαίων μεταβλητών. Για να γίνει αυτό σαφές, δίνουμε το παρακάτω παράδειγμα.

Παράδειγμα 1. *Έστω X τυχαία μεταβλητή Bernoulli με $P(X = 0) = P(X = 1) = \frac{1}{2}$. Έστω $X_n = X$, $n = 0, 1, \dots$. Τότε είναι προφανές ότι $X_n \xrightarrow{d} X$. Από την άλλη μεριά αν θέσουμε $Y = 1 - X$ τότε έχουμε $X_n \xrightarrow{d} Y$, αφού η Y έχει ακριβώς την ίδια κατανομή με την X, αλλά ισχύει το παράδοξο να έχουμε $|X_n - Y| = 1$ για κάθε $n = 0, 1, \dots$ που αποκλείει οποιαδήποτε από τις άλλες συγκλίσεις (με πιθανότητα 1, σε κάποιον L^p ή κατά πιθανότητα).*

Στο υπόλοιπο της παραγράφου αυτής θα διερευνήσουμε τη σχέση των διαφόρων τύπων συγκλίσεων που μόλις ορίστηκαν. Θα δούμε δηλαδή ποιές συγκλίσεις συνεπάγονται άλλες συγκλίσεις. Πολύ πρόχειρα, θα μπορούσαμε να πούμε ότι οι πιο ισχυρές έννοιες σύγκλισης είναι η σύγκλιση με πιθανότητα 1 και η σύγκλιση σε κάποιον L^p και ακολουθεί η σύγκλιση κατά πιθανότητα και κατόπιν η σύγκλιση κατά κατανομή. Επίσης η σύγκλιση σε κάποιον L^p είναι πιο ισχυρή καθώς το p αυξάνει. Τα αντίστοιχα αποτελέσματα θα διατυπωθούν καθαρά παρακάτω. Πριν όμως από τη διατύπωσή τους είναι σημαντικό να δούμε δυο βασικά εργαλεία για την απόδειξή τους. Το πρώτο εργαλείο είναι η έννοια της Cauchy ακολουθίας τυχαίων μεταβλητών και το δεύτερο τα λήμματα Borel-Cantelli.

Υπενθυμίζουμε από τη μαθηματική ανάλυση, ότι μια ακολουθία πραγματικών αριθμών $(x_n)_{n\geq 0}$ λέγεται Cauchy ή βασική αν για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε για κάθε $n, m \geq n_0$ να ισχύει $|x_n - x_m| < \epsilon$. Η έννοια αυτή μεταφέρεται και στις ακολουθίες τυχαίων μεταβλητών, δηλαδή μια ακολουθία τυχαίων μεταβλητών $(X_n)_{n\geq 0}$ λέγεται Cauchy ή βασική σχεδόν παντού, αν $P(\{\omega \in \Omega : (X_n(\omega))_{n\geq 0} \text{ είναι Cauchy}\}) = 1$. Μια ακολουθία πραγματικών αριθμών είναι συγκλίνουσα αν και μόνο αν είναι Cauchy. Το αποτέλεσμα αυτό έχει το στοχαστικό του αντίστοιχο σύμφωνα με το οποίο, μια ακολουθία τυχαίων μεταβλητών συγκλίνει με πιθανότητα 1 αν και μόνο αν είναι Cauchy σχεδόν παντού. Αυτή η ισοδυναμία είναι αρκετά χρήσιμη σε διάφορες αποδείξεις.

Τα λήμματα Borel-Cantelli συνιστούν ένα ισχυρό εργαλείο για να αποδείξει κανείς ότι κάποιο ενδεχόμενο έχει πιθανότητα 1. Για το λόγο αυτό χρησιμοποιούνται συχνά γα να αποδειχθεί ότι μια ακολουθία τυχαίων μεταβλητών συγκλίνει με πιθανότητα 1.

Πρόταση 16. *1ο Λήμμα Borel-Cantelli: Αν $(A_n)_{n\geq 0}$ είναι μια ακολουθία ενδεχομένων με $\sum_{n=0}^{\infty} P(A_n) < \infty$, τότε $P(\sum_{n=0}^{\infty} 1_{A_n} < \infty) = 1$.*

2ο Λήμμα Borel-Cantelli: Αν $(A_n)_{n\geq 0}$ είναι μια ακολουθία ενδεχομένων με $\sum_{n=0}^{\infty} P(A_n) = \infty$ και επιπλέον τα $(A_n)_{n\geq 0}$ είναι ανεξάρτητα, τότε $P(\sum_{n=0}^{\infty} 1_{A_n} = \infty) = 1$.

Διαισθητικά, το 1ο λήμμα Borel-Cantelli λέει ότι αν οι πιθανότητες να συμβούν μια σειρά από ενδεχόμενα είναι αρκετά μικρές (τόσο μικρές ώστε η αντίστοιχη σειρά να συγκλίνει), τότε σχεδόν όλα τα σημεία του δειγματικού χώρου ανήκουν σε πεπερασμένα το πλήθος από αυτά τα ενδεχόμενα. Αντίστοιχα, το 2ο λήμμα Borel-Cantelli λέει ότι αν οι πιθανότητες να συμβούν μια σειρά από ενδεχόμενα είναι αρκετά μεγάλες (τόσο μεγάλες ώστε η αντίστοιχη σειρά να αποκλίνει), τότε σχεδόν όλα τα σημεία του δειγματικού χώρου ανήκουν σε άπειρα το πλήθος από αυτά τα ενδεχόμενα. Όμως, το 2ο λήμμα Borel-Cantelli χρειάζεται επιπλέον την υπόθεση της ανεξαρτησίας των ενδεχομένων για να δουλέψει.

Με βάση αυτά μπορεί να αποδειχθεί ο επόμενος χαρακτηρισμός της σύγκλισης με πιθανότητα 1.

Πρόταση 17. *Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών και X τυχαία μεταβλητή. Τα ακόλουθα είναι ισοδύναμα:*

(1) $X_n \overset{a.s.}{\to} X$.

(2) *Για κάθε $\epsilon > 0$ ισχύει $\lim_{m\to\infty} P(|X_n - X| > \epsilon, \text{ για κάποιο } n \geq m) = 0$.*

Επιπλέον, η επόμενη πρόταση δίνει μια ικανή συνθήκη για τη σύγκλιση με πιθανότητα 1.

Πρόταση 18. *Έστω* $(X_n)_{n\geq 0}$ *ακολουθία τυχαίων μεταβλητών και* X *τυχαία μεταβλητή ώστε για κάθε* $\epsilon > 0$ *να ισχύει* $\sum_{n=0}^{\infty} P(|X_n - X| > \epsilon) < \infty$. *Τότε* $X_n \xrightarrow{a.s.} X$.

Είμαστε τώρα σε θέση να διατυπώσουμε την παρακάτω πρόταση που δείχνει τη σχέση ισχύος μεταξύ των διαφόρων τύπων συγκλίσεων τυχαίων μεταβλητών.

Πρόταση 19. *Έστω* $(X_n)_{n\geq 0}$ *ακολουθία τυχαίων μεταβλητών και* X *τυχαία μεταβλητή. Τότε*

(1) $X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$.

(2) $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X$.

(3) *Για κάθε* $p \geq 1$, $X_n \xrightarrow{L^p} X \Rightarrow X_n \xrightarrow{p} X$.

(4) *Για κάθε* $p > q \geq 1$, $X_n \xrightarrow{L^p} X \Rightarrow X_n \xrightarrow{L^q} X$.

Γενικά δεν ισχύουν οι αντίστροφες συνεπαγωγές. Το γεγονός αυτό πιστοποιείται από μια σειρά αντιπαραδειγμάτων.

Παράδειγμα 2. *Έστω* X *τυχαία μεταβλητή Bernoulli με* $P(X=0) = P(X=1) = \frac{1}{2}$. *Έστω* $X_n = X$, $n = 0, 1, \ldots$ *και* $Y = 1 - X$. *Τότε έχουμε* $F_{X_n}(x) = F_Y(x)$, *οπότε και* $X_n \xrightarrow{d} Y$, *αφού η* Y *έχει ακριβώς την ίδια κατανομή με την* X, *αλλά ισχύει* $|X_n - Y| = 1$ *για κάθε* $n = 0, 1, \ldots$. *Επομένως* $X_n \xrightarrow{d} Y$ *αλλά* $X_n \not\xrightarrow{p} Y$.

Παράδειγμα 3. *Έστω* X_n *ανεξάρτητες με*

$$X_n = \begin{cases} 1 & \text{με πιθανότητα } \frac{1}{n} \\ 0 & \text{με πιθανότητα } 1 - \frac{1}{n}. \end{cases}$$

Τότε έχουμε $X_n \xrightarrow{p} 0$ *αλλά* $X_n \not\xrightarrow{a.s.} 0$.

Παράδειγμα 4. *Έστω* X_n *ανεξάρτητες με*

$$X_n = \begin{cases} n^3 & \text{με πιθανότητα } \frac{1}{n^2} \\ 0 & \text{με πιθανότητα } 1 - \frac{1}{n^2}. \end{cases}$$

Τότε έχουμε $X_n \xrightarrow{p} 0$ *αλλά* $X_n \not\xrightarrow{L^p} 0$, *για κάθε* $p \geq 1$.

Παράδειγμα 5. *Έστω* $p > q \geq 1$. *Ορίζουμε* X_n *ανεξάρτητες με*

$$X_n = \begin{cases} n & \text{με πιθανότητα } n^{-\frac{1}{2}(p+q)} \\ 0 & \text{με πιθανότητα } 1 - n^{-\frac{1}{2}(p+q)}. \end{cases}$$

Τότε έχουμε $X_n \xrightarrow{L^q} 0$ *αλλά* $X_n \not\xrightarrow{L^p} 0$.

Κάτω από κάποιες επιπλέον συνθήκες μπορούν να αποδειχθούν κάποια μερικά αντίστροφα των σχέσεων που είδαμε για τις διάφορες συγκλίσεις. Συγκεκριμένα, έχουμε την ακόλουθη πρόταση.

Πρόταση 20. *Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών και X τυχαία μεταβλητή. Τότε*

(1) *Αν $X_n \xrightarrow{d} c$, όπου c σταθερα, τότε $X_n \xrightarrow{p} c$.*

(2) *Αν $X_n \xrightarrow{p} X$ και υπάρχει σταθερά k τέτοια ώστε $P(|X_n| \leq k) = 1$ για κάθε $n = 0, 1, 2, \ldots$, τότε $X_n \xrightarrow{L^p} X$, για κάθε $p \geq 1$.*

Επιπλέον, θα πρέπει να διευκρινήσουμε ότι η σύγκλιση με πιθανότητα 1 και η σύγκλιση στον L^1 δεν συγκρίνονται. Δηλαδή, μπορεί να έχουμε όλες τις δυνατές περιπτώσεις. Αυτό γίνεται φανερό με τα παρακάτω δυο παραδείγματα.

Παράδειγμα 6. *Έστω X_n ανεξάρτητες με*

$$X_n = \begin{cases} n^3 & \text{με πιθανότητα } \frac{1}{n^2} \\ 0 & \text{με πιθανότητα } 1 - \frac{1}{n^2}. \end{cases}$$

Τότε έχουμε $X_n \xrightarrow{a.s.} 0$ αλλά $X_n \not\xrightarrow{L^p} 0$, για κάθε $p \geq 1$. Πρόκειται για την ίδια ακολουθία που δείχνει ότι $X_n \xrightarrow{p} 0 \not\Rightarrow X_n \xrightarrow{L^p} 0$.

Παράδειγμα 7. *Έστω X_n ανεξάρτητες με*

$$X_n = \begin{cases} 1 & \text{με πιθανότητα } \frac{1}{n} \\ 0 & \text{με πιθανότητα } 1 - \frac{1}{n}. \end{cases}$$

Τότε έχουμε $X_n \xrightarrow{L^1} 0$ αλλά $X_n \not\xrightarrow{a.s.} 0$. Πρόκειται για την ίδια ακολουθία που δείχνει ότι $X_n \xrightarrow{p} 0 \not\Rightarrow X_n \xrightarrow{a.s.} 0$.

Τέλος, είναι πολλές φορές βολικό να δουλεύουμε με τη σύγκλιση με πιθανότητα 1 παρόλο που οι υποθέσεις που έχουμε μας εξασφαλίζουν κάποια ασθενή σύγκλιση, όπως τη σύγκλιση κατά πιθανότητα ή κατά κατανομή. Οι παρακάτω δυο προτάσεις αποτελούν χρήσιμα εργαλεία που μας επιτρέπουν να 'μεταφράζουμε' τις ασθενείς συγκλίσεις σε κάποιας μορφής σύγκλιση με πιθανότητα 1.

Πρόταση 21. *Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών με $X_n \xrightarrow{p} X$. Τότε υπάρχει υπακολουθία $(X_{n_i})_{i\geq 0}$ τέτοια ώστε $X_{n_i} \xrightarrow{a.s.} X$.*

Πρόταση 22. *Θεώρημα αναπαράστασης Skorokhod: Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών με $X_n \xrightarrow{d} X$. Τότε υπάρχει ακολουθία τυχαίων μεταβλητών $(Y_n)_{n\geq 0}$ και τυχαία μεταβλητή Y ορισμένες στον ίδιο χώρο πιθανότητας ώστε για κάθε n οι X_n και Y_n να έχουν την ίδια κατανομή και οι X, Y να έχουν την ίδια κατανομή και $Y_n \xrightarrow{a.s.} Y$.*

2.6. **Βασικά οριακά αποτελέσματα στη Θεωρία Πιθανοτήτων: Ο νόμος των μεγάλων αριθμών, το κεντρικό οριακό θεώρημα, ο νόμος του διπλολογαρίθμου.** Στη Θεωρία Πιθανοτήτων, τα αθροίσματα ανεξάρτητων και ισόνομων τυχαίων μεταβλητών έχουν μελετηθεί εκτενώς, λόγω των πολλών εφαρμογών που έχουν στη στατιστική. Αν και δεν χρησιμοποιούνται στα παρακάτω, τα αναφέρουμε εδώ για την πληρότητα της παρούσας επισκόπησης.

Πρόταση 23. *Ισχυρός νόμος των μεγάλων αριθμών: Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $E[X_n] = 0$, $n \geq 1$, και έστω η ακολουθία των μερικών αθροισμάτων της, $(S_n)_{n\geq 1}$, με $S_n = \sum_{i=1}^n X_i$. Τότε*

$$\frac{1}{n} S_n \stackrel{a.s.}{\to} 0.$$

Πρόταση 24. *Κεντρικό οριακό θεώρημα: Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $E[X_n] = 0$, $Var[X_n] = 1$, $n \geq 1$, και έστω η ακολουθία των μερικών αθροισμάτων της, $(S_n)_{n\geq 1}$, με $S_n = \sum_{i=1}^n X_i$. Τότε*

$$\frac{1}{\sqrt{n}} S_n \stackrel{d}{\to} X,$$

όπου η X ακολουθεί την τυποποιημένη κανονική κατανομή με συνάρτηση κατανομής

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

Πρόταση 25. *Νόμος του διπλολογαρίθμου: Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $E[X_n] = 0$, $Var[X_n] = 1$, $n \geq 1$, και έστω η ακολουθία των μερικών αθροισμάτων της, $(S_n)_{n\geq 1}$, με $S_n = \sum_{i=1}^n X_i$. Τότε*

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2n \log\log n}} S_n = 1, \ a.s.$$

και

$$\liminf_{n\to\infty} \frac{1}{\sqrt{2n \log\log n}} S_n = -1, \ a.s.$$

Συνδυάζοντας τα τρία αυτά θεωρήματα αποκτά κανείς μια καλή διαίσθηση για το πώς διακυμαίνονται τα μερικά αθροίσματα μιας ακολουθίας ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Μεταξύ των αξιοσημείωτων γεγονότων, είναι ενδιαφέρον να παρατηρήσει ότι η μέση τιμή και η διασπορά των αντίστοιχων τυχαίων μεταβλητών αρκούν για να καθορίσουν την οριακή συμπεριφορά αυτών των μερικών αθροισμάτων.

3. Δεσμευμενη μεση τιμη

Από τη στοιχειώδη Θεωρία Πιθανοτήτων γνωρίζουμε μια έκδοση της έννοιας της δεσμευμένης μέσης τιμής μιας τυχαίας μεταβλητής, δοθείσης μιας άλλης τυχαίας μεταβλητής. Συγκεκριμένα, αν η (X, Y) είναι μια διακριτή ή απόλυτα συνεχής διδιάστατη τυχαία μεταβλητή με συνάρτηση πιθανότητας ή συνάρτηση πυκνότητας πιθανότητας $f_{X,Y}(x, y)$ ορίζουμε της αντίστοιχη δεσμευμένη συνάρτηση πιθανότητας ή συνάρτηση πυκνότητας πιθανότητας της X δοθείσης της Y που δίνεται από τον τύπο $f_{X|Y}(x, y) = f_{X,Y}(x, y)/f_Y(y)$, όπου $f_Y(y)$ η αντίστοιχη περιθώρια (φυσικά η $f_{X|Y}(x, y)$ ορίζεται μονοσήμαντα μόνο για y με $f_Y(y) \neq 0$). Και ακολούθως ορίζεται η δεσμευμένη μέση τιμή της X δοθέντος ότι $Y = y$ που δίνεται από τη σχέση

$$E[X|Y = y] = \begin{cases} \sum_x x f_{X|Y}(x, y) & \text{αν η } X \text{ διακριτή,} \\ \int_{-\infty}^{\infty} x f_{X|Y}(x, y) dx & \text{αν η } X \text{ συνεχής,} \end{cases}$$

για y με $f_Y(y) \neq 0$, ενώ ορίζεται αυθαίρετα για y με $f_Y(y) = 0$. Κατόπιν η τυχαία μεταβλητή $g(Y)$ που είναι συνάρτηση της Y με $g(y) = E[X|Y = y]$, ορίζεται να είναι η μέση τιμή της X δοθείσης της Y και συμβολίζεται με $E[X|Y]$. Αποδεικνύεται τότε ότι η τυχαία μεταβλητή $E[X|Y]$ είναι η 'καλύτερη' συνάρτηση της Y που προσεγγίζει τη X, με την έννοια ότι ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα $E[(X - Z)^2]$ μέσα στην κλάση των τυχαίων μεταβλητών Z που είναι $\sigma(Y)$-μετρησιμες (και άρα της μορφής $Z = g(Y)$ για κάποια $g : \mathbb{R} \to \mathbb{R}$ $\mathcal{B}(\mathbb{R})$-μετρήσιμη). Ο ορισμός αυτός δείχνει επίσης ότι η $E[X|Y]$ δεν ορίζεται μονοσήμαντα παντού (αφού η $E[X|Y = y]$ ορίζεται αυθαίρετα για y με $f_Y(y) = 0$), αλλά είναι μοναδική σχεδόν παντού. Αυτό είναι κάτι που ισχύει και για τη γενικότερη έννοια της δεσμευμένης μέσης τιμής που θα δούμε παρακάτω. Αν και είναι λίγο ενοχλητικό, είναι αναπόφευκτο και στην πραγματικότητα δεν δημιουργεί τεχνικά προβλήματα, εφόσον δουλεύουμε με αριθμήσιμο πλήθος ενδεχομένων καθένα από τα οποία έχει πιθανότητα 1.

Θα ορίσουμε τώρα τη δεσμευμένη μέση τιμή $E[X|\mathcal{G}]$ μιας τυχαίας μεταβλητής X ως προς μια σ-υποάλγεβρα $\mathcal{G}$ της σ-άλγεβρας $\mathcal{F}$ ως προς την οποία είναι μετρήσιμη η X. Η ιδέα είναι ανάλογη με τα παραπάνω, δηλαδή με κάποια έννοια η $E[X|\mathcal{G}]$ είναι η 'καλύτερη' $\mathcal{G}$-μετρήσιμη τυχαία μεταβλητή που προσεγγίζει τη X. Στην πραγματικότητα λοιπόν ο τελεστής $E[\cdot|\mathcal{G}]$ μπορεί να θεωρηθεί ως ένας τελεστής προβολής από το χώρο των τυχαίων μεταβλητών ως προς τη σ-άλγεβρα $\mathcal{F}$ του χώρου πιθανότητας στο χώρο των τυχαίων μεταβλητών που είναι $\mathcal{G}$-μετρήσιμες. Εδώ η έννοια της 'καλύτερης ' $\mathcal{G}$-μετρήσιμης τυχαίας μεταβλητής που προσεγγίζει τη X αποτυπώνεται στον παρακάτω ορισμό.

Ορισμός 12. *Έστω χώρος πιθανότητας $(\Omega, \mathcal{F}, P)$, X τυχαία μεταβλητή με $E[|X|] < \infty$ και $\mathcal{G}$ μια σ-υποάλγεβρα της $\mathcal{F}$. Θα λέμε ότι η Y είναι μια δεσμευμένη μέση τιμή της X ως προς την $\mathcal{G}$, αν*

(1) *η Y είναι $\mathcal{G}$-μετρήσιμη.*

(2) *Για κάθε $A \in \mathcal{G}$, έχουμε $E[X 1_A] = E[Y 1_A]$.*

Στην περίπτωση αυτή θα γράφουμε την Y και ως $E[X|\mathcal{G}]$.

Η έννοια του ορισμού αυτού είναι ότι η $E[X|\mathcal{G}]$ είναι η 'καλύτερη' $\mathcal{G}$-μετρήσιμη τυχαία μεταβλητή που προσεγγίζει την X αν ενδιαφερόμαστε για τον υπολογισμό της μέσης τιμής της X πάνω σε $\mathcal{G}$-μετρήσιμα σύνολα (δηλαδή σε σύνολα που ανήκουν στην $\mathcal{G}$). Πραγματικά ο ορισμός απαιτεί η μέση τιμή της X πάνω σε κάθε τέτοιο σύνολο να δίνεται από την αντίστοιχη μέση τιμή της $E[X|\mathcal{G}]$. Αλλά πού είναι το όφελος;

Το όφελος βρίσκεται στο ότι η $E[X|\mathcal{G}]$ είναι απλούστερο αντικείμενο από την X, ακριβώς επειδή είναι μετρήσιμη ως προς τη μικρότερη σ-άλγεβρα $\mathcal{G}$ και επομένως οι μέσες τιμές της $E[X|\mathcal{G}]$ υπολογίζονται ευκολότερα.

Μια άλλη διαισθητική προσέγγιση λέει ότι η δεσμευμένη μέση τιμή $E[X|\mathcal{G}]$ είναι η καλύτερη εκτίμηση της X δεδομένης της πληροφορίας που φέρει η σ-άλγεβρα $\mathcal{G}$. Το να φτιάξουμε την $E[X|\mathcal{G}]$ σημαίνει να δημιουργήσουμε μια προσέγγιση της X, αποδίδωντας τιμές σε κάθε $\omega \in \Omega$ για το $E[X|\mathcal{G}](\omega)$, αλλά στην κατασκευή αυτή μπορούμε να χρησιμοποιήσουμε μόνο τα σύνολα της $\mathcal{G}$ για να ξεχωρίσουμε τα στοιχεία του Ω.

Έτσι, αν $\mathcal{G} = \{\emptyset, \Omega\}$ είναι η τετριμμένη σ-άλγεβρα, τότε $E[X|\mathcal{G}] = E[X]$. Πραγματικά, στην περίπτωση αυτή η δομή της σ-άλγεβρας $\mathcal{G}$ είναι πολύ φτωχή και δεν μας επιτρέπει να ξεχωρίσουμε καθόλου τα στοιχεία του Ω. Άρα αναγκαστικά για κάθε $\omega \in \Omega$ θα πρέπει να δώσουμε την ίδια τιμή στο $E[X|\mathcal{G}](\omega)$. Πιο αυστηρά οι μόνες $\mathcal{G}$-μετρήσιμες τυχαίες μεταβλητές είναι οι σταθερές. Αν εφαρμόσουμε τον ορισμό βλέπουμε ότι αναγκαστικά η σταθερά πρέπει να είναι η $E[X]$.

Στο άλλο άκρο, αν $\mathcal{G} = \mathcal{F}$, τότε $E[X|\mathcal{G}] = X$. Πραγματικά, στην περίπτωση αυτή η ίδια η X είναι $\mathcal{G}$-μετρήσιμη και ικανοποιεί τον ορισμό.

Ως τελευταίο παράδειγμα, ας θεωρήσουμε την περίπτωση που η $\mathcal{G}$ είναι η παραγόμενη σ-άλγεβρα από μια αριθμήσιμη ακολουθία ενδεχομένων $(A_n)_{n \geq 0}$ θετικής πιθανότητας. Τότε, οι $\mathcal{G}$-μετρήσιμες συναρτήσεις είναι οι συναρτήσεις της μορφής $\sum_{n=0}^{\infty} c_n 1_{A_n}$. Συνεπώς, στην περίπτωση αυτή η δομή της σ-άλγεβρας $\mathcal{G}$ είναι τέτοια ώστε η πληροφορία που φέρει είναι σε ποιό A_n ανήκει κάθε ω. Τα σημεία ω που ανήκουν, όμως, στο ίδιο A_n δεν ξεχωρίζονται και στα σημεία αυτά θα πρέπει να αποδοθεί η ίδια τιμή στα $E[X|\mathcal{G}](\omega)$. Επομένως έχουμε ότι $E[X|\mathcal{G}] = \sum_{n=0}^{\infty} c_n 1_{A_n}$. Εφαρμόζοντας τον ορισμό, επιλέγοντας κάθε φορά για A ένα σύνολο από τα A_n, παίρνουμε $E[X 1_{A_n}] = c_n P(A_n)$ και τελικά

$$E[X|\mathcal{G}] = \sum_{n=0}^{\infty} \frac{E[X 1_{A_n}]}{P(A_n)} 1_{A_n}.$$

Ο ορισμός της μέσης τιμής τυχαίας μεταβλητής ως προς σ-άλγεβρα γενικεύει το στοιχειώδη ορισμό της δεσμευμένης μέσης τιμής. Πραγματικά αν δίνονται κάποιες τυχαίες μεταβλητές $Y_1, Y_2, \ldots, Y_n$, τότε η $E[X|\sigma(Y_1, Y_2, \ldots, Y_n)]$ ταυτίζεται με την τυχαία μεταβλητή $E[X|Y_1, Y_2, \ldots, Y_n]$ που προκύπτει από το στοιχειώδη ορισμό που δώσαμε στην αρχή της παραγράφου (όπου η Y είναι τώρα μια n-διάστατη τυχαία μεταβλητή).

Αν $P(Y = Y') = 1$ και η Y είναι μια δεσμευμένη μέση τιμή της X ως προς μια σ-άλγεβρα $\mathcal{G}$ και η Y' είναι $\mathcal{G}$-μετρήσιμη, τότε και η Y' είναι μια δεσμευμένη μέση τιμή της X ως προς τη σ-άλγεβρα $\mathcal{G}$. Επομένως η δεσμευμένη μέση τιμή δεν είναι μοναδική, αλλά από την άλλη μεριά, οποιεσδήποτε δυο τυχαίες μεταβλητές που είναι μέσες τιμές της X ως προς μια σ-άλγεβρα $\mathcal{G}$ είναι ίσες σχεδόν παντού. Για το λόγο αυτό αναφερόμαστε σε πολλές εκδοχές της δεσμευμένης μέσης τιμής της X ως προς τη σ-άλγεβρα $\mathcal{G}$. Η ασάφεια αυτή δεν προκαλεί προβλήματα όταν έχουμε διάφορα επιχειρήματα, αρκεί να περιοριζόμαστε σε αριθμήσιμο πλήθος ιδιοτήτων που ισχύουν με πιθανότητα 1 (γιατί η τομή αριθμήσιμου πλήθους ενδεχομένων που το καθένα έχει πιθανότητα 1 έχει επίσης πιθανότητα 1).

Η πιο σύντομη απόδειξη της ύπαρξης της δεσμευμένης μέσης τιμής γίνεται με το θεώρημα Radon-Nikodym από τη Θεωρία Μέτρου. Εναλλακτικά μπορεί κανείς να θεμελιώσει σταδιακά τη δεσμευμένη μέση τιμή πρώτα για τυχαίες μεταβλητές στον $L^2(\Omega, \mathcal{F}, P)$ και κατόπιν να επεκτείνει τον ορισμό σε μη-αρνητικές τυχαίες μεταβλητές και τελικά σε τυχαίες μεταβλητές στον $L^1(\Omega, \mathcal{F}, P)$. Η διαδικασία αυτή γίνεται αξιοποιώντας αποτελέσματα από τη θεωρία προβολών σε χώρους Hilbert. Συγκεκριμένα, ο χώρος $L^2(\Omega, \mathcal{F}, P)$, εφοδιασμένος με το εσωτερικό γινόμενο $\langle X, Y\rangle = E[XY]$ είναι χώρος Hilbert. Ο $L^2(\Omega, \mathcal{G}, P)$ είναι κλειστός υπόχωρος του $L^2(\Omega, \mathcal{F}, P)$ για κάθε σ-υποάλγεβρα $\mathcal{G}$ της $\mathcal{F}$. Έχουμε τότε την ακόλουθη πρόταση.

Πρόταση 26. *Η συνάρτηση $T : L^2(\Omega, \mathcal{F}, P) \to L^2(\Omega, \mathcal{G}, P)$ με $T(X) = E[X|\mathcal{G}]$ είναι η ορθογώνια προβολή στον υπόχωρο $L^2(\Omega, \mathcal{G}, P)$.*

Το παρακάτω θεώρημα δίνει τις βασικές ιδιότητες της δεσμευμένης μέσης τιμής.

Πρόταση 27. *Έστω χώρος πιθανότητας $(\Omega, \mathcal{F}, P)$ και $\mathcal{G}, \mathcal{H}$ σ-υποάλγεβρες της $\mathcal{F}$. Όσες τυχαίες μεταβλητές εμφανίζονται παρακάτω θεωρούνται ότι έχουν πεπερασμένη μέση τιμή. Επίσης οι ισότητες που αφορούν δεσμευμένες μέσες τιμές θα πρέπει να εννοούνται με πιθανότητα 1. Ισχύουν τα ακόλουθα:*

(1) *Θεώρημα διπλής μέσης τιμής:* $E[E[X|\mathcal{G}]] = E[X]$.
(2) *Γραμμικότητα:* $E[a_1X_1 + a_2X_2|\mathcal{G}] = a_1E[X_1|\mathcal{G}] + a_2E[X_2|\mathcal{G}]$.
(3) *Μη-αρνητικότητα: Αν* $X \geq 0$, *τότε* $E[X|\mathcal{G}] \geq 0$.
(4) *Μονοτονία: Αν* $X \leq Y$ *με πιθανότητα 1, τότε* $E[X|\mathcal{G}] \leq E[Y|\mathcal{G}]$.

(5) *Ιδιότητα πύργου: Αν* $\mathcal{H} \subseteq \mathcal{G}$ *τότε* $E[E[X|\mathcal{G}]|\mathcal{H}] = E[E[X|\mathcal{H}]|\mathcal{G}] = E[X|\mathcal{H}]$.

(6) *Τα γνωστά βγαίνουν έξω: Αν η* Z *είναι* $\mathcal{G}$*-μετρήσιμη και* $E[|ZX|] < \infty$ *τότε* $E[ZX|\mathcal{G}] = ZEX|\mathcal{G}]$.

(7) *Αν η* X *είναι* $\mathcal{G}$*-μετρήσιμη, τότε* $E[X|\mathcal{G}] = X$.

(8) *Τα ανεξάρτητα παραλείπονται από τη δέσμευση: Αν η* $\mathcal{H}$ *είναι ανεξάρτητη της* $\sigma(\sigma(X), \mathcal{G})$ *τότε* $E[X|\sigma(\mathcal{G}, \mathcal{H})] = E[X|\mathcal{G}]$.

(9) *Αν η* X *είναι ανεξάρτητη της* $\mathcal{H}$*, τότε* $E[X|\mathcal{H}] = E[X]$.

Πέραν των παραπάνω ιδιοτήτων, τα θεωρήματα μονότονης και κυριαρχημένης συγκλισης, το λήμμα Fatou και η ανισότητα Jensen ισχύουν αν αντικαταστήσουμε τις αντίστοιχες μέσες τιμές με δεσμευμένες μέσες τιμές ως προς μια σ-άλγεβρα $\mathcal{G}$. Μια επιπλέον χρήσιμη ανισότητα (που προκύπτει άμεσα από την ανισότητα Jensen) είναι και η ακόλουθη.

Πρόταση 28. *Έστω χώρος πιθανότητας* $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$ *σ-υποάλγεβρα της* $\mathcal{F}$ *και* $X \in L^1(\Omega, \mathcal{F}, P)$. *Για κάθε* $p \geq 1$ *ισχύει*

$$\| E[X|\mathcal{G}] \|_p \leq \| X \|_p .$$

4. Martingales, submartingales, supermartingales: Ορισμοι, παραδειγματα και βασικες ιδιοτητες

Τα martingales, τα submartingales και τα supermartingales είναι στοχαστικές διαδικασίες που έχουν τη ρίζα τους στα στοιχηματικά παιχνίδια. Θα μπορούσαμε να πούμε ότι μοντελοποιούν την εξέλιξη της περιουσίας ενός παίκτη που παίζει ένα τυχερό παιχνίδι που είναι αντίστοιχα δίκαιο, συμφέρον ή ασύμφορο για τον παίκτη και ποντάρει μια χρηματική μονάδα σε κάθε γύρο του παιχνιδιού. Πραγματικά, για ένα δίκαιο (αντίστοιχα συμφέρον, ασύμφορο) για τον παίκτη παιχνίδι, απαιτούμε η μέση τιμή της περιουσίας του παίκτη σε κάποια μελλοντική χρονική στιγμή, δεδομένης της πληροφορίας που υπάρχει στο παρόν να είναι ίση με (αντίστοιχα μεγαλύτερη από, μικρότερη από) τη σημερινή αξία της περιουσίας του παίκτη. Η έννοια της πληροφορίας που αυξάνει με το χρόνο μοντελοποιείται στη Θεωρία Πιθανοτήτων με την έννοια της διήθησης σ-αλγεβρών. Πιο συγκεκριμένα έχουμε τον ακόλουθο ορισμό.

Ορισμός 13. *Έστω* $(\Omega, \mathcal{F}, P)$ *ένας χώρος πιθανότητας. Μια αύξουσα ακολουθία σ-αλγεβρών* $(\mathcal{F}_n)_{n\geq 0}$ *(που έχει δηλαδή την ιδιότητα* $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, $n \geq 0$*) με* $\mathcal{F}_n \subseteq \mathcal{F}$, $n \geq 0$, *λέγεται διήθηση (filtration) στον* $(\Omega, \mathcal{F}, P)$.

Η σ-άλγεβρα $\mathcal{F}_n$ μπορεί να θεωρηθεί ότι περιέχει τα ενδεχόμενα του Ω τα οποία μπορούμε να αποφανθούμε αν πραγματοποιήθηκαν ή όχι μέχρι κάποια στιγμή n και επομένως εκφράζει την συσσωρευμένη πληροφορία μέχρι τη στιγμή n.

Ορισμός 14. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$.*

(1) *Μια ακολουθία τυχαίων μεταβλητών $(X_n)_{n\geq 0}$ λέγεται προσαρμοσμένη (adapted) στη διήθηση αν για κάθε $n = 0, 1, \ldots$ η X_n είναι $\mathcal{F}_n$-μετρήσιμη.*
(2) *Μια ακολουθία τυχαίων μεταβλητών $(X_n)_{n\geq 1}$ λέγεται προβλέψιμη (non-anticipating) ως προς τη διήθηση αν για κάθε $n = 1, 2 \ldots$ η X_n είναι $\mathcal{F}_{n-1}$-μετρήσιμη.*

Χρησιμοποιώντας αυτές τις έννοιες μπορούμε τώρα να ορίσουμε τα martingales, submartingales και supermartingales.

Ορισμός 15. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$. Μια ακολουθία τυχαίων μεταβλητών $(M_n)_{n\geq 0}$ λέγεται $(\mathcal{F}_n)_{n\geq 0}$-martingale αν*

(1) *$E[|M_n|] < \infty$, για κάθε $n \geq 0$,*
(2) *η $(M_n)_{n\geq 0}$ είναι προσαρμοσμένη στην $(\mathcal{F}_n)_{n\geq 0}$,*
(3) *$E[M_n|\mathcal{F}_m] = M_m$ a.s. για κάθε $0 \leq m \leq n$.*

Η ακολουθία λέγεται submartingale αν αντί της (3) ισχύει ότι $E[M_n|\mathcal{F}_m] \geq M_m$ a.s. για κάθε $0 \leq m \leq n$. Τέλος, αν ισχύει ότι $E[M_n|\mathcal{F}_m] \leq M_m$ a.s. για κάθε $0 \leq m \leq n$, η ακολουθία λέγεται supermartingale.

Σε πολλά συγγράμματα αντί της ιδιότητας (3) του ορισμού απαιτείται η

(3′) $E[M_{m+1}|\mathcal{F}_m] = M_m$ *a.s.* για κάθε $m \geq 0$.

(αντίστοιχα $E[M_{m+1}|\mathcal{F}_m] \geq M_m$ για submartingales ή $E[M_{m+1}|\mathcal{F}_m] \leq M_m$ για supermartingales). Είναι άμεσο ότι η (3) συνεπάγεται την (3′) και επίσης μπορούμε εύκολα να αποδείξουμε την ισχύ της (3) ξεκινώντας από την (3′) και χρησιμοποιώντας επαγωγή. Συνεπώς οι (3) και (3′) είναι ισοδύναμες και οδηγούν στην ίδια έννοια martingale. Το πλεονέκτημα της (3) είναι ότι μπορεί να γενικευτεί και στην περίπτωση των martingales συνεχούς χρόνου, πράγμα που δεν ισχύει για την (3′) (στο συνεχή χρόνο δεν υπάρχει αμέσως επόμενη χρονική στιγμή). Για το λόγο αυτό είναι προτιμητέα για τον ορισμό. Έχουμε επίσης την εξής πρόταση.

Πρόταση 29. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$. Αν μια στοχαστική διαδικασία $(M_n)_{n\geq 0}$ είναι $(\mathcal{F}_n)_{n\geq 0}$-martingale (αντίστοιχα submartingale ή supermartingale), τότε είναι martingale (αντίστοιχα submartingale ή supermartingale) ως προς τη διήθηση $(\sigma(M_k : 1 \leq k \leq n))_{n\geq 0}$.*

Απόδειξη. Πραγματικά οι συνθήκες (1) και (2) του ορισμού συνεχίζουν να ισχύουν και ως προς τη διήθηση $(\sigma(M_k : 1 \leq k \leq n))_{n\geq 0}$. Σχετικά με την (3) έχουμε για $0 \leq m \leq n$:

$$\begin{aligned} E[M_n|\sigma(M_k : 1 \leq k \leq m)] &= E[E[M_n|\mathcal{F}_m]|\sigma(M_k : 1 \leq k \leq m)] \\ &= E[M_m|\sigma(M_k : 1 \leq k \leq m)] = M_m. \end{aligned}$$

Στην πρώτη ισότητα έχουμε χρησιμοποιήσει το γεγονός ότι $\sigma(M_k : 1 \leq k \leq m) \subseteq \mathcal{F}_m$, $m \geq 0$, (που ισχύει επειδή η $(M_n)_{n\geq 0}$ είναι προσαρμοσμένη στην $(\mathcal{F}_n)_{n\geq 0}$) και την ιδιότητα πύργου της δεσμευμένης μέσης τιμής. Στη δεύτερη έχουμε χρησιμοποιήσει την συνθήκη (3) του ορισμού που ισχύει ως προς τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$. Τέλος, στην τρίτη έχουμε χρησιμοποιήσει την ιδιότητα 'τα γνωστά βγαίνουν έξω'. Για την περίπτωση που η $(M_n)_{n\geq 0}$ είναι submartingale ή supermartingale, έχουμε τα ίδια επιχειρήματα. ∎

Μια άλλη οπτική είναι να βλέπουμε τα martingales ως ακολουθίες τυχαίων μεταβλητών που 'παραμένουν σταθερές δοθέντος του παρόντος ', ενώ τα submartingales (αντίστοιχα τα supermartingales) είναι ακολουθίες τυχαίων μεταβλητών που 'είναι αύξουσες (αντίστοιχα φθίνουσες) δοθέντος του παρόντος '.

Μερικές σχεδόν προφανείς ιδιότητες των martingales, submartingales και supermartingales δίνονται στην επόμενη πρόταση.

Πρόταση 30. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας, $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$ και $(M_n)_{n\geq 0}$ μια ακολουθία τυχαίων μεταβλητών.*

(1) *$(M_n)_{n\geq 0}$ είναι martingale αν και μόνο αν είναι submartingale και supermartingale.*
(2) *Αν $(M_n)_{n\geq 0}$ είναι submartingale (αντίστοιχα supermartingale) τότε $(-M_n)_{n\geq 0}$ είναι supermartingale (αντίστοιχα submartingale).*
(3) *Αν $(M_n)_{n\geq 0}$ είναι martingale (αντίστοιχα submartingale, supermartingale), τότε η ακολουθία $(E[M_n])_{n\geq 0}$ είναι σταθερή (αντίστοιχα αύξουσα, φθίνουσα).*

Απόδειξη. Οι (1) και (2) είναι άμεσες από τον ορισμό. Για την (3) αρκεί να πάρουμε μέση τιμή στην (3) του ορισμού και να χρησιμοποιήσουμε το θεώρημα διπλής μέσης τιμής. Πράγματι, θα έχουμε για $0 \leq m \leq n$:

$$E[M_n] = E[E[M_n|\mathcal{F}_m]] = E[M_m].$$

∎

Είμαστε τώρα σε θέση να δούμε αρκετά παραδείγματα martingales submartingales και supermartingales.

Παράδειγμα 8. *Ο συμμετρικός τυχαίος περίπατος: Έστω η ακολουθία των ανεξάρτητων τυχαίων μεταβλητών $(X_n)_{n\geq 1}$ με $E[X_n] = 0$, $n \geq 1$. Θεωρούμε την ακολουθία των μερικών* *αθροισμάτων τους,* $(S_n)_{n\geq 0}$, *που ορίζεται θέτοντας*

$$\begin{aligned} S_0 &= 0, \\ S_n &= \sum_{i=1}^{n} X_i,\ n\geq 1. \end{aligned}$$

Θεωρούμε επίσης τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ *που ορίζεται θέτοντας*

$$\begin{aligned} \mathcal{F}_0 &= \{\emptyset, \Omega\}, \\ \mathcal{F}_n &= \sigma(X_k : 1\leq k\leq n),\ n\geq 1. \end{aligned}$$

Τότε η $(S_n)_{n\geq 0}$ *είναι ένα* $(\mathcal{F}_n)_{n\geq 0}$*- martingale. Πράγματι, η* $(S_n)_{n\geq 0}$ *είναι προφανώς προσαρμοσμένη στην* $(\mathcal{F}_n)_{n\geq 0}$*. Επιπλέον,* $E[|S_n|]\leq nE[|X_1|]<\infty$, $n\geq 0$*. Τέλος, για κάθε* $n\geq 0$ *έχουμε*

$$E[S_{n+1}|\mathcal{F}_n] = E[X_{n+1}+S_n|\mathcal{F}_n] = E[X_{n+1}|\mathcal{F}_n]+E[S_n|\mathcal{F}_n] = E[X_{n+1}]+S_n = S_n.$$

Στην πρώτη ισότητα χρησιμοποιήσαμε τον ορισμό της $(S_n)_{n\geq 0}$*, στη δεύτερη ισότητα χρησιμοποιήσαμε τη γραμμικότητα της μέσης τιμής, στην τρίτη τις ιδιότητες 'τα ανεξάρτητα παραλείπονται από τη δέσμευση' και τα 'γνωστά βγαίνουν έξω' και στην τέταρτη την υπόθεση ότι* $E[X_n]=0$*.*

Ειδική περίπτωση αυτού του παραδείγματος, την οποία θα δούμε με μεγαλύτερη λεπτομέρεια είναι η περίπτωση του συμμετρικού απλού τυχαίου περίπατου, όπου $P(X_n=1)=P(X_n=-1)=1/2$, $n\geq 1$.

Παράδειγμα 9. *Το τετράγωνο του συμμετρικού τυχαίου περίπατου: Έστω η ακολουθία των ανεξάρτητων τυχαίων μεταβλητών* $(X_n)_{n\geq 1}$ *με* $E[X_n]=0$ *και* $Var[X_n]=\sigma^2$, $n\geq 1$*. Θεωρούμε την ακολουθία* $(M_n)_{n\geq 0}$ *που ορίζεται ως*

$$\begin{aligned} M_0 &= 0, \\ M_n &= S_n^2 - n\sigma^2,\ n\geq 1, \end{aligned}$$

όπου $(S_n)_{n\geq 0}$ *είναι η ακολουθία των μερικών αθροισμάτων της* (X_n) *(βλέπε παράδειγμα συμμετρικού τυχαίου περίπατου). Θεωρούμε και πάλι τη διήθηση* $(\mathcal{F}_n)_{n\geq 0}$ *που ορίζεται θέτοντας* $\mathcal{F}_0=\{\emptyset,\Omega\}$ *την τετριμένη σ-άλγεβρα και* $\mathcal{F}_n=\sigma(X_k : 1\leq k\leq n)$, $n\geq 1$*. Τότε η* $(M_n)_{n\geq 0}$ *είναι ένα* $(\mathcal{F}_n)_{n\geq 0}$*-martingale. Πράγματι, έχουμε άμεσα ότι η* $(M_n)_{n\geq 0}$ *είναι προσαρμοσμένη στην* $(\mathcal{F}_n)_{n\geq 0}$*. Επίσης,*

$$E[|M_n|]\leq E[S_n^2]+n\sigma^2\leq nE[X_1^2]+n\sigma^2<\infty,\ n\geq 0.$$

Τέλος, για κάθε $n \geq 0$ έχουμε

$$\begin{aligned} E[M_{n+1}|\mathcal{F}_n] &= E[S_{n+1}^2 - (n+1)\sigma^2|\mathcal{F}_n] = E[(X_{n+1}+S_n)^2 - (n+1)\sigma^2|\mathcal{F}_n] \\ &= E[X_{n+1}^2|\mathcal{F}_n] + E[2X_{n+1}S_n|\mathcal{F}_n] + E[S_n^2|\mathcal{F}_n] - E[(n+1)\sigma^2|\mathcal{F}_n] \\ &= E[X_{n+1}^2] + 2S_nE[X_{n+1}] + S_n^2 - (n+1)\sigma^2 = \sigma^2 + 0 + S_n^2 - (n+1)\sigma^2 \\ &= M_n. \end{aligned}$$

Στην πρώτη ισότητα χρησιμοποιήσαμε τον ορισμό της $(M_n)_{n\geq 0}$, στη δεύτερη τον ορισμό της $(S_n)_{n\geq 0}$, στην τρίτη τη γραμμικότητα της μέσης τιμής, στην τέταρτη τις ιδιότητες 'τα ανεξάρτητα παραλείπονται από τη δέσμευση' και τα 'γνωστά βγαίνουν έξω' και στην πέμπτη τις υποθέσεις $E[X_n] = 0$ και $Var[X_n] = \sigma^2$.

Παράδειγμα 10. *Το γινόμενο ανεξάρτητων τυχαίων μεταβλητών με μέση τιμή 1: Έστω η ακολουθία των ανεξάρτητων τυχαίων μεταβλητών $(X_n)_{n\geq 1}$ με $X_n \geq 0$ και $E[X_n] = 1$, $n \geq 1$. Θέτουμε*

$$\begin{aligned} R_0 &= 1, \\ R_n &= \prod_{i=1}^{n} X_i, \ n \geq 1. \end{aligned}$$

Θεωρούμε και πάλι τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ που ορίζεται θέτοντας $\mathcal{F}_0 = \{\emptyset, \Omega\}$ την τετριμένη σ-άλγεβρα και $\mathcal{F}_n = \sigma(X_k : 1 \leq k \leq n)$, $n \geq 1$. Τότε η $(R_n)_{n\geq 0}$ είναι ένα $(\mathcal{F}_n)_{n\geq 0}$- martingale. Πραγματικά η ικανοποίηση των συνθηκών του ορισμού μπορεί εύκολα να ελεγχθεί όπως και στα προηγούμενα παραδείγματα.

Αυτό το martingale έχει πολλές χρήσιμες εφαρμογές που βοηθούν να εμφυτεύσουμε martingales σε διάφορα προβλήματα και κατόπιν να επικαλεστούμε την πλούσια θεωρία των martingales για να βγάλουμε διάφορα συμπεράσματα. Μια τέτοια σημαντική εφαρμογή είναι όταν έχουμε μια ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών Y_n με ροπογεννήτρια

$$\phi(\lambda) = E[exp(\lambda Y_n)] < \infty.$$

Τότε οι τυχαίες μεταβλητές $X_n = exp(\lambda Y_n)/\phi(\lambda)$ έχουν μέση τιμή 1 και επομένως το γινόμενό τους μας οδηγεί σε μια ολόκληρη παραμετρική οικογένεια martingales με παράμετρο λ. Πραγματικά, αν ορίσουμε

$$R_n = exp(\lambda \sum_{i=1}^{n} Y_i)/\phi(\lambda)^n,$$

τότε έχουμε ότι η (R_n) είναι martingale ως προς τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ που ορίζεται θέτοντας $\mathcal{F}_0 = \{\emptyset, \Omega\}$ την τετριμένη σ-άλγεβρα και $\mathcal{F}_n = \sigma(Y_k : 1 \leq k \leq n)$, $n \geq 1$.

Στην περίπτωση που υπάρχει κάποιο $\lambda_0 \neq 0$ τέτοιο ώστε $\phi(\lambda_0) = 1$, τότε παίρνουμε ένα ιδιαίτερα χρήσιμο martingale της οικογένειας. Σε αυτή την περίπτωση, θέτοντας $S_n = \sum_{i=1}^{n} Y_i$, έχουμε ότι το

$$R_n = e^{\lambda_0 S_n},\ n \geq 0,$$

είναι martingale. Αυτό το martingale μας επιτρέπει να συνάγουμε χρήσιμες πληροφορίες για τα μερικά αθροίσματα S_n μιας ακολουθίας τυχαίων μεταβλητών.

Παράδειγμα 11. *Το martingale του Doob: Έστω X μια τυχαία μεταβλητή με $E[|X|] < \infty$ και $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση. Ορίζουμε*

$$X_n = E[X|\mathcal{F}_n],\ n \geq 0.$$

Τότε η $(X_n)_{n\geq 0}$ είναι $(\mathcal{F}_n)_{n\geq 0}$-martingale. Πράγματι, έχουμε άμεσα ότι η $(X_n)_{n\geq 0}$ είναι προσαρμοσμένη στην $(\mathcal{F}_n)_{n\geq 0}$. Επιπλέον

$$E[|X_n|] = E[|E[X|\mathcal{F}_n]|] \leq E[|X|] < \infty.$$

Τέλος, χρησιμοποιώντας την ιδιότητα του πύργου και τον ορισμό του martingale έχουμε για $n \geq 0$

$$E[X_{n+1}|\mathcal{F}_n] = E[E[X|\mathcal{F}_{n+1}]|\mathcal{F}_n] = E[X|\mathcal{F}_n] = X_n.$$

Πρόταση 31. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας, $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$ και $(M_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών. Έστω, επίσης, f συνάρτηση ορισμένη σε ένα διάστημα $I \subseteq \mathbb{R}$ με $P(M_n \in I) = 1$, $E[|f(M_n)|] < \infty$, $n \geq 0$.*

(1) *Αν $(M_n)_{n\geq 0}$ είναι martingale και η f είναι κυρτή τότε η $(f(M_n))_{n\geq 0}$ είναι submartingale.*
(2) *Αν $(M_n)_{n\geq 0}$ είναι submartingale και η f είναι αύξουσα και κυρτή τότε η $(f(M_n))_{n\geq 0}$ είναι submartingale.*

Απόδειξη. Για το (1) αρκεί να αποδείξουμε ότι η συνθήκη (3) του ορισμού των submartingales ισχύει, μια και οι συνθήκες (1) και (2) είναι άμεσα επαληθεύσιμες. Για $n \geq 0$ έχουμε

$$E[f(M_{n+1})|\mathcal{F}_n] \geq f(E[M_{n+1}|\mathcal{F}_n]) = f(M_n),$$

όπου στην ανισότητα χρησιμοποιήσαμε την ανισότητα Jensen και στην ισότητα των ορισμό των martingale. Ομοίως αποδεικνύεται και το (2). ∎

Πόρισμα 1. *Έστω $(M_n)_{n\geq 0}$ ένα martingale και $X_n = |M_n|$ και $Y_n = M_n^2$, $n \geq 0$. Τότε οι $(X_n)_{n\geq 0}$ και $(Y_n)_{n\geq 0}$ είναι submartingales.*

Θεώρημα 1. *Θεώρημα ανάλυσης του Doob: Έστω $X = (X_n)_{n\geq 0}$ ένα submartingale. Τότε υπάρχει ένα martingale $(M_n)_{n\geq 0}$ και μια αύξουσα σχεδόν παντού, προβλέψιμη ακολουθία $(A_n)_{n\geq 0}$ (δηλαδή $A_n \leq A_{n+1}$ a.s. για κάθε n και A_{n+1} να είναι $\mathcal{F}_n$-μετρήσιμη για κάθε n), τέτοια ώστε:*

$$X_n = X_0 + M_n + A_n,\ n \geq 0 \text{ με } M_0 = A_0 = 0.$$

Επιπλέον αυτή η ανάλυση είναι μοναδική.

Απόδειξη. Ορίζουμε

$$\begin{aligned} A_0 &= 0, \\ A_n &= \sum_{k=1}^{n} E[X_k - X_{k-1}|\mathcal{F}_{k-1}],\ n \geq 1. \end{aligned}$$

Προφανώς η A_n είναι $\mathcal{F}_{n-1}$-μετρήσιμη για $n \geq 1$ και επομένως η $(A_n)_{n\geq 0}$ είναι προβλέψιμη. Επίσης, επειδή η X είναι submartingale έχουμε ότι $E[X_k - X_{k-1}|\mathcal{F}_{k-1}] \geq 0$ σχεδόν παντού για κάθε $k \geq 1$ και επομένως η $(A_n)_{n\geq 0}$ είναι αύξουσα σχεδόν παντού. Επίσης έχουμε ότι

$$E[X_n|\mathcal{F}_{n-1}] - X_{n-1} = E[X_n - X_{n-1}|\mathcal{F}_{n-1}] = A_n - A_{n-1}$$

και επομένως

$$E[X_n|\mathcal{F}_{n-1}] - A_n = X_{n-1} - A_{n-1}.$$

Επειδή $A_n \in \mathcal{F}_{n-1}$, έχουμε και ότι

$$E[X_n - A_n|\mathcal{F}_{n-1}] = X_{n-1} - A_{n-1},$$

που δείχνει ότι η $(M_n)_{n\geq 0}$ με $M_n = X_n - A_n$, $n \geq 0$ είναι martingale και έχουμε την απόδειξη της ύπαρξης της ανάλυσης.

Έστω τώρα δυο αναλύσεις αυτού του τύπου:

$$\begin{aligned} X_n &= X_0 + M_n + A_n,\ n \geq 0, \\ X_n &= X_0 + M_n' + A_n',\ n \geq 0. \end{aligned}$$

Αφαιρώντας έχουμε ότι $M_n' - M_n = A_n - A_n'$. Επειδή οι A_n, A_n' είναι $\mathcal{F}_{n-1}$-μετρήσιμες, έχουμε ότι και η $M_n' - M_n$ είναι $\mathcal{F}_{n-1}$-μετρήσιμη. Επομένως

$$M_n' - M_n = E[M_n' - M_n|\mathcal{F}_{n-1}] = M_{n-1}' - M_{n-1} = A_{n-1} - A_{n-1}'$$

σχεδόν παντού. Συνεχίζοντας αναδρομικά, έχουμε $M_n' - M_n = M_0' - M_0 = 0$ σχεδόν παντού, αφού $M_0' = M_0 = 0$. Συνεπώς, έχουμε $M_n' = M_n$ σχεδόν παντού για κάθε $n \geq 0$ και μετά $A_n = A_n'$ που δείχνει ότι η ανάλυση είναι μοναδική. ∎

Πόρισμα 2. *Έστω $X = (X_n)_{n\geq 0}$ ένα supermartingale. Τότε υπάρχει ένα martingale $(M_n)_{n\geq 0}$ και μια αύξουσα σχεδόν παντού, προβλέψιμη ακολουθία $(A_n)_{n\geq 0}$ (δηλαδή $A_n \leq A_{n+1}$ a.s. για κάθε n και A_{n+1} να είναι $\mathcal{F}_n$-μετρήσιμη για κάθε n), τέτοια ώστε:*

$$X_n = X_0 + M_n - A_n,\ n \geq 0 \text{ με } M_0 = A_0 = 0.$$

Επιπλέον αυτή η ανάλυση είναι μοναδική.

Απόδειξη. Έστω $Y_n = -X_n$. Η $(Y_n)_{n\geq 0}$ είναι submartingale, οπότε υπάρχει η ανάλυση Doob ώστε

$$Y_n = Y_0 + M'_n + A'_n,\ n \geq 0.$$

Τότε $X_n = X_0 - M'_n - A'_n$. Θέτοντας $M_n = -M'_n$ και $A_n = A'_n$ παίρνουμε τη ζητούμενη ανάλυση. ■

5. Χρονοι στασης και martingales

Ορισμός 16. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας και $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$. Μια τυχαία μεταβλητή $T : \Omega \to \bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$ λέγεται χρόνος στάσης ως προς την $(\mathcal{F}_n)_{n\geq 0}$ αν $\{T \leq n\} \in \mathcal{F}_n$ για κάθε $n \geq 0$.*

Κάθε σταθερή τυχαία μεταβλητή είναι χρόνος στάσης. Συνήθως, όμως, οι χρόνοι στάσης που χρησιμοποιούνται στις εφαρμογές σχετίζονται με την πρώτη πραγματοποίηση ενός γεγονότος που σχετίζεται με κάποια ακολουθία τυχαίων μεταβλητών. Αν το γεγονός δεν συμβεί ποτέ, τότε ο αντίστοιχος χρόνος στάσης είναι άπειρος. Αυτός είναι άλλωστε και ο λόγος που επιτρέπουμε την τιμή ∞ στο πεδίο τιμών των χρόνων στάσης. Ο ορισμός του χρόνου στάσης λέει ουσιαστικά ότι η πραγματοποίηση ή μη του ενδεχομένου $\{T \leq n\}$ καθορίζεται από την πληροφορία που έχουμε ως το χρόνο n (την οποία 'εκφράζει' η σ-άλγεβρα $\mathcal{F}_n$).

Παράδειγμα 12. *Ένα παράδειγμα χρόνου διακοπής δίνει ο χρόνος πρώτης εισόδου σε κάποιο συγκεκριμένο υποσύνολο μια στοχαστικής διαδικασίας. Π.χ. αν $(X_n)_{n\geq 0}$ είναι μια ακολουθία τυχαίων μεταβλητών και $A \subseteq \mathbb{R}$, τότε η τυχαία μεταβλητή*

$$T = \inf\{n \geq 0 : X_n \in A\}$$

είναι χρόνος διακοπής ως προς την $(\sigma(X_k : 1 \leq k \leq n))_{n\geq 0}$.

Ο όρος χρόνος στάσης προέρχεται από τις στοιχηματικές διαδικασίες. Ένας χρόνος στάσης μπορεί εκεί να εννοηθεί ως ο χρόνος που πρέπει να σταματήσει να παίζει ένας παίκτης. Το ενδεχόμενο να σταματήσει να παίζει ο παίκτης μέχρι τη στιγμή n είναι λογικό να απαιτήσουμε να εξαρτάται μόνο από την πληροφορία που είναι διαθέσιμη τη στιγμή n που αντιστοιχεί στη σ-άλγεβρα $\mathcal{F}_n$. Αυτός είναι και ο λόγος που απαιτούμε στον ορισμό $\{T \leq n\} \in \mathcal{F}_n$ για κάθε $n \geq 0$.

Όπως έχουμε ήδη σημειώσει, η σ-άλγεβρα $\mathcal{F}_n$ σε μια διήθηση, αντιστοιχεί στην πληροφορία που υπάρχει μέχρι τη χρονική στιγμή n, δηλαδή περιλαμβάνει όλα εκείνα τα ενδεχόμενα για τα οποία μπορούμε να αποφανθούμε κατά πόσο πραγματοποιήθηκαν ή όχι μέχρι τη χρονική στιγμή n. Έχοντας έναν χρόνο διακοπής T, μας ενδιαφέρουν τα ενδεχόμενα για τα οποία μπορούμε να αποφανθούμε κατά πόσο πραγματοποιήθηκαν ή όχι μέχρι την τυχαία χρονική στιγμή T. Ο κατάλληλος ορισμός είναι ο επόμενος.

Ορισμός 17. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας, $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$ και T χρόνος στάσης ως προς την $(\mathcal{F}_n)_{n\geq 0}$. Ορίζουμε*

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq n\} \in \mathcal{F}_n,\ n \geq 0\}.$$

Πρόταση 32. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας, $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$ και T χρόνος στάσης ως προς την $(\mathcal{F}_n)_{n\geq 0}$. Η $\mathcal{F}_T$ είναι σ-άλγεβρα.*

Απόδειξη. Άσκηση. ∎

Πρόταση 33. *Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας, $(\mathcal{F}_n)_{n\geq 0}$ μια διήθηση στον $(\Omega, \mathcal{F}, P)$ και T, S χρόνοι στάσης ως προς την $(\mathcal{F}_n)_{n\geq 0}$ με $S \leq T$. Τότε $\mathcal{F}_S \subseteq \mathcal{F}_T$.*

Απόδειξη. Άσκηση. ∎

Ορισμός 18. *Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών, προσαρμοσμένη στη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και T χρόνος στάσης ως προς την $(\mathcal{F}_n)_{n\geq 0}$. Αν $\Pr[T < \infty] = 1$, τότε ορίζουμε την τυχαία μεταβλητή*

$$X_T = \sum_{n=0}^{\infty} X_n 1_{\{T=n\}}.$$

Πρόταση 34. *Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών, προσαρμοσμένη στη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και T χρόνος στάσης ως προς την $(\mathcal{F}_n)_{n\geq 0}$ με $\Pr[T < \infty] = 1$. Τότε η X_T είναι $\mathcal{F}_T$-μετρήσιμη.*

Απόδειξη. Άσκηση. ∎

6. Διακριτο στοχαστικο ολοκληρωμα

Οι τζογαδόροι όλων των εποχών ελπίζουν ότι με κάποια έξυπνη στρατηγική μπορούν να νικήσουν τα καζίνο, μετατρέποντας ένα δίκαιο παιχνίδι (ή ακόμα καλύτερα ένα ασύμφορο γι αυτούς παιχνίδι) σε συμφέρον. Μια τέτοια κλασική στρατηγική φαίνεται να είναι ο διπλασιασμός του στοιχήματος σε κάθε γύρο του παιχνιδιού μέχρι ο παίκτης να κερδίσει. Η διαίσθησή μας βέβαια λέει ότι αυτό δεν μπορεί να είναι δυνατό. Πραγματικά, αν υποθέσουμε ότι η περιουσία ενός παίκτη είναι πεπερασμένη και επομένως δεν του επιτρέπεται να στοιχηματίζει πάνω από ένα ποσό και ο παίκτης δεν έχει τη δυνατότητα να προβλέπει το μέλλον (και επομένως η στρατηγική του βασίζεται σε ό,τι έχει δει μέχρι τη στιγμή που αποφασίζει πόσο θα στοιχηματίσει στον επόμενο γύρο), τότε ένα δίκαιο παιχνίδι παραμένει δίκαιο, ό,τι στρατηγική και να ακολουθήσει. Αυτό είναι το αποτέλεσμα που παραθέτουμε παρακάτω ως θεώρημα διατήρησης της ιδιότητας του martingale του διακριτού στοχαστικού ολοκληρώματος. Για να το διατυπώσουμε χρειάζεται πρώτα να δώσουμε τον ορισμό του διακριτού στοχαστικού ολοκληρώματος.

Ορισμός 19. *Έστω $(M_n)_{n\geq 0}$ μια στοχαστική διαδικασία προσαρμοσμένη σε μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και $(A_n)_{n\geq 1}$ μια προβλέψιμη ακολουθία τυχαίων μεταβλητών ως προς τη ίδια διήθηση. Η στοχαστική διαδικασία $((A \bullet M)_n)_{n\geq 0}$ που ορίζεται θέτοντας*

$$\begin{aligned}(A \bullet M)_0 &= M_0,\\ (A \bullet M)_n &= M_0 + \sum_{i=1}^{n} A_i(M_i - M_{i-1}),\ n \geq 1,\end{aligned}$$

λέγεται διακριτό ολοκλήρωμα της $(A_n)_{n\geq 1}$ ως προς την $(M_n)_{n\geq 0}$.

Σκεφτόμενοι με όρους στοιχημάτων, μπορούμε να θεωρήσουμε ότι η στοχαστική διαδικασία $(M_n)_{n\geq 0}$ περιγράφει την εξέλιξη της περιουσίας του παίκτη όταν παίζει ένα τυχερό παιχνίδι και ποντάρει 1 χρηματική μονάδα σε κάθε γύρο. Στην περίπτωση αυτή η διαφορά $M_n - M_{n-1}$ αντιστοιχεί στην επίδραση που έχει ο n-οστός γύρος του παιχνιδιού στην εξέλιξη της περιουσίας του παίκτη. Η ιδιότητα της προβλεψιμότητας της $(A_n)_{n\geq 1}$ μας λέει ότι η A_n εξαρτάται μόνο από την πληροφορία $\mathcal{F}_{n-1}$ που είναι διαθέσιμη πριν τον n-οστό γύρο του παιχνιδιού. Συνεπώς μπορούμε να σκεφτόμαστε την $(A_n)_{n\geq 1}$ σαν τη στρατηγική πονταρίσματος του παίκτη που αποφασίζει το πόσες χρηματικές μονάδες να ποντάρει σε κάθε γύρο. Επομένως η $A_n(M_n - M_{n-1})$ αντιστοιχεί στην επίδραση που έχει ο n-οστός γύρος του παιχνιδιού στην εξέλιξη της περιουσίας του παίκτη αν ακολουθήσει τη στρατηγική πονταρίσματος που υπαγορεύει η $(A_n)_{n\geq 1}$ και η $((A \bullet M)_n)_{n\geq 0}$ περιγράφει την εξέλιξη της περιουσίας του παίκτη όταν παίζει σύμφωνα με τη στρατηγική αυτή.

Θεώρημα 2. *Θεώρημα διατήρησης της ιδιότητας του martingale, submartingale, supermartingale του διακριτού στοχαστικού ολοκληρώματος: Έστω $(M_n)_{n\geq 0}$ μια στοχαστική διαδικασία προσαρμοσμένη στη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και $(A_n)_{n\geq 1}$ μια προβλέψιμη ακολουθία τυχαίων μεταβλητών ως προς την ίδια διήθηση.*

(1) *Αν η* $(M_n)_{n\geq 0}$ *είναι martingale και κάθε* A_n *είναι φραγμένη τυχαία μεταβλητή, τότε η* $((A \bullet M)_n)_{n\geq 0}$ *είναι martingale.*

(2) *Αν η* $(M_n)_{n\geq 0}$ *είναι submartingale (αντίστοιχα supermartingale) και κάθε* A_n *είναι μη-αρνητική φραγμένη τυχαία μεταβλητή, τότε η* $((A \bullet M)_n)_{n\geq 0}$ *είναι submartingale (αντίστοιχα supermartingale).*

Απόδειξη. Έχουμε εύκολα ότι $E[|(A \bullet M)_n|] < \infty$, $n \geq 0$, από το γεγονός ότι η $(A_n)_{n\geq 1}$ είναι φραγμένη σε συνδυασμό με το ότι $E[|M_n|] < \infty$, $n \geq 0$. Επίσης έχουμε ότι η $((A \bullet M)_n)_{n\geq 0}$ είναι προσαρμοσμένη στην $(\mathcal{F}_n)_{n\geq 0}$, αφού η $(M_n)_{n\geq 0}$ είναι προσαρμοσμένη σε αυτήν και η $(A_n)_{n\geq 1}$ προβλέψιμη. Τέλος, για κάθε $n \geq 0$, έχουμε ότι

$$
\begin{aligned}
E[(A \bullet M)_{n+1}|\mathcal{F}_n] - (A \bullet M)_n &= E[(A \bullet M)_{n+1} - (A \bullet M)_n|\mathcal{F}_n] \\
&= E[A_{n+1}(M_{n+1} - M_n)|\mathcal{F}_n] \\
&= A_{n+1}(E[M_{n+1}|\mathcal{F}_n] - M_n).
\end{aligned}
$$

Στην περίπτωση (1) που η $(M_n)_{n\geq 0}$ είναι martingale η τελευταία ποσότητα είναι 0 και επομένως έχουμε ότι η $((A \bullet M)_n)_{n\geq 0}$ είναι martingale. Ανάλογα έχουμε και την περίπτωση (2). ∎

Θεώρημα 3. *Θεώρημα διατήρησης της ιδιότητας του martingale, submartingale, supermartingale της σταματημένης διαδικασίας: Έστω* $(M_n)_{n\geq 0}$ *μια στοχαστική διαδικασία προσαρμοσμένη στη διήθηση* $(\mathcal{F}_n)_{n\geq 0}$ *και* T *ένας χρόνος στάσης ως προς την ίδια διήθηση. Αν η διαδικασία* $(M_n)_{n\geq 0}$ *είναι martingale (αντίστοιχα submartingale, supermartingale) τότε και η 'σταματημένη' διαδικασία σε χρόνο* T $(M_{n\wedge T})_{n\geq 0}$ *είναι επίσης martingale (αντίστοιχα submartingale, supermartingale).*

Απόδειξη. Θέτοντας $A_n = 1_{\{T\geq n\}}$, $n \geq 1$. Επειδή η T είναι χρόνος στάσης ως προς τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$, έχουμε ότι $\{T \geq n\} = \{T \leq n-1\}^c \in \mathcal{F}_{n-1}$, $n \geq 1$ και άρα η $A_n = 1_{\{T\geq n\}}$ είναι $\mathcal{F}_{n-1}$-μετρήσιμη, $n \geq 1$. Επομένως η ακολουθία $(A_n)_{n\geq 1}$ είναι προβλέψιμη ως προς την $(\mathcal{F}_n)_{n\geq 0}$. Εφαρμόζουμε το θεώρημα διατήρησης της ιδιότητας του martingale, submartingale, supermartingale του διακριτού στοχαστικού ολοκληρώματος για την ακολουθία αυτή. Επειδή

$$(A \bullet M)_n = M_{n\wedge T},\ n \geq 0,$$

έχουμε ότι αν $(M_n)_{n\geq 0}$ είναι martingale (αντίστοιχα submartingale, supermartingale) τότε και η $(M_{n\wedge T})_{n\geq 0}$ είναι επίσης martingale (αντίστοιχα submartingale, supermartingale). ∎

Πόρισμα 3. *Έστω* $(M_n)_{n\geq 0}$ *ένα martingale (αντίστοιχα submartingale, supermartingale) ως προς τη διήθηση* $(\mathcal{F}_n)_{n\geq 0}$ *και* T *ένας χρόνος στάσης ως προς την ίδια διήθηση. Τότε η ακολουθία*

$(E[M_{n\wedge T}])_{n\geq 0}$ είναι σταθερή (αντίστοιχα αύξουσα, φθίνουσα). Ειδικότερα έχουμε $E[M_{n\wedge T}] = E[M_0]$ (αντίστοιχα $E[M_{n\wedge T}] \geq E[M_0]$, $E[M_{n\wedge T}] \leq E[M_0]$).

7. Το θεωρημα επιλεκτικης στασης

Όπως έχουμε δει, ένα martingale $(M_n)_{n\geq 0}$ έχει σταθερή μέση τιμή αν το θεωρήσουμε οποιαδήποτε σταθερή χρονική στιγμή n, δηλαδή $E[M_n] = E[M_0]$, $n \geq 0$. Τι γίνεται όμως σε μια χρονική στιγμή που καθορίζεται από κάποιο χρόνο στάσης T; Ισχύει $E[M_T] = E[M_0]$; Η απάντηση είναι ότι αυτό δεν μπορεί να ισχύει γενικά. Πραγματικά, αν (M_n) είναι ένα martingale με $M_0 = a$, θέσουμε $T = \inf\{n \geq 0 : M_n = b\}$, $b \neq a$ και ισχύει $\Pr[M < \infty] = 1$ τότε έχουμε $E[M_T] = b \neq a = E[M_0]$. Το θεώρημα επιλεκτικής στάσης μας δίνει συνθήκες που μας επιτρέπουν να συμπεράνουμε ότι $E[M_T] = E[M_0]$ για ένα martingale $(M_n)_{n\geq 0}$ και ένα χρόνο στάσης T. Ανάλογα αποτελέσματα ισχύουν και όταν $(M_n)_{n\geq 0}$ είναι submartingale ή supermartingale.

Έχουμε ήδη δείξει στο τέλος της προηγούμενης παραγράφου ότι αν $(M_n)_{n\geq 0}$ είναι ένα martingale (αντίστοιχα submartingale, supermartingale) ως προς διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και T ένας χρόνος στάσης ως προς την ίδια διήθηση τότε $E[M_{n\wedge T}] = E[M_0]$ (αντίστοιχα $E[M_{n\wedge T}] \geq E[M_0]$, $E[M_{n\wedge T}] \leq E[M_0]$). Επιπλέον, στο σύνολο $\{T < \infty\}$ έχουμε ότι $\lim_{n\to\infty} M_{n\wedge T} = M_T$ (αυτό ισχύει για οποιαδήποτε ακολουθία τυχαίων μεταβλητών, χωρίς καμιά ιδιαίτερη υπόθεση για τον τρόπο εξάρτησης των μελών της). Συνεπώς, η συνθήκη

$$P(T < \infty)$$

συνεπάγεται ότι

$$M_{n\wedge T} \stackrel{a.s.}{\to} M_T.$$

Συνεπώς για ένα martingale $(M_n)_{n\geq 0}$ και ένα χρόνο στάσης T με $P(T < \infty)$ θα μπορούσαμε να ισχυριστούμε ότι $E[M_T] = E[M_0]$ ως ακολούθως:

$$E[M_0] = \lim_{n\to\infty} E[M_{n\wedge T}] = E[\lim_{n\to\infty} M_{n\wedge T}] = E[M_T].$$

Πραγματικά η πρώτη ισότητα ισχύει διότι όπως έχουμε δείξει στο τέλος της προηγούμενης παραγράφου ισχύει $E[M_{n\wedge T}] = E[M_0]$, ενώ η τρίτη ισότητα έπεται από το ότι $M_{n\wedge T} \stackrel{a.s.}{\to} M_T$. Το μόνο που μένει να αιτιολογηθεί είναι η δεύτερη ισότητα, δηλαδή η εναλλαγή ορίου και μέσης τιμής. Η ισότητα αυτή δεν ισχύει γενικά, αλλά μόνο κάτω από ορισμένες συνθήκες. Οι συνθήκες που παρουσιάζονται στη βιβλιογραφία είναι ποικίλες. Παρουσιάζουμε κάποιες από αυτές στο επόμενο λήμμα.

Λήμμα 1. *Συνθήκες εναλλαγής ορίου και μέσης τιμής για σταματημένη ακολουθία: Έστω $(X_n)_{n\geq 0}$ ακολουθία τυχαίων μεταβλητών με $E[|X_n|] < \infty$, $n \geq 0$ και T τυχαία μεταβλητή με τιμές στο $\{0,1,2,\ldots\} \cup \{\infty\}$. Αν ικανοποιείται μια από τις παρακάτω συνθήκες*

(1) $P(T < \infty) = 1$, $E[|X_T|] < \infty$ *και* $\lim_{n\to\infty} E[X_n 1_{\{T>n\}}] = 0$.

(2) *Υπάρχει* $c \in \mathbb{R}$ *ώστε* $P(T \leq c) = 1$.

(3) $P(T < \infty) = 1$ *και υπάρχει* $c \in \mathbb{R}$ *ώστε* $|X_n| \leq c$, $n \geq 0$.

(4) $E[T] < \infty$, $E[|X_0|] < \infty$ *και υπάρχει* $c \in \mathbb{R}$ *ώστε* $|X_n - X_{n-1}| \leq c$, $n \geq 1$.

τότε

$$\lim_{n\to\infty} E[X_{n\wedge T}] = E[X_T].$$

Απόδειξη. (1) Έχουμε ότι

$$X_T = X_{n\wedge T} + (X_T - X_n)1_{\{T>n\}},\ n \geq 0.$$

Συνεπώς, παίρνοντας μέσες τιμές, έχουμε

$$E[X_T] = E[X_{n\wedge T}] + E[X_T 1_{\{T>n\}}] - E[X_n 1_{\{T>n\}}],\ n \geq 0.$$

Για να αποδείξουμε ότι $\lim_{n\to\infty} E[X_{n\wedge T}] = E[X_T]$, αρκεί να δείξουμε ότι $\lim_{n\to\infty} E[X_T 1_{\{T>n\}}] = 0$ και $\lim_{n\to\infty} E[X_n 1_{\{T>n\}}] = 0$. Η δεύτερη ισχύει από την τρίτη υπόθεση της (1). Για την πρώτη έπεται από την $E[|X_T|] < \infty$. Πράγματι, έχουμε ότι

$$E[|X_T| 1_{\{T>n\}}] = \sum_{k=n+1}^{\infty} E[|X_k| 1_{\{T=k\}}] \to 0,\ n \to \infty,$$

αφού

$$\sum_{k=0}^{\infty} E[|X_k| 1_{\{T=k\}}] = E[|X_T|] < \infty.$$

(2) Είναι εύκολο να δούμε ότι οι συνθήκες (2) συνεπάγονται τις (1) και συνεπώς το συμπέρασμα ισχύει. Μπορούμε, όμως, και άμεσα να δούμε και ότι το συμπέρασμα ισχύει. Πράγματι ο c μπορεί να θεωρηθεί χωρίς βλάβη ακέραιος και έχουμε $X_{n\wedge T} = X_T$ για $n \geq c$. Συνεπώς $E[X_{n\wedge T}] = E[X_T]$ για $n \geq c$ και έχουμε $\lim_{n\to\infty} E[X_{n\wedge T}] = E[X_T]$.

(3) Είναι πάλι εύκολο να δούμε ότι οι συνθήκες (3) συνεπάγονται τις (1) και συνεπώς το συμπέρασμα ισχύει. Εναλλακτικά, χρησιμοποιούμε το θεώρημα κυριαρχημένης σύγκλισης για την $(X_{n\wedge T})_{n\geq 0}$. Έχουμε ότι $X_{n\wedge T} \stackrel{a.s.}{\to} X_T$ αφού $P(T < \infty) = 1$. Επιπλέον, $E[|X_n|] \leq c$, $n \geq 0$. Συνεπώς $\lim_{n\to\infty} E[X_{n\wedge T}] = E[X_T]$.

(4) Χρησιμοποιούμε το θεώρημα κυριαρχημένης σύγκλισης για την $(X_{n\wedge T})_{n\geq 0}$. Έχουμε ότι $E[T]<\infty$ που δίνει $P(T<\infty)=1$ και συνεπώς $X_{n\wedge T}\overset{a.s.}{\to} X_T$. Επιπλέον

$$|X_{n\wedge T}| = |X_0 + \sum_{k=1}^{n\wedge T}(X_k - X_{k-1})| \leq |X_0| + cT, n\geq 0,$$

και $E[|X_0|+cT]<\infty$ από τις υποθέσεις. Συνεπώς $\lim_{n\to\infty} E[X_{n\wedge T}] = E[X_T]$. ∎

Είμαστε τώρα σε θέση να διατυπώσουμε και να αποδείξουμε το θεώρημα επιλεκτικής στάσης.

Θεώρημα 4. *Θεώρημα επιλεκτικής στάσης: Έστω $(M_n)_{n\geq 0}$ ένα martingale (αντίστοιχα submartingale, supermartingale) ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και T ένας χρόνος στάσης ως προς την ίδια διήθηση. Αν ικανοποιείται μια από τις παρακάτω συνθήκες*

(1) $P(T<\infty)=1$, $E[|M_T|]<\infty$ *και* $\lim_{n\to\infty} E[M_n 1_{\{T>n\}}]=0$.

(2) *Υπάρχει* $c\in\mathbb{R}$ *ώστε* $P(T\leq c)=1$.

(3) $P(T<\infty)=1$ *και υπάρχει* $c\in\mathbb{R}$ *ώστε* $|M_n|\leq c$, $n\geq 0$.

(4) $E[T]<\infty$, $E[|M_0|]<\infty$ *και υπάρχει* $c\in\mathbb{R}$ *ώστε* $|M_n - M_{n-1}|\leq c$, $n\geq 1$.

τότε

$$E[M_T]=E[M_0] \text{ (αντίστοιχα } E[M_T]\geq E[M_0], E[M_T]\leq E[M_0]).$$

Απόδειξη. Οι συνθήκες εξασφαλίζουν ότι $\lim_{n\to\infty} E[M_{n\wedge T}] = E[M_T]$. Επίσης έχουμε δεί ότι όταν $(M_n)_{n\geq 0}$ είναι martingale (αντίστοιχα submartingale, supermartingale) τότε $E[M_{n\wedge T}] = E[M_0]$ (αντίστοιχα $E[M_{n\wedge T}]\geq E[M_0]$, $E[M_{n\wedge T}]\leq E[M_0]$) και το συμπέρασμα έπεται. ∎

Το θεώρημα επιλεκτικής στάσης χρησιμοποιείται για τον υπολογισμό πιθανοθεωρητικών χαρακτηριστικών που αφορούν ή σχετίζονται με έναν χρόνο στάσης T (π.χ. της μέσης τιμής του T, της κατανομής του T, της κατανομής μιας τυχαίας τυχαίας μεταβλητής που εξαρτάται από το T κλπ.). Η ιδέα είναι ότι χρησιμοποιούμε τη σχέση $E[M_T]=E[M_0]$ για ένα κατάλληλο martingale $(M_n)_{n\geq 0}$ για το οποίο είναι εύκολος ο υπολογισμός της $E[M_0]$. Από την άλλη μεριά η $E[M_T]$ αναπτύσσεται συναρτήσει του πιθανοθεωρητικού χαρακτηριστικού που μας ενδιαφέρει. Έτσι προκύπτει κάποια εξίσωση την οποία λύνουμε και υπολογίζεται το αντίστοιχο πιθανοθεωρητικό χαρακτηριστικό. Η ιδέα αυτή γίνεται περισσότερο σαφής με το ακόλουθο παράδειγμα.

Παράδειγμα 13. *Θεωρούμε τον απλό συμμετρικό τυχαίο περίπατο που απεικονίζει την εξέλιξη της περιουσίας ενός παίκτη που παίζει ένα τίμιο παιχνίδι, π.χ. στρίβει ένα δίκαιο νόμισμα και σε κάθε γύρο ποντάρει 1 χρηματική μονάδα. Για να μοντελοποιήσουμε αυτή την κατάσταση θεωρούμε μια ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $(X_n)_{n\geq 0}$ με την ιδιότητα*

$$P(X_n=1)=P(X_n=-1)=\frac{1}{2},\ n\geq 0.$$

Αν S_0 είναι η αρχική περιουσία του παίκτη, τότε μετά από n γύρους η περιουσία του θα δίνεται από την ποσότητα

$$S_n = S_0 + \sum_{i=1}^{n} X_i, \ n \geq 0.$$

Η τυχαία μεταβλητή $S_n - S_0$ αντιστοιχεί στο καθαρό κέρδος του παίκτη (ή στη ζημία αν παίρνει αρνητικές τιμές) μετά από n γύρους του παιχνιδιού. Ας υποθέσουμε ότι ενδιαφερόμαστε για την πιθανότητα ο παίκτης να κερδίσει A χρηματικές μονάδες πριν χάσει B. Για το σκοπό αυτό ορίζουμε την τυχαία μεταβλητή

$$T = \inf\{n \geq 0 : S_n = A \text{ ή } S_n = -B\}.$$

Ο χρόνος T είναι ο χρόνος πρώτης εισόδου της στοχαστικής διαδικασίας $(S_n)_{n\geq 0}$ στο συνολο $\{A, -B\}$ και επομένως είναι χρόνος στάσης. Η ζητούμενη πιθανότητα είναι η $P(S_T = A)$. Έχουμε ότι

$$E[S_0] = 0$$

και

$$E[S_T] = AP(S_T = A) - BP(S_T = -B),$$

μια και οι μόνες δυνατές τιμές της S_T εκ του ορισμού της T είναι οι A και $-B$.

Επιβεβαιώνοντας ότι $P(T < \infty) = 1$, $(S_n)_{n\geq 0}$ είναι $(\mathcal{F}_n)_{n\geq 0}$-martingale ως προς τη διήθηση $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$, $n \geq 1$ και οι συνθήκες του θεωρήματος επιλεκτικής στάσης επαληθεύονται παίρνουμε ότι

$$AP(S_T = A) - BP(S_T = -B) = AP(S_T = A) - B[1 - P(S_T = A)] = 0$$

και λύνοντας για το $P(S_T = A)$ έχουμε

$$P(S_T = A) = \frac{B}{A+B}.$$

Θεώρημα 5. *Θεώρημα επιλεκτικής δειγματοληψίας του Doob: Έστω $(M_n)_{n\geq 0}$ ένα martingale ως προς τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και S, T χρόνοι στάσης ως προς την ίδια διήθηση, c σταθερά ώστε $P(S \leq T \leq c) = 1$. Τότε*

$$E[M_T | \mathcal{F}_S] = M_S.$$

Απόδειξη. Μπορούμε να υποθέσουμε, χωρίς βλάβη της γενικότητας, ότι η σταθερή c είναι ακέραιος. Έχουμε ότι

$$|M_T|, |M_S| \leq \sum_{n=0}^{c} |M_n|$$

και επομένως $E[|M_T|], E[|M_S|] < \infty$. Επιπλέον, η M_S είναι $\mathcal{F}_S$-μετρήσιμη από το σχετικό θεώρημα που έχουμε δει για τους χρόνους στάσης και το γεγονός οτι η $(M_n)_{n\geq 0}$ είναι προσαρμοσμένη στην $(\mathcal{F}_n)_{n\geq 0}$.

Μένει, επομένως να αποδείξουμε ότι για κάθε $A \in \mathcal{F}_S$ ισχύει

$$E[M_T 1_A] = E[M_S 1_A].$$

Ας θεωρήσουμε επομένως ένα $A \in \mathcal{F}_S$. Θεωρούμε την τυχαία μεταβλητή

$$R = S1_A + T1_{A^c}.$$

Είναι εύκολο να δούμε ότι η R είναι χρόνος στάσης. Επιπλέον, το θεώρημα επιλεκτικής στάσης είναι εφαρμόσιμο, επειδή οι R, T είναι φραγμένοι από τη c (συνθήκες (2) του θεωρήματος επιλεκτικής στάσης). Συνεπώς έχουμε $E[M_R] = E[M_T] = E[M_0]$. Όμως

$$\begin{aligned} E[M_R] &= E[M_S 1_A + M_T 1_{A^c}], \\ E[M_T] &= E[M_T 1_A + M_T 1_{A^c}]. \end{aligned}$$

Αφαιρώντας κατά μέλη έχουμε $E[M_S 1_A] - E[M_T 1_A] = 0$ και έχουμε το ζητούμενο. ∎

Γνωρίζουμε ότι αν $(M_n)_{n\geq 0}$ είναι martingale τότε $E[M_n] = E[M_0]$, $n \geq 0$. Προφανώς, το αντίστροφο δεν ισχύει, δηλαδή $E[M_n] = E[M_0]$, $n \geq 0$ δεν συνεπάγεται ότι $(M_n)_{n\geq 0}$ είναι martingale. Η παρακάτω πρόταση παρέχει ένα είδος αντιστρόφου αποτελέσματος.

Πρόταση 35. *Έστω $(M_n)_{n\geq 0}$ μια ακολουθία τυχαίων μεταβλητών προσαρμοσμένη στη διήθηση $(\mathcal{F}_n)_{n\geq 0}$. Υποθέτουμε ότι $E[|M_n|] < \infty$ για κάθε $n \geq 0$ και ότι $E[M_T] = E[M_0]$ για κάθε φραγμένο χρόνο στάσης T ως προς τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$. Τότε η $(M_n)_{n\geq 0}$ είναι martingale ως προς τη διήθηση $(\mathcal{F}_n)_{n\geq 0}$.*

Απόδειξη. Άσκηση. ∎

8. Βασικες ανισοτητες των martingales

Θα δούμε τώρα τρεις σημαντικές ανισότητες που αναφέρονται σε martingales, submartingales και supermartingales και οι οποίες είναι πολύ χρήσιμες στην ανάπτυξη της θεωρίας και των εφαρμογών τους. Πρόκειται για τη μεγιστική ανισότητα , την L^p-ανισότητα και την ανισότητα των προς -τα-πάνω-διασχίσεων του Doob.

Αρχίζουμε με μια στοιχειώδη ανισότητα που είναι συχνά χρήσιμη ως βοηθητικό αποτέλεσμα.

Πρόταση 36. *Η ανισότητα μεταφοράς δείκτη για submartingale: Έστω $(M_n)_{n\geq 0}$ ένα submartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$. Τότε*

$$E[M_m 1_A] \leq E[M_n 1_A],\ A \in \mathcal{F}_m,\ n \geq m.$$

Απόδειξη. Αφού $(M_n)_{n\geq 0}$ είναι submartingale ως προς την $(\mathcal{F}_n)_{n\geq 0}$, για $n \geq m$ έχουμε ότι $M_m \leq E[M_n|\mathcal{F}_m]$. Έστω, τώρα, ένα $A \in \mathcal{F}_m$. Πολλαπλασιάζοντας και τα δυο μέλη της τελευταίας ισότητας με 1_A και παίρνοντας μέση τιμή έχουμε

$$E[M_m 1_A] \leq E[E[M_n|\mathcal{F}_m]1_A] = E[E[M_n 1_A|\mathcal{F}_m]] = E[M_n 1_A].$$

Η πρώτη ισότητα ισχύει επειδή $A \in \mathcal{F}_m$, οπότε η 1_A είναι $\mathcal{F}_m$-μετρήσιμη και η δεύτερη από το θεώρημα διπλής μέσης τιμής. ■

Συνεχίζουμε με έναν ορισμό.

Ορισμός 20. *Δοθείσης μιας ακολουθίας τυχαίων μεταβλητών $(X_n)_{n\geq 0}$, η ακολουθία $(X_n^*)_{n\geq 0}$ με*

$$X_n^* = \sup_{0\leq m\leq n} X_m,\ n \geq 0,$$

λέγεται ακολουθία μεγίστων της $(X_n)_{n\geq 0}$.

Έχουμε την παρακάτω εκλέπτυνση της ανισότητας Markov που ισχύει για μη-αρνητικά submartingales.

Θεώρημα 6. *Μεγιστική ανισότητα του Doob: Έστω $(M_n)_{n\geq 0}$ μη-αρνητικό submartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$, με ακολουθία μεγίστων $(M_n^*)_{n\geq 0}$ και $c > 0$. Τότε*

$$P(M_n^* \geq c) \leq \frac{E[M_n 1_{\{M_n^* \geq c\}}]}{c} \leq \frac{E[M_n]}{c}.$$

Απόδειξη. Έστω ο χρόνος στάσης

$$T = \inf\{m \geq 0 : M_m \geq c\},$$

που αντιστοιχεί στο χρόνο πρώτης υπέρβασης της τιμής c από το submartingale. Τότε έχουμε ότι τα ενδεχόμενα $\{M_n^* \geq c\}$ και $\{T \leq n\}$ είναι ισοδύναμα και άρα

$$c1_{\{M_n^* \geq c\}} \leq M_T 1_{\{M_n^* \geq c\}} = M_T 1_{\{T\leq n\}} = \sum_{m=0}^{n} M_m 1_{\{T=m\}}.$$

Παίρνουμε, τώρα μέσες τιμές στην παραπάνω σχέση, λαμβάνοντας υπόψη και την ανισότητα

$$E[M_m 1_{\{T=m\}}] \leq E[M_n 1_{\{T=m\}}],\ n \geq m,$$

που ισχύει εφαρμόζοντας την ανισότητα μεταφοράς δείκτη για το submartingale $(M_n)_{n\geq 0}$ και για το ενδεχόμενο $\{T = m\} \in \mathcal{F}_m$. Έτσι παίρνουμε

$$cP(M_n^* \geq c) \leq \sum_{m=0}^{n} E[M_m 1_{\{T=m\}}] \leq \sum_{m=0}^{n} E[M_n 1_{\{T=m\}}] = E[M_n 1_{\{T\leq n\}}] = E[M_n 1_{\{M_n^* \geq c\}}] \leq E[M_n],$$

από το οποίο προκύπτει άμεσα η ανισότητα. ■

Στην πλειονότητα των εφαρμογών της μεγιστικής ανισότητας του Doob, αυτά που χρειάζονται είναι οι 'εξωτερικές ' ποσότητες $P(M_n^* \geq c)$ και $\frac{E[M_n]}{c}$, που είναι αυτές που δείχνουν και ότι η ανισότητα αυτή είναι κατά κάποιο τρόπο εκλέπτυνση της ανισότητας Markov. Όμως η μεσαία ποσότητα χρειάζεται για την απόδειξη της L^p ανισότητας του Doob και γι αυτό και αναφέρεται.

Πόρισμα 4. *Έστω $(M_n)_{n\geq 0}$ μη-αρνητικό submartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$, με ακολουθία μεγίστων $(M_n^*)_{n\geq 0}$, $c>0$ και $p\geq 1$. Τότε*

$$P(M_n^* \geq c) \leq \frac{E[M_n^p]}{c^p}.$$

Απόδειξη. Η συνάρτηση x^p είναι αύξουσα και κυρτή και συνεπώς αφού $(M_n)_{n\geq 0}$ μη-αρνητικό submartingale, έχουμε και ότι η διαδικασία $(M_n^p)_{n\geq 0}$ είναι και αυτή μη-αρνητικό submartingale. Εφαρμόζουμε τη μεγιστική ανισότητα του Doob και έχουμε το συμπέρασμα. ∎

Θα χρησιμοποιήσουμε τώρα την ανισότητα Hölder για να αποδείξουμε το παρακάτω λήμμα που σε συνδυασμό με τη μεγιστική ανισότητα του Doob δίνει σχεδόν αμέσως την L^p ανισότητα του Doob.

Λήμμα 2. *Αν X, Y είναι μη-αρνητικές τυχαίες μεταβλητές με $E[Y^p]<\infty$ για κάποιο $p>1$ και*

$$cP(X\geq c) \leq E[Y1_{\{X\geq c\}}],\ c\geq 0,$$

τότε

$$\| X \|_p \leq \frac{p}{p-1} \| Y \|_p .$$

Απόδειξη. Αποδεικνύουμε πρώτα το λήμμα κάτω από την επιπλέον υπόθεση ότι $E[X^p]<\infty$. Έχουμε ότι για οποιοδήποτε $z\geq 0$ ισχύει

$$z^p = p\int_0^z x^{p-1}dx = p\int_0^\infty x^{p-1}1_{\{z\geq x\}}dx,$$

οπότε για μια μη αρνητική τυχαία μεταβλητή X έχουμε

$$X^p = p\int_0^X x^{p-1}dx = p\int_0^\infty x^{p-1}1_{\{X\geq x\}}dx.$$

Αν επιπλέον έχουμε $E[X^p] < \infty$ τότε παίρνοντας μέσες τιμές έχουμε:

$$\begin{aligned}
E[X^p] &= p\int_0^\infty x^{p-1}P(X \geq x)dx \\
&\leq p\int_0^\infty x^{p-2}E[Y1_{\{X\geq x\}}]dx \\
&= pE\left[Y\int_0^\infty x^{p-2}1_{\{X\geq x\}}dx\right] \\
&= \frac{p}{p-1}E\left[Y(p-1)\int_0^\infty x^{p-2}1_{\{X\geq x\}}dx\right] \\
&= \frac{p}{p-1}E[YX^{p-1}].
\end{aligned}$$

Χρησιμοποιούμε τώρα την ανισότητα Hölder υψώνοντας το Y στην p και το X^{p-1} στο συζυγή εκθέτη $q = p/(p-1)$ για τον οποίο ισχύει $1/p + 1/q = 1$. Έχουμε

$$\| X \|_p^p = E[X^p] \leq \frac{p}{p-1}E[YX^{p-1}] \leq \frac{p}{p-1} \| Y \|_p \| X \|_p^{p-1} .$$

Διαιρώντας με $\| X \|_p^{p-1}$ έχουμε το συμπέρασμα.

Αν αφαιρέσουμε την υπόθεση $E[X^p] < \infty$, τότε θέτοντας $X_n = \min(X, n)$ έχουμε ότι οι X_n ικανοποιούν τη συνθήκη $E[X_n^p] < \infty$ και τη συνθήκη του λήμματος και επομένως η ανισότητα ισχύει, δηλαδή

$$\| X_n \|_p \leq \frac{p}{p-1} \| Y \|_p,\ n \geq 0.$$

Χρησιμοποιώντας το λήμμα Fatou έχουμε το συμπέρασμα. ■

Είμαστε τώρα σε θέση να αποδείξουμε την ανισότητα L^p του Doob.

Θεώρημα 7. *Ανισότητα L^p του Doob: Έστω $(M_n)_{n\geq 0}$ μη-αρνητικό submartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$, με ακολουθία μεγίστων $(M_n^*)_{n\geq 0}$ και $p > 1$. Τότε*

$$\| M_n^* \|_p \leq \frac{p}{p-1} \| M_n \|_p,\ n \geq 0.$$

Απόδειξη. Από τη μεγιστική ανισότητα του Doob έχουμε ότι

$$cP(M_n^* \geq c) \leq E[M_n 1_{\{M_n^* \geq c\}}],\ c \geq 0.$$

Εφαρμόζοντας τώρα το προηγούμενο λήμμα με $X = M_n^*$ και $Y = M_n$, παίρνουμε την ανισότητα.

■

Μια τελευταία ανισότητα η οποία είναι το κλειδί για την απόδειξη το βασικού θεωρήματος σύγκλισης των martingales είναι η ανισότητα των προς -τα-πάνω-διασχίσεων του Doob. Πρέπει, όμως, πρώτα να δώσουμε τον ορισμό μιας προς -τα-πάνω-διάσχισης.

Ορισμός 21. *Αν* $(c_i)_{0\leq i\leq n}$ *είναι μια ακολουθία πραγματικών αριθμών, ορίζουμε ως αριθμό των προς -τα-πάνω-διασχίσεων (upcrossings) ενός διαστήματος* $[a,b]$*, από την ακολουθία, το μεγαλύτερο αριθμό* k *για τον οποίο υπάρχουν* $2k$ *ακέραιοι* $0 \leq i_1 < j_1 < i_2 < j_2 < \cdots < i_k < j_k \leq n$ *για τους οποίους τα στοιχεία* $c_{i_1}, c_{i_2}, \ldots, c_{i_k}$ *είναι μικρότερα ή ίσα του* a *και τα στοιχεία* $c_{j_1}, c_{j_2}, \ldots, c_{j_k}$ *είναι μεγαλύτερα ή ίσα του* b.

Ο αριθμός των προς -τα-πάνω-διασχίσεων ενός διαστήματος $[a,b]$ από την ακολουθία $(c_i)_{0\leq i\leq n}$ δείχνει ακριβώς το πόσες φορές η ακολουθία 'διασχίζει' το σύνολο $[a,b]$, με την έννοια ότι θεωρούμε ότι μια διάσχιση έγινε, όταν η ακολουθία ξεκινήσει κάποια στιγμή από μια τιμή μικρότερη ή ίση του a και βρεθεί σε κάποια μελλοντική στιγμή σε μια τιμή μεγαλύτερη ή ίση του b. Ένα submartingale έχει όπως έχουμε πει κατά κάποιο τρόπο αύξουσα συμπεριφορά και επομένως περιμένουμε ότι υπάρχει κάποιο φράγμα στο μέσο πλήθος των προς -τα-πάνω-διασχίσεων ενός διαστήματος $[a,b]$ που μπορεί να κάνει. Ένα τέτοιο φράγμα δίνει η ανισότητα των προς -τα-πάνω-διασχίσεων του Doob. Πριν τη διατυπώσουμε, αποδεικνύουμε το εξής λήμμα που θα χρησιμοποιήσουμε στην απόδειξή της.

Λήμμα 3. *Αν* $(M_n)_{n\geq 0}$ *είναι μη-αρνητικό submartingale ως προς μια διήθηση* $(\mathcal{F}_n)_{n\geq 0}$ *και* $(A_n)_{n\geq 1}$ *είναι μια προβλέψιμη ακολουθία τυχαίων μεταβλητών ως προς την ίδια διήθηση και οι* A_n*,* $n \geq 1$*, παίρνουν μόνο τις τιμές 0 και 1, τότε*

$$E[(A \bullet M)_n] \leq E[M_n],\ n \geq 0.$$

Απόδειξη. Από το γεγονός ότι $A_k \in \mathcal{F}_{k-1}$, $A_k \in \{0,1\}$ και το ότι η $(M_n)_{n\geq 0}$ είναι submartingale έχουμε ότι

$$E[A_k(M_k - M_{k-1})|\mathcal{F}_{k-1}] = A_k E[M_k - M_{k-1}|\mathcal{F}_{k-1}] \leq E[M_k - M_{k-1}|\mathcal{F}_{k-1}].$$

Παίρνουμε τώρα μέσες τιμές στην ανισότητα και έχουμε

$$E[A_k(M_k - M_{k-1})] \leq E[M_k - M_{k-1}]$$

και αθροίζονας για $k = 1, 2, \ldots, n$ παίρνουμε το ζητούμενο. ■

Το λήμμα αυτό αποδεικνύει κάτι διαισθητικά προφανές: Σε ένα συμφέρον παιχνίδι, αν έχεις την επιλογή να ποντάρεις 1 μονάδα ή τίποτα σε κάθε γύρο, είναι προτιμότερο να ποντάρεις 1 μονάδα σε κάθε γύρο.

Είμαστε τώρα σε θέση να διατυπώσουμε και να αποδείξουμε την ανισότητα των προς -τα -πάνω διασχίσεων του Doob.

Θεώρημα 8. *Ανισότητα των προς -τα-πάνω-διασχίσεων του Doob: Έστω $(M_n)_{n\geq 0}$ ένα submartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$ και $a < b$. Τότε, για το πλήθος $U_n(a,b)$ των προς -τα-πάνω-διασχίσεων του $[a,b]$ από την $(M_i)_{0\leq i\leq n}$ ισχύει η ανισότητα*

$$E[U_n(a,b)] \leq \frac{(M_n - a)_+}{b-a}.$$

Απόδειξη. Καταρχήν, έχουμε ότι η $((M_n - a)_+)_{n\geq 0}$ είναι ένα μη-αρνητικό submartingale αφού η $(M_n)_{n\geq 0}$ είναι submartingale και η $f(x) = (x-a)_+$ είναι κυρτή αύξουσα συνάρτηση. Επίσης είναι φανερό ότι ο αριθμός των προς -τα-πάνω-διασχίσεων του $[a,b]$ από την $(M_i)_{0\leq i\leq n}$ ισούται με τον αριθμό των προς -τα-πάνω-διασχίσεων του $[0, b-a]$ από την $((M_i - a)_+)_{0\leq i\leq n}$.

Υποθέτουμε ότι υιοθετούμε την ακόλουθη στρατηγική πονταρίσματος (προβλέψιμη ακολουθία): Αρχίζουμε να ποντάρουμε μια μονάδα σε κάθε γύρο του 'παιχνιδιού' όταν η ακολουθία $((M_i - a)_+)_{0\leq i\leq n}$ πιάνει για πρώτη φορά την τιμή 0 και σταματάμε να ποντάρουμε όταν φτάσει ή ξεπεράσει την τιμή $b-a$. Συνεχίζουμε με τον ίδιο κανόνα, δηλαδή αρχίζουμε ξανά να ποντάρουμε μια μονάδα σε κάθε γύρο όταν η ακολουθία $((M_i - a)_+)_{0\leq i\leq n}$ ξαναπιάσει για πρώτη φορά την τιμή 0 και σταματάμε να ποντάρουμε όταν ξαναφτάσει ή ξαναξεπεράσει την τιμή $b-a$ κ.ο.κ. Προφανώς έχουμε ότι αυτή η στρατηγική πονταρίσματος $(A_n)_{n\geq 1}$ είναι προβλέψιμη ως προς την ακολουθία $\sigma((M_n - a)_+)_{n\geq 0})$ και συνεπώς και ως προς την ακολουθία $(\mathcal{F}_n)_{n\geq 0}$ ως προς την οποία είναι προσαρμοσμένη η $(M_n)_{n\geq 0}$. Συνεπώς, το διακριτό στοχαστικό ολοκλήρωμα της $((M_n - a)_+)_{n\geq 0}$ ως προς την $(A_n)_{n\geq 1}$ είναι submartingale, αφού η $(A_n)_{n\geq 1}$ είναι μη-αρνητική και φραγμένη. Επιπλέον ισχύει με επίκληση του λήμματος που μόλις αποδείξαμε ότι

$$E[(A \bullet M)_n] \leq E[(M_n - a)_+].$$

Όμως η στρατηγική πονταρίσματος είναι τέτοια ώστε εγγυάται κέρδη τουλάχιστον $(b-a)U_n(a,b)$, δηλαδή $(b-a)U_n(a,b) \leq (A \bullet M)_n$. Παίρνουμε μέσες τιμές και τελικά έχουμε

$$(b-a)E[U_n(a,b)] \leq E[(A \bullet M)_n] \leq E[(M_n - a)_+].$$

■

9. Σύγκλιση των martingales

Το κυριότερο αποτέλεσμα σύγκλισης για martingales, submartingales και supermartingales και ένα από τα σημαντικότερα στη Θεωρία Πιθανοτήτων είναι η ισχυρή σύγκλιση των ομοιόμορφα φραγμένων κατά μέση τιμή martingales, submartingales και supermartingales. Συγκεκριμένα, έχουμε το ακόλουθο αποτέλεσμα.

Θεώρημα 9. *Θεώρημα σύγκλισης φραγμένων martingales, submartingales και supermartingales του Doob:*

(1) *Έστω $(M_n)_{n\geq 0}$ submartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$ με $\sup_{n\geq 0} E[M_n^+] < \infty$. Τότε υπάρχει τυχαία μεταβλητή M_∞ τέτοια ώστε $M_n \stackrel{a.s}{\to} M_\infty$. Επιπλέον $M_\infty \in L^1(\Omega, \mathcal{F}, P)$.*

(2) *Έστω $(M_n)_{n\geq 0}$ supermartingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$ με $\sup_{n\geq 0} E[M_n^-] < \infty$. Τότε υπάρχει τυχαία μεταβλητή M_∞ τέτοια ώστε $M_n \stackrel{a.s}{\to} M_\infty$. Επιπλέον $M_\infty \in L^1(\Omega, \mathcal{F}, P)$.*

(3) *Έστω $(M_n)_{n\geq 0}$ martingale ως προς μια διήθηση $(\mathcal{F}_n)_{n\geq 0}$, ώστε $\sup_{n\geq 0} E[M_n^+] < \infty$ ή $\sup_{n\geq 0} E[M_n^-] < \infty$. Τότε υπάρχει τυχαία μεταβλητή M_∞ τέτοια ώστε $M_n \stackrel{a.s}{\to} M_\infty$. Επιπλέον $M_\infty \in L^1(\Omega, \mathcal{F}, P)$.*

Απόδειξη. Για το (1), για να αποδείξουμε ότι υπάρχει τυχαία μεταβλητή M_∞ τέτοια ώστε $M_n \stackrel{a.s}{\to} M_\infty$, αρκεί και πρέπει να αποδείξουμε ότι για το σύνολο

$$\Lambda = \{\omega \in \Omega : \limsup_n M_n(\omega) > \liminf_n M_n(\omega)\}$$

έχουμε $P(\Lambda) = 0$. Όμως από την πυκνότητα των ρητών στους πραγματικούς έχουμε ότι

$$\Lambda = \cup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \Lambda_{a,b},$$

με

$$\Lambda_{a,b} = \{\omega \in \Omega : \limsup_n M_n(\omega) \geq b, \liminf_n M_n(\omega) \leq a\}.$$

Συνεπώς, αρκεί να αποδείξουμε ότι $P(\Lambda_{a,b}) = 0$ για κάθε $a, b \in \mathbb{Q}$. Όμως $\omega \in \Lambda_{a,b}$ συνεπάγεται ότι το πλήθος των προς -τα-πάνω-διασχίσεων $U(a, b)$ του $[a, b]$ από τη $(M_n(\omega))_{n\geq 0}$ είναι άπειρο. Συνεπώς, $P(\Lambda_{a,b}) \leq P(U(a, b) = \infty)$. Αρκεί επομένως να αποδείξουμε ότι $P(U(a, b) = \infty) = 0$. Για το σκοπό αυτό θα αποδείξουμε το ισχυρότερο αποτέλεσμα ότι $E[U(a, b)] < \infty$, χρησιμοποιώντας την ανισότητα των προς -τα-πάνω-διασχίσεων του Doob.

Πράγματι, αν $U_n(a, b)$ είναι το πλήθος των προς -τα-πάνω-διασχίσεων του $[a, b]$ από την $(M_n(\omega))_{n\geq 0}$ μέχρι τη χρονική στιγμή n, τότε η ακολουθία $(U_n(a, b))_{n\geq 0}$ είναι αύξουσα και συνεπώς το όριο $U(a, b) = \lim_{n\to\infty} U_n(a, b)$ υπάρχει. Επιπλέον, επειδή η $(U_n(a, b))_{n\geq 0}$ είναι αύξουσα και μη-αρνητική μπορούμε να επικαλεστούμε το θεώρημα μονότονης σύγκλισης. Λαμβάνοντας υπόψη και την ανισότητα των προς -τα-πάνω-διασχίσεων του Doob, καθώς και το γεγονός ότι $\sup_{n\geq 0} E[M_n^+] < \infty$ έχουμε:

$$\begin{aligned} E[U(a,b)] &= \lim_{n\to\infty} E[U_n(a,b)] \\ &\leq \frac{1}{b-a} \sup_{n\geq 0} E[(M_n - a)^+] \\ &\leq \frac{1}{b-a} \left(\sup_{n\geq 0} E[M_n^+] + |a| \right) \\ &< \infty. \end{aligned}$$

Επομένως, αποδείξαμε ότι υπάρχει το με πιθανότητα 1 όριο M_∞ της $(M_n)_{n\geq 0}$. Θα αποδείξουμε τώρα ότι $E[|M_\infty|] < \infty$. Πράγματι, αφού η $(M_n)_{n\geq 0}$ είναι submartingale έχουμε ότι $E[M_0] \leq E[M_n]$, $n \geq 0$ και επομένως έχουμε

$$E[|M_n|] = E[M_n^+] + E[M_n^-] = 2E[M_n^+] - E[M_n] \leq 2E[M_n^+] - E[M_0].$$

Χρησιμοποιώντας, τώρα, το λήμμα Fatou σε συνδυασμό με την υπόθεση $\sup_{n\geq 0} E[M_n^+] < \infty$, έχουμε

$$E[\lim_{n\to\infty} |M_n|] \leq \liminf_{n\to\infty} E[|M_n|] \leq 2\sup_{n\geq 0} E[M_n^+] - E[M_0] < \infty.$$

Συνεπώς $E[|M_\infty|] = E[\lim_{n\to\infty} |M_n|] < \infty$ και έχουμε $M_\infty \in L^1(\Omega, \mathcal{F}, P)$.

Για το (2), αν η $(M_n)_{n\geq 0}$ είναι supermartingale με $\sup_{n\geq 0} E[M_n^-] < \infty$, τότε έχουμε ότι η $(-M_n)_{n\geq 0}$ είναι submartingale με $\sup_{n\geq 0} E[(-M_n)^+] < \infty$ και εφαρμόζοντας το (1) του θεωρήματος έχουμε το αποτέλεσμα.

Για το (3), αρκεί να συνδυάσουμε τα (1) και (2). ■

Θα πρέπει να τονιστεί ότι το παραπάνω θεώρημα δεν λέει ότι $M_n \xrightarrow{L^1} M_\infty$, αλλά μόνο ότι η οριακή με πιθανότητα 1 τυχαία μεταβλητή M_∞ ανήκει στον $L^1(\Omega, \mathcal{F}, P)$. Για να εξασφαλίσουμε σύγκλιση στον $L^1(\Omega, \mathcal{F}, P)$, χρειαζόμαστε μια ισχυρότερη συνθήκη και για το λόγο αυτό εισάγουμε την έννοια της ομοιόμορφης ολοκληρωσιμότητας μιας ακολουθίας τυχαίων μεταβλητών.

Ορισμός 22. *Μια οικογένεια τυχαίων μεταβλητών $\{X_i\}_{i\in I}$ λέγεται ομοιόμορφα ολοκληρώσιμη, αν*

$$\lim_{c\to\infty} \sup_{i\in I} E[1_{\{|X_i|\geq c\}}] = 0.$$

Η συνθήκη του ορισμού της ομοιόμορφης ολοκληρωσιμότητας δεν είναι εύκολα ελέγξιμη και για το λόγο αυτό είναι χρήσιμες απλούστερες ικανές συνθήκες. Δυο τέτοιες δίνονται στο επόμενο αποτέλεσμα.

Πρόταση 37. *Έστω $\{X_i\}_{i\in I}$ μια οικογένεια τυχαίων μεταβλητών.*

(1) *Αν $\sup_{i\in I} E[|X_i|^p] < \infty$ για κάποιο $p > 1$, τότε η $\{X_i\}_{i\in I}$ είναι ομοιόμορφα ολοκληρώσιμη.*
(2) *Αν υπάρχει τυχαία μεταβλητή Y τέτοια ώστε $P(|X_i| \leq Y) = 1$ για κάθε $i \in I$ και $E[Y] < \infty$ τότε η $\{X_i\}_{i\in I}$ είναι ομοιόμορφα ολοκληρώσιμη.*

Το θεώρημα σύγκλισης των martingales του Doob μπορεί να επεκταθεί σε μια πιο ισχυρή έκδοση που δίνει σύγκλιση στον $L^1(\Omega, \mathcal{F}, P)$ στην περίπτωση που το martingale είναι ομοιόμορφα ολοκληρώσιμο. Έχουμε

Θεώρημα 10. *Θεώρημα L^1-σύγκλισης των martingales:*

(1) *Έστω* $(M_n)_{n\geq 0}$ *ένα ομοιόμορφα ολοκληρώσιμο martingale ως προς μια διήθηση* $(\mathcal{F}_n)_{n\geq 0}$*. Τότε υπάρχει τυχαία μεταβλητή* M_∞ *τέτοια ώστε*

$$M_n \stackrel{a.s.}{\to} M_\infty,$$

$$M_\infty \in L^1(\Omega, \mathcal{F}, P),$$

$$M_n \stackrel{L^1}{\to} M_\infty,$$

$$M_n = E[M_\infty | \mathcal{F}_n].$$

(2) *Αν* $Y \in L^1(\Omega, \mathcal{F}, P)$ *και* $(\mathcal{F}_n)_{n\geq 0}$ *είναι μια διήθηση στον* $(\Omega, \mathcal{F}, P)$*, τότε η ακολουθία* $(M_n)_{n\geq 0}$ *με* $M_n = E[Y|\mathcal{F}_n]$*,* $n \geq 0$ *είναι ομοιόμορφα ολοκληρώσιμο martingale.*

Τα martingales έχουν μια πλούσια οριακή θεωρία που ασφαλώς δεν σταματάει εδώ. Πέρα από το οριακό θεώρημα για την ισχυρή σύγκλιση που διατυπώσαμε και αποδείξαμε υπάρχουν και οριακά θεωρήματα που αφορούν την κατά κατανομή σύγκλιση (ανάλογα του κεντρικού οριακού θεωρήματος).

10. Βιβλιογραφια - Πηγες

(1) Κουμουλλή, Γ. και Νεγρεπόντη, Σ. (1991) *Θεωρία Μέτρου.* Συμμετρία, Αθήνα.

(2) Χελιώτη, Δ. (2012) *Εισαγωγή στο Στοχαστικό Λογισμό.* Σημειώσεις, Αθήνα.

(3) Grimmett, G. and Stirzaker, D. (2001) *Probability and Random Processes, 3rd. Edition.* Oxford University Press, Oxford.

(4) Jacod, J. and Protter, P. (2003) *Probability Essentials, 2nd Edition.* Springer, Berlin.

(5) Steele, M.J. (2001) *Stochastic Calculus and Financial Applications.* Springer, New York.

(6) Williams, D. (1991) *Probability with Martingales.* Cambridge University Press, Cambridge.