הטכניון - מכון טכנולוגי לישראל

הפקולטה להנדסת חשמל

יסודות תהליכים אקראיים 048868

משה זכאי, עפר זיתוני, אדם שורץ

מהדורת תשס"ט 2008/9. גרסת 2/6/2009.

תוכן ענינים

1 מבוא

תורת התקשורת והבקרה מבוססות במידה רבה על אותות ועל מערכות.
מטרת קורס זה היא לתת בסיס מדויק למושגי "רעש" ומערכות דינמיות רועשות.
כך, נגדיר "תהליך אקראי" ונדון במאפייניו. נטפל בפרוט בתהליך אקראי המשמש מודל חשוב לרעשים - הוא תנועת בראון (הקשור ל"רעש לבן").

את מושגי המערכת וה"מסנן", שתוארו על ידי משוואות דיפרנציאליות, נחליף בהרחבה, שהיא משוואה דיפרנציאלית סטוכסטית. לצורך כך נפתח את האינטגרל הסטוכסטי ונדון בנוסחת איטו.

דוגמאות לשימושים בתקשורת ובקרה יובאו להלן. עיקרן:

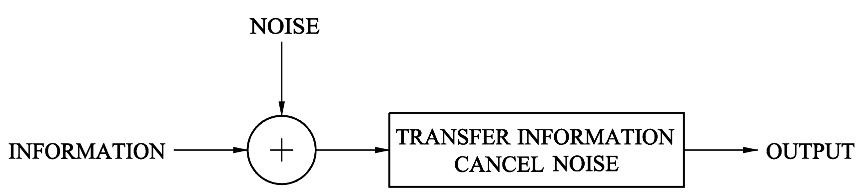
זהוי אות רצוי מתוך מדידה רועשת ("סינון").

הערכת השפעת רעש על ביצוע מערכות - לצורך תכנון בקרה.

עיקר הקורס עוסק בבנית מודלים מתמטיים, שיאפשרו <u>ניסוח</u> של בעיות אלו בצורה מדויקת, ומתן כלים בסיסיים לטיפול בהן.

<u>דוגמאות</u>

<u>תקשורת</u>



איור 1.1: מודל תקשורת

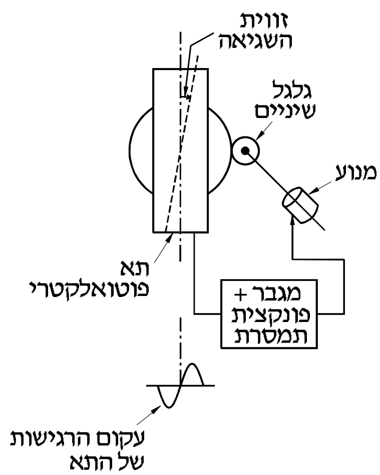
המודל הפשוט ביותר מניח $y(t) = x(t) + n(t)$ כאשר $n(t)$ הוא הרעש.

במציאות - לתווך יש תכונות דנמיות, ולכן הקשר מסובך הרבה יותר.

<u>השאלה:</u> שחזר את $x(t)$ בצורה "מיטבית", בהנתן $\{y(s),\ \ 0 \le s \le t\}$.

<u>מודל עקיבה:</u>

מקור הרעשים:



איור 1.2: מערכת עקיבה

התווך שבן הגוף למשקפת

רעשים בתא הפוטו-אלקטרי

רעשים במערכת הבקרה ובמנוע

<u>שאלות:</u>

בהנתן מודל לתנועת הגוף, כמה זמן בממוצע עד ליציאת הגוף משדה הראיה?

מה סיכוי האיבוד בזמן נתון?

כיצד לתכנן מערכת בקרה טובה?

<u>מהלך הקורס</u>

- חזרה והרחבה - מרחבים ומרחבי הילברט (לימוד עצמי)
- חזרה והרחבה - מרחבי הסתברות, מ"א, תוחלת, תהליכים. הגדרה מדויקת - תוחלת מותנית.
- תהליכי מרטינגל, תנועת בראון ("רעש לבן")
- אינטגרל סטוכסטי.
- משוואות דיפרנציאליות סטוכסטיות
- חישובים - הקשור למשוואות חלקיות

מה <u>לא</u> נעסוק:

- תהליכים גאוסיים.
- תכונות מסדר שני.
- תהליכי קפיצות.

ובחזרה לדוגמאות:

האותות בהם אנו עוסקים, בהנחה <u>שאין</u> רעשים, מתוארים על ידי מערכות :

$$\frac{dx}{dt} = f\left(x\left(t\right)\right)$$

ובנוכחות רעש לבן $n(t)$

$$\frac{dx}{dt} = f\left(x\left(t\right)\right) + n\left(t\right)$$

וכדי להבין ולטפל במשוואה כזו, יש ללמוד על משוואות סטוכסטיות.

הקורס נותן בסיס לטפל בשאלות כמו:

נניח שמודדים את $x\left(t\right)$ דרך מערכת מדידה רועשת. ואז מודל אות

$$\frac{dx\left(t\right)}{dt} = f\left(x\left(t\right)\right) + n_1\left(t\right)$$

מודל מדידה

$$\frac{dy\left(t\right)}{dt} = h\left(x\left(t\right)\right) + n_2\left(t\right)$$

שאלת הסינון היא: חשב משערך אופטימלי (במובן שגיאה ריבועית ממוצעת) עבור $x(t)$ בהנתן המדידה $\{y(s) : s \le t\}$. כלומר חשב תוחלת מותנית

$$\mathbb{E}(x(t) \mid y(s), s \le t). \tag{1.1}$$

מודל העקיבה יהיה שוב מהצורה

$$\frac{dx\left(t\right)}{dt}=f\left(x\left(t\right)\right)+n\left(t\right)$$

כאשר $x(t)$ החא שגיאת העקיבה, או ווקטור ששגיאת העקיבה היא אחת מרכיביו. מסתבר שניתן לחשב את תוחלת הזמן לניתוק נעילה, בהנתן המצב ההתחלתי, על ידי פתרון משוואה דיפרנציאלית חלקית מסדר שני. בדרך כלל הפיתרון המעשי הוא נומרי.

קורס זה יכין אותנו גם לטיפול בשאלות של רעש "קטן": נניח

$$\frac{dx\left(t\right)}{dt}=f\left(x\left(t\right)\right)+\varepsilon n\left(t\right)$$

כאשר ε הוא קטן. זהו אכן המצב במערכות הנדסיות רבות. למשל במערכות תקשורת, הרעש בחוג נעילת פאזה PLL קטן יחסית, והרעש המתווסף לאות קטן גם הוא. עור ε קטן נוכל לעיתים לקבל נוסחאות אנליטיות מפורשות עבור הגדלים המעניינים, כגון שגיאת הסינון, זמן בריחה מנעילה וכד'.

קורס זה נותן את התשתית המתמטית לתחום ההסתברות ותהליכים אקראיים, במיוחד בתחומי התקשורת והבקרה. קורסים קשורים כוללים תנודות רחבות משווי משקל (Large deviations).

2 הסתברות

בפרק זה נקים את הבסיס המתמטי לתורת ההסתברות, נלמד מושגי התכנסות וקירובים, ונפתח אינטואיציה עבור מושגים כגון תוחלת מותנית.

מדוע צריך בסיס מתמטי? אנו רוצים להגדיר "מאורעות" - אשר במובן הנדסי הם ניתנים לצפיה או למדידה. לשם כך יש צורך במבנה, המתואר להלן כ"אוסף המאורעות". מסתבר כי מבחינה מתמטית, לא ניתן להגדיר הסתברות כך שגם תהיה בעלת מבנה סביר, המתאר תכונות אינטאי-טיביות, וגם תהיה מוגדרת עבור כל קבוצת מקרים. כדי להגדיר מבנה מתמטי הכולל הסתברות, תוחלת ועוד, יש לכן צורך במבנה כמתואר בסעיף הבא.

2.1 $\sigma-$ שדות, מרחבים מדידים ומרחב הסתברות

יהי Ω אוסף של אלמנטים ω ;$\Omega=\{\omega\}$ הוא מרחב הדגמים.

הגדרה 2.1 יהי $\mathcal{F}$ אוסף תת קבוצות של Ω המהווה *sigma-field* (סימון שקול---*σ-field*), כלומר מקיים:

(א) $A\in\mathcal{F}$ גורר $A^C\in\mathcal{F}$ (המשלים של A) $\Omega\in\mathcal{F}$.

(ב) $A_n\in\mathcal{F}, n=1,2,3...$ גורר $\bigcup_{n=1}^{\infty}A_n\in\mathcal{F}$ (איחוד בן מניה).

קל לראות כי מ-(א) ו-(ב) נובע גם כי $A\in\mathcal{F}$ וכן $B\in\mathcal{F}$ גורר $A\cap B\in\mathcal{F}$.

$\mathcal{F}$ הוא מרחב המאורעות, והזוג $(\Omega,\mathcal{F})$ נקרא מרחב מדיד *(measurable space)* .

בהנתן אוסף קבוצות $A_\alpha\subset\Omega$, כאשר $\alpha\in I$ קבוצת אינדקסים כלשהי, קיים σ-field מינימלי הכולל את כל ה- A_α והוא יסומן ע"י $\sigma\left(A_\alpha,\ \alpha\in I\right)$

תרגיל 2.2 יהיו $\mathcal{F}_1$ ו- $\mathcal{F}_2$ *σ-fields* על Ω . נגדיר: $\mathcal{F}=\{A:A\in\mathcal{F}_1\ \ and\ A\in\mathcal{F}_2\}$. הוכח כי $\mathcal{F}$ היא *σ-field*.

בעזרת תרגיל זה והלמה של $Zorn$ ניתן להוכיח קיום ה- σ-field המינימלי.

הגדרה 2.3 במרחב טופולוגי, ה- σ שדה המינימלי המכיל את כל הקבוצות הסגורות של המרחב נקרא "שדה בורל ", ויסומן ב- $\mathbb{B}$ או ב- $\mathbb{B}\left(\mathbb{R}\right)$

בשל ההגדרה של σ-field, אפשר בצורה שקולה להגדיר את שדה בורל להיות זה המכיל את כל הקבוצות הפתוחות.

תרגיל 2.4 הוכח על סמך הגדרת ה- $\sigma - field$ המינימלי:

$\mathbb{B}(\mathbb{R}) =$

$=$ (כל האינטרוולים מהצורה (a, b)) σ

$=$ (כל האינטרוולים מהצורה $[a, b]$) σ

$=$ (כל האינטרוולים מהצורה $(-\infty, b]$) σ

(כל האינטרוולים מהצורה $(-\infty, b], b \in \mathbb{Q}$) σ

כאשר $\mathbb{Q}$ הוא אוסף המספרים הרציונליים (אוסף זה הוא בן מניה). רמז: מההגדרה של קבוצה פתוחה, כל קבוצה פתוחה על הישר היא איחוד של אינטרוולים פתוחים, וקל להראות שאפשר להסתפק באיחוד של אינטרוולים שקצותיהם רציונליים.

כעת נגדיר סיגמה-שדה $\mathcal{F}$ להיות הסיגמה שדה המינימלי המכיל את כל האינטרוולים הסגורים שקצותיהם מספרים שלמים. הראה כי הקבוצות $k \leq x < k + 0.5$ הן מאורעות על $\mathbb{B}$ אך לא על $\mathcal{F}$.

אפשר לחשוב על Ω כאוסף כל ה"מאורעות האלמנטריים" שיכולים לקרות, ועל $\mathcal{F}$ כאוסף התוצאות האפשריות של מדידות שביכולתינו לבצע. במובן זה $\mathcal{F}$ מתאר את המידע שעומד לרשותינו.

הגדרה 2.5 <u>מרחב הסתברות</u> הוא שלשה $(\Omega, \mathcal{F}, \mathbb{P})$ כאשר $(\Omega, \mathcal{F})$ הוא מרחב מדיד ($measurable\ space$) כלומר מרחב עם σ-$field$ ו- $\mathbb{P}$ פונקצית הסתברות, המוגדרת על כל $A \in \mathcal{F}$, כך ש-

(א) $0 \leq \mathbb{P}(A) \leq 1$

(ב) $\mathbb{P}(\Omega) = 1$

(ג) אם $A_n \in \mathcal{F}, n = 1, 2, ..$ ו- $A_n \cap A_m = \emptyset$ עבור כל $n \neq m$, אזי

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

תכונה (ג) נקראת countable additivity. היא אקוויולנטית לזוג הדרישות: ראשית, finite additivity, **כלומר תכונה (ג) מתקיימת עבור אוסף סופי, ושנית אם** $B_n \in \mathcal{F}, n = 1, 2, \ldots$, $B_n \subset B_{n+1}$ **ו-** $\bigcup_{n=1}^{\infty} B_n = B$ **אזי** $\mathbb{P}(B_n) \to \mathbb{P}(B)$.

תהי g פונקציה המוגדרת בין שני מרחבים מדידים $g:(X,B)\to\left(Y,\bar{B}\right)$ כאשר $B,\bar{B}$ הם ה-σ-שדות בתחום ובטווח בהתאמה.

הגדרה 2.6 הפונקציה $g:(X,B)\to\left(Y,\bar{B}\right)$ נקראת <u>מדידה</u> אם לכל קבוצה $\bar{b}\in\bar{B}$ מתקיים

$$\{x\in X:g(x)\in\bar{b}\}=g^{-1}\left(\bar{b}\right)\in B$$

אם $B,\bar{B}$ הם ה- σ - שדות של בורל אזי g נקראת "מדידה בורל" או בקיצור "פונקציית בורל".

תרגיל 2.7 הראה כי כל פונקצית מדרגות היא מדידה בורל.

<u>פונקציה רציפה</u> בין מרחבים טופולוגיים היא פונקציה $g:X\to Y$ המקיימת כי לכל קבוצה פתוחה $\bar{b}\subset Y$, הקבוצה $g^{-1}\left(\bar{b}\right)\subset X$ גם כן פתוחה. קל לראות שהגדרה זו דומה להגדרת פונקציה מדידה.

תרגיל 2.8 (א) הראה כי במקרה כי X,Y הם מרחבים מטריים (הגדרה *8.2*), ההגדרה להלן זהה להגדרה של פונקציה רציפה מחדו"א (לפי גבולות).

(ב) הוכח כי פונקציה רציפה היא מדידה בורל.

הגדרה 2.9 <u>משתנה אקראי</u> הוא פונקציה ממשית על Ω $X(\omega),\omega\in\Omega$ המקיימת את התנאי הבא:

לכל α ממשי מתקיים: $\{\omega:X(\omega)\leq\alpha\}\in\mathcal{F}$ כלומר הקבוצה היא מאורע.

הגדרה זו תלויה במרחב המדיד $(\Omega,\mathcal{F})$ אך לא ב- $\mathbb{P}$.

במילים אחרות, משתנה אקראי היא פונקציה מדידה ממרחב מדיד לשדה בורל:

$X:(\Omega,\mathcal{F})\to(\mathbb{R},\mathbb{B})$.

אפשר לחשוב על מ"א כעל דרך לבצע מדידות שהן "אפשריות" מבחינת המידע העומד לרשותינו באוסף המאורעות $\mathcal{F}$.

דוגמה 2.10 יהיה $\Omega=\{1,2,3\}$ ו- $\mathcal{F}_1=\{\emptyset,\Omega,\{1\},\{2,3\}\}$. אזי הפונקציה $Y(\omega)=\omega$ איננה מ"א על $(\Omega,\mathcal{F}_1)$. למעשה, כל מ"א על מרחב מדיד זה הוא בהכרח מהצורה:

$$X(\omega)=\begin{cases}\alpha & \omega=1\\ \beta & \omega=2\ or\ \omega=3\end{cases}$$

דוגמה 2.11 נבחר $A\in\mathcal{F}$. הפונקציה $I_A(\omega)$ המוגדרת

$$I_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & otherwise\end{cases}$$

נקראת פונקציה מציינת *(Indicator Function)* והיא מ"א (הוכח זאת).

אם נתונים $A_n\in\mathcal{F}$ ו- α_n ממשיים אזי הפונקציה

$$X(\omega)=\sum_{n=1}^{N}\alpha_n I_{A_n}(\omega)$$

נקראת פונקציה פשוטה *(Elementary Function)* והיא מ"א.

טענה 2.12 אם $X_n(\omega)$ סדרת מ"א, ואם $X_n(\omega)\to X(\omega)$ לכל ω ו- $X(\omega)$ סופי לכל ω אזי $X(\omega)$ מ"א (יוכח בתרגיל הבית).

טענה 2.13 לכל מ"א $X(\omega)$ קיימת סדרה של מ"א פשוטים $X_n(\omega)$ כך ש- $X_n(\omega)\to X(\omega)$ לכל ω

הוכחה: נניח תחילה כי $X(\omega)\geq 0$. נגדיר

$$X_n(\omega)=\begin{cases}k/2^n & X(\omega)\in[k/2^n,\ (k+1)/2^n),\ k\geq 0\\ 2^n & X(\omega)\geq 2^n\end{cases}$$

לכן X_n מ"א פשוט, ו-$X_n(\omega)\leq X_{n+1}(\omega)\leq X(\omega)$. כעת, לכל ω קיים l כך ש-$X(\omega)\leq 2^l$. מההגדרה לעיל נובע כי

$$\sup_{n>l}|X_n(\omega)-X(\omega)|\leq 2^{-l}\ . \tag{2.1}$$

ומכאן $X_n\uparrow X$.

כעת ל- X כללי נרשום $X=X^+-X^-$ כאשר

$X^+(\omega) = \max(X(\omega), 0)$ ו- $X^-(\omega) = \max(-X(\omega), 0)$

ונקרב בנפרד את X^+ ואת X^- .

תרגיל 2.14 יהי Ω כלשהוא ו- $X(\omega) = \sum_{n=1}^{N} \alpha_n I_{A_n}(\omega)$ כאשר A_n קבוצות כלשהן ב- Ω ו- $\alpha_i \neq \alpha_j$ לכל $i \neq j$.

מהו ה- $\sigma - field$ המינימלי $\mathcal{F}$ כך ש- X הוא מ"א על $(\Omega, \mathcal{F})$?

שדה זה יסומן בד"כ ב- $\sigma(X)$

תרגיל 2.15 על המרחב $(\mathbb{R}, \mathbb{B})$ נגדיר מ"א:

$X(\omega) = \omega^2$. מהו $\sigma(X)$?

הגדרה 2.16 שני מ"א X ו- Y נקראים זהים או "שווים בהסתברות 1" אם

$$\mathbb{P}\left(X\left(\omega\right) \neq Y\left(\omega\right)\right) := \mathbb{P}\left(\left\{\omega : X\left(\omega\right) \neq Y\left(\omega\right)\right\}\right) \;=\; 0$$

מקובל לסמן זאת ע"י $X \;=\; Y \; a.s. \; (= almost \; sure)$

או $X \;=\; Y \; a.e. \; (= almost \; everywhere)$

או $X \;=\; Y \; w.p.1 \; (= with \; probability \; 1)$.

אנו נאמר שתכונה מסוימת (לדוגמא, $X(\omega) \geq 0$) מתקיימת עבור "כמעט כל ω " אם קבוצת ה- ω שעבורם התכונה אינה מתקיימת הסתברותה אפס. להבא לא נקפיד לכתוב $a.s.$ וכו', אך כל שויון יהיה במובן של $a.s.$.

טענה 2.17 יהיו X_n מ"א. אזי:

αX_n מ"א לכל α ממשי.

$X_1 + X_2$ מ"א.

$X_1 \cdot X_2$ מ"א.

אם g פונקציית בורל על $\mathbb{R}$ אזי $g\left(X\left(\omega\right)\right)$ מ"א (יוכח בתרגיל הבית).

וכן טענו לעיל כי גבול הסדרה (אם קיים) מגדיר מ"א.

המסקנה מהטענות לעיל היא כי ניתן לבנות מ"א ממשתנה נתון ע"י פעולות חיבור, כפל והרכבת פונקציות (בורל).

הגדרה 2.18 יהי $(\Omega, \mathcal{F})$ מרחב מדיד ו- $Y(\omega)$ מ"א. נגדיר

$$\mathcal{F}_Y = \sigma\left(\{\omega :\ Y\left(\omega\right) \leq \alpha\},\ -\infty < \alpha < \infty\right)$$

אזי בהגדרה $\mathcal{F}_Y \subset \mathcal{F}$. $\mathcal{F}_Y$ הוא *sub-σ-field* של $\mathcal{F}$.

לכל פונקצית בורל מתקיים $\mathcal{F}_{g(Y)} \subset \mathcal{F}_Y$ **לדוגמה,**

$\mathcal{F}_{Y^2} \subset \mathcal{F}_Y$, **אבל ייתכן** $\mathcal{F}_{Y^2} \neq \mathcal{F}_Y$ **(בדוק!).**

מסתבר שגם ההפך נכון:

משפט 2.19 *(Breiman, p.69)* אם $Z(\omega)$ מ"א על $(\Omega, \mathcal{F}_Y)$ (או במילים אחרות, Z מדיד על $\mathcal{F}_Y$), אזי קיימת פונקצית בורל $g(\Theta)$ כך ש- $Z(\omega) = g(Y(\omega))$.

יהיה $(\Omega, \mathcal{F}, \mathbb{P})$ **מרחב הסתברות, ויהיה** X **מ"א.**

הגדרה 2.20 $\mathbb{P}_X$ יקרא חוק הפילוג של X אם לכל קבוצת בורל (ממשית) B, מתקיים

$$\mathbb{P}(\omega : X(\omega) \in B) = \mathbb{P}_X(B)$$

אם למשתנה ערכים במרחב כללי S הפילוג יוגדר בצורה דומה, עבור קבוצות מדידות ב-σ-שדה על S.

דוגמה 2.21 הראה כי $(\mathbb{R}, \mathbb{B}, \mathbb{P}_X)$ הוא מרחב הסתברות; $\mathbb{R}$ הוא הישר הממשי $(-\infty, \infty)$, $\mathbb{B}$ שדה בורל על $\mathbb{R}$.

הערה 2.22 הגדרה זו נתנת להרחבה מיידית למ"א ב- $\mathbb{R}^{\mathbb{N}}$, ואף למרחבים כלליים יותר.

הגדרה 2.23 <u>פונקצית הפילוג</u> F_X של מ"א (ממשי) X מוגדרת ע"י

$$F_X(\alpha) = \mathbb{P}\left(\omega : X(\omega) \leq \alpha\right) = \mathbb{P}_X((-\infty, \alpha])$$

למ"א X יש <u>צפיפות</u> f_X , אם לכל α, $F_X(\alpha) = \int_{-\infty}^{\alpha} f(x)dx$.

הפילוג של ווקטור של מ"א, שהוא הפילוג המשותף של רכיבי ווקטור זה, מוגדר בצורה זהה. מ"א X,Y (עם ערכים במרחבים כלליים) נקראים בלתי תלויים סטטיסטית אם לכל זוג קבוצות מדי-דות A,B ההסתברות שהמ"א יקבלו ערכים בשניהם היא מכפלת ההסתברויות:

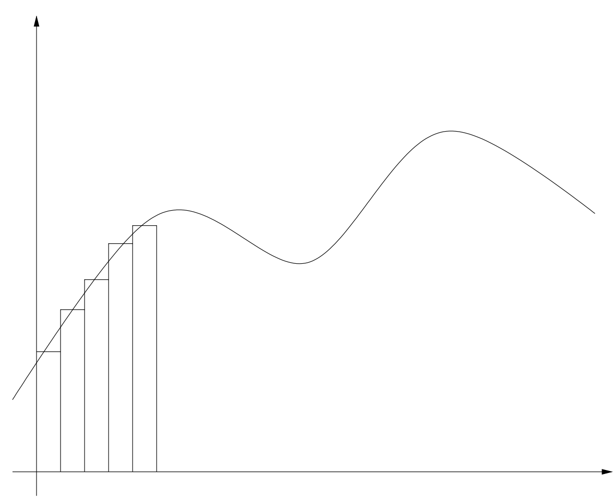
$$\mathbb{P}\left\{X \in A,\ Y \in B\right\} = \mathbb{P}\left\{X \in A\right\}\mathbb{P}\left\{Y \in B\right\} = \mathbb{P}_X(A)\,\mathbb{P}_Y(B) \tag{2.2}$$

הגדרה 2.24 שני סיגמה-שדות $\mathcal{F}_1, \mathcal{F}_2$ נקראים בת"ס אם כל משתנה אקראי שמדיד $\mathcal{F}_1$ בהכרח בת"ס בכל משתנה אקראי שמדיד $\mathcal{F}_2$.

תרגיל 2.25 הוכח כי אם X בת"ס ב-Y אזי $\mathcal{F}_X$ בת"ס ב-$\mathcal{F}_Y$.

2.3 תוחלת של משתנה אקראי (ממוצע)

ניזכר באינטגרל רימן, המוגדר על ידי קירובים על אינטרוולים, כלומר קירובים המוגדרים על תחום ההגדרה של הפונקציה-כמו באיור 2.1. ההגדרה בה נשתמש לתוחלת היא שונה. יהי $X(\omega)$



איור 2.1: אינטגרל רימן: קירוב על ידי חלוקה של תחום האינטגרציה

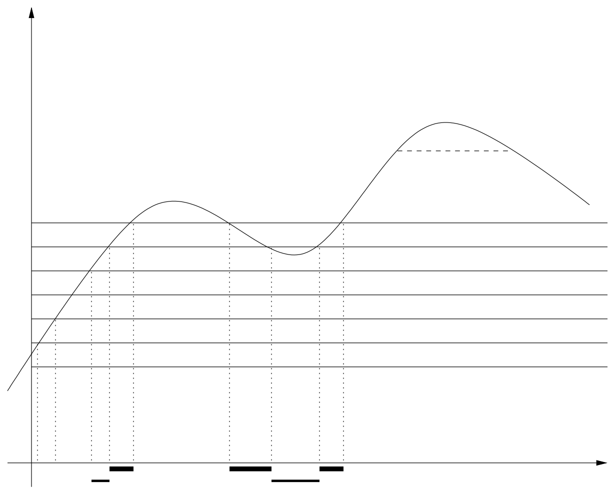
מ"א אי שלילי. נגדיר תוחלת ע"י

$$\mathbb{E}\,X(\omega) = \lim_{\delta\to 0}\lim_{N\to\infty}\sum_{n=1}^{N}(n-1)\delta\,\mathbb{P}\{\omega : (n-1)\delta \le X(\omega) < n\delta\}$$

(ראה איור 2.2) ונסמן זאת באחת מהצורות הבאות:

$$\mathbb{E}\, X = \int_\Omega X(\omega) P(d\omega) = \int_\Omega X(\omega) dP(\omega)$$

גבול זה קיים כיוון שצד ימין הוא פונקציה עולה כאשר N עולה וכאשר δ יורד. אינטגרל זה נקרא 'אינטגרל לבג' (Lebesgue) והוא שונה מאינטגרל רימן: שימו לב כי יתכן (ומוגדר היטב) המקרה $\mathbb{E}\, X = \infty$.



איור 2.2: אינטגרל לבג: קירוב על ידי הגדרת קבוצות בהן לפונקציה (מ"א) ערכים דומים

הקבוצות בעזרתן מוגדר אינטגרל זה אינן "אוסף של אינטרוולים": אין משמעות ל"אינטרוול" במרחב Ω.

תרגיל 2.26 הראה כי $\{\omega : (n-1)\delta \leq X(\omega) < n\delta\}$ הוא מאורע.

עבור פונקציה או מ"א כללי (שאינו בהכרח אי שלילי) נפרק $X = X^+ - X^-$ כאשר:

$$X^+(\omega) = \max(X(\omega), 0), \quad X^-(\omega) = \max(-X(\omega), 0)$$

ונגדיר: $\mathbb{E}\, X = \mathbb{E}\, X^+ - \mathbb{E}\, X^-$ בתנאי שלפחות אחד מהמחוברים שונה מ- ∞. בפרט, נובע מההגדרה כי אם $X(\omega) \leq 0$ אזי $\mathbb{E}\, X = -\,\mathbb{E}\, X^-$.

משתנה אקראי נקרא אינטגרבילי או בעל תוחלת סופית אם $\mathbb{E}\,|X| < \infty$.

תרגיל 2.27 חשב בעזרת ההגדרה הכללית תוחלת של פונקציה מציינת ושל מ"א פשוט.

הערה 2.28 ניתן לחשב את התוחלת דרך ההסתברות, הפילוג או פונקצית הפילוג: למשתנה ממשי

$$\mathbb{E}(X) = \int_\Omega X(\omega)\, d\,\mathbb{P}(\omega) = \int_{-\infty}^{\infty} x\, d\,\mathbb{P}_X(x) = \int_{-\infty}^{\infty} x\, dF_X(x) \tag{2.3}$$

ובצורה דומה, למשתנה שערכיו במרחב כללי S ולפונקציה ממשית g

$$\mathbb{E}(g(X)) = \int_\Omega g(X(\omega))\, d\,\mathbb{P}(\omega) = \int_S g(x)\, d\,\mathbb{P}_X(x) = \int_{-\infty}^{\infty} \alpha\, dF_{g(X)}(\alpha) \tag{2.4}$$

את האנטגרל כנגד פונקציית הפילוג אפשר לפרש בצורה שהגדרנו-כלומר שזוהי מידת הסתברות-או כאינטגרל סטליצ'ס-לבג, אותו לא נגדיר כאן.

תכונות התוחלת

יהיו X ו- Y מ"א אינטגרביליים ו-$\alpha,\ \beta$ ו-C קבועים.

א. לינאריות: $\mathbb{E}(\alpha X + \beta Y) = \alpha\,\mathbb{E}\,X + \beta\,\mathbb{E}\,Y$

ב. אם $X(\omega) = C$ **בהסתברות 1 אזי** $\mathbb{E}\,X = C$

ג. אם $X \geq Y$ **בהסתברות 1 אזי** $\mathbb{E}\,X \geq \mathbb{E}\,Y$

ד. אי שיויון ינסן (Jensen inequality)**: אם** g **פונקציה קונווקסית אזי** $\mathbb{E}[g(X)] \geq g(\mathbb{E}\,X)$

ה. אי שוויון הולדר Holder**: אם** $0 < p, 0 < q$ **ו-** $\frac{1}{p} + \frac{1}{q} = 1$ **אזי**

$$\mathbb{E}\,|XY| \leq ||X||_p\,||Y||_q, \qquad ||X||_p \doteq (\mathbb{E}\,|X|^p)^{1/p} \tag{2.5}$$

ו. אי שוויון קושי שוורץ הוא מקרה פרטי בו $q = p = 2$ **ואז מתקבל**

$$\mathbb{E}\,|XY| \leq ||X||_2\,||Y||_2 \qquad (\mathbb{E}\,|XY|)^2 \leq \mathbb{E}\,X^2\,\mathbb{E}\,Y^2 \tag{2.6}$$

ז. אי שוויון צ'בישב: עבור פונקציה ממשית וחיובית ϕ וקבוצה (בורל) A

$$\inf\{\phi(y) : y \in A\} \cdot \mathbb{P}\{X \in A\} \leq \mathbb{E}\left[\phi(X)\mathbf{1}_{\{X\in A\}}\right] \leq \mathbb{E}\,\phi(X) \tag{2.7}$$

הוכחה:

$$\inf\{\phi(y) : y \in A\}\mathbf{1}_{\{X\in A\}} \leq \phi(X)\mathbf{1}_{\{X\in A\}} \leq \phi(X) \tag{2.8}$$

וכעת ניקח תוחלת.

כמקרה פרטי נקח $\phi(x) = x^2$ ו-$A = \{x : |x| \geq a\}$ ונקבל את אי שוויון צ'בישב

$$a^2 \, \mathbb{P}\{|X| \geq a\} \leq \mathbb{E}\, X^2 \tag{2.9}$$

תרגיל 2.29 הוכח תכונות א-ג במקרה ש- X ו- Y מ"א פשוטים.

התכונה הבאה היא תוצאה של ההגדרה, אך יכולה לשמש כהגדרה חלופית. משתנים ממשיים X, Y נקראים בלתי תלויים סטטיסטית (בת"ס) אם לכל θ_1, θ_2

$$\mathbb{E}\, e^{i(\theta_1 X + \theta_2 Y)} = \mathbb{E}\, e^{i\theta_1 X}\, \mathbb{E}\, e^{i\theta_2 Y} \ . \tag{2.10}$$

מכאן נובעת תכונה דומה עבור פונקציות כלליות יותר, והיא משתמשת להגדרת אי תלות עבור משתנים עם ערכים כלשהם. המשתנים X, Y נקראים בלתי תלויים סטטיסטית (בת"ס) אם לכל זוג פונקציות ממשיות f, g

$$\mathbb{E}\, f(X)g(Y) = \mathbb{E}\, f(X)\, \mathbb{E}\, g(Y) \ . \tag{2.11}$$

2.4 מושגי התכנסות למ"א

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, ויהיו X_n מ"א.

הגדרה 2.30 הסדרה X_n מתכנסת בהסתברות 1 ל- X אם קיים מאורע A עם $\mathbb{P}(A) = 1$ כך ש- $X_n(\omega) \to X(\omega)$ לכל $\omega \in A$. נסמן זאת $X_n \overset{a.s.}{\to} X$.

הערה 2.31 לא הנחנו כי X הוא מ"א! אולם תמיד קיים מ"א Y כך שהקבוצה $N = \{\omega : X(\omega) \neq Y(\omega)\}$ שאיננה בהכרח מאורע, מקיימת $N \subset B$ כאשר B מאורע ו- $\mathbb{P}(B) = 0$, וכן $X_n \overset{a.s.}{\to} Y$. אנו נבחר תמיד Y כזה, ולכן נניח תמיד שהגבול הוא מ"א.

דוגמה 2.32 יהי $\Omega = \{1, 2, 3\}$, $\mathcal{F} = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$, $\mathbb{P}(\{1\}) = 1$. תהי $A = \{1\}$ ו- $X_n(1) = 1$, ונגדיר $X(\omega) = \omega$. אזי X איננו מ"א, ולכל $\omega \in A$, $X_n(\omega) = X(\omega)$. המשתנה האקראי $Y(\omega) = 1$ מקיים את התכונות שבהערה לעיל. ברור כי בעיה כזו לא תתעורר אם $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות <u>שלם</u>, כלומר $\mathcal{F}$ מכיל כל תת קבוצה של כל מאורע שהסתברותו אפס.

הגדרה 2.33 סדרה X_n היא סדרת קושי או סדרה מתכנסת הדדית בהסתברות 1 אם המ"א $a_n(\omega) = \sup_{i,j\geq n} |X_i(\omega) - X_j(\omega)|$ מקיימים $a_n \xrightarrow{a.s.} 0$.

הערה 2.34 a_n הוא מ"א (הוכח!).

לפי קריטריון קושי לסדרות ממשיות, לכל ω (פרט למאורע שהסתברותו אפס) קיים הגבול $X_n(\omega) \to X(\omega)$, ומהטיעון לעיל נוכל להניח ש- X מ"א.

מסקנה: סדרה X_n מתכנסת a.s. אם ורק אם היא מתכנסת הדדית a.s.

הגדרה 2.35 $X_n \xrightarrow{P} X$, הסדרה מתכנסת בהסתברות אם לכל $\epsilon > 0$, $\mathbb{P}\{|X_n - X| > \epsilon\} \to 0$ כאשר $n \to \infty$. X_n מתכנס הדדית בהסתברות אם לכל $\epsilon > 0$,

$$\sup_{i,j\geq n} \mathbb{P}\{|X_i - X_j| > \epsilon\} \to 0 \quad as\ n \to \infty\ .$$

טענה 2.36

א. X_n מתכנסת בהסתברות אם"ם היא מתכנסת הדדית בהסתברות.

ב. אם $X_n \xrightarrow{a.s.} X$ אזי $X_n \xrightarrow{P} X$.

ג. אם $X_n \xrightarrow{P} X$ אזי קיימת תת סדרה n_i עבורה $X_{n_i} \xrightarrow{a.s.} X$.

תרגיל 2.37

1. הוכח את ב לעיל.

2. תן דוגמא נגדית לטענה: $X_n \xrightarrow{P} X$ גורר $X_n \xrightarrow{a.s.} X$. רמז:

יהי $\Omega = [0,1]$ ועבור $0 \leq a < b \leq 1$ יהי $\mathbb{P}([a,b]) = b - a$.

נגדיר $t_n = \sum_{i=2}^n 1/i$. נסמן ב- $[s]$ את השלם הגדול ביותר הקטן מ- s. יהי

$$X_n(\omega) = \begin{cases} 1 & \omega = s - [s], s \in [t_n t_{n+1}] \\ 0 & otherwise \end{cases}$$

כדי להוכיח את ג. אנו זקוקים ללמה חשובה וידועה. נעזר בתרגיל הבא:

תרגיל 2.38 עבור סדרות מאורעות A_k נגדיר $B_n = \bigcup_{k=n}^{\infty} A_k$, ו- B_n $A^{\infty} = \bigcap_{n=1}^{\infty} B_n$. הוכח:

$A^{\infty} \ni \omega$ אם ורק אם ω שייך ל- A_k אין סוף פעמים.

למה 2.39 בורל קנטלי *(Borel Cantelli)*: בסימונים של התרגיל לעיל, אם $\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty$ אזי $\mathbb{P}(A^{\infty}) = 0$, כלומר הסיכוי שקרו אין סוף A_k הוא אפס.

הוכחה: לכל n - $\mathbb{P}(A_{\infty}) \leq \mathbb{P}(B_n) \leq \sum_{k=n}^{\infty} \mathbb{P}(A_k)$. **אבל צד ימין שואף ל-0 כאשר** $n \to \infty$ **תחת ההנחה.**

הערה 2.40 החצי השני של הלמה הוא: אם A_i בת"ס ו- $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ אזי ההסתברות כי A_i קורה אין סוף פעמים היא 1, כלומר: $\mathbb{P}(A^{\infty}) = 1$.

<u>**הוכחת סעיף ג'**</u> **בטענה** 2.36: **נבחר סדרה כלשהי** $a_i \to 0$. **נגדיר תת סדרה** n_i **על ידי הדרישה** $\mathbb{P}(\omega : |X_{n_i}(\omega) - X| > a_i) < 2^{-i}$. **הגדרה זו אפשרית בגלל ההתכנסות בהסתברות. נגדיר** $A_i = \{\omega : |X_{n_i}(\omega) - X| > a_i\}$. **אזי** $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \sum_{i=1}^{\infty} 2^{-i} = 1 < \infty$ **ולכן** $\mathbb{P}(A^{\infty}) = 0$, **כלומר** $\mathbb{P}(|X_{n_i}(\omega) - X| > a_i \; i.o.) = 0$ **כאשר** i.o. = infinitely often. **לכן, בהסתברות 1,** $|X_{n_i}(\omega) - X| > a_i$ **מספר סופי של פעמים ומכיוון ש** $a_i \to 0$ **נסיק כי** $X_{n_i}(\omega) \xrightarrow{a.s.} X(\omega)$.

תרגיל 2.41 הוכח את א. רמז: להוכחת הדדית -> התכנסות, מצא תת סדרה מתכנסת $a.s.$ כמו בהוכחת ג, ולכן קיים X. הוכח התכנסות.

<u>**התכנסות בממוצע**</u>

התכנסות בממוצע והקשר שלה להתכנסויות אחרות הם מהכלים המרכזים בהם נשתמש בה-משך.

הגדרה 2.42 עבור $1 \leq r < \infty$, המרחב $L^r(\Omega, \mathcal{F}, \mathbb{P})$ הוא אוסף המ"א על $(\Omega,\ \mathcal{F})$ כך ש-$\mathbb{E}(|X|^r) < \infty$.

נשתמש בסימון המקוצר: L^r.

הגדרה 2.43 סדרה $\{X_n\}$ מתכנסת בממוצע מסדר r למ"א X אם X_n, X שייכים ל- L^r ו- $\mathbb{E}\,|X_n - X|^r \to 0$.

נסמן זאת על ידי $X_n \xrightarrow{r.m.} X$. המקרה $r = 2$ נקרא התכנסות בממוצע ריבועי $q.m. = quadratic\ mean$.

טענה 2.44 [ללא הוכחה] המרחב $L^r(\Omega, \mathcal{F}, \mathbb{P})$ הוא מרחב שלם תחת הנורמה $||X||_r = (E|X|^r)^{1/r}$, עבור כל $1 \leq r < \infty$. זהו מרחב הילברט כאשר $r = 2$ כאשר המכפלה הפנימית היא $(X, Y) = \mathbb{E}\,XY$. נסמן $H = L^2(\Omega, \mathcal{F}, \mathbb{P})$. יהי $\mathcal{F}_1$ *sub sigma-field* של $\mathcal{F}$, ויהי H_1 אוסף המ"א המדידים על $\mathcal{F}_1$ ומקיימים $\mathbb{E}\,X^2 < \infty$. אזי H ו- H_1 מרחבים וקטוריים לינאריים, וכן הם מרחבי הילברט, כאשר H_1 תת מרחב הילברט של H.

משפט 2.45 [אי שוויון מרקוב] תהי f פונקציה חיובית לא יורדת (בורל). אזי לכל מ"א X:

$$\mathbb{P}\left\{|X(\omega)| \geq \varepsilon\right\} \leq \frac{1}{f(\varepsilon)} \cdot \mathbb{E}(f(|X|))$$

הוכחה: זהו מקרה פרטי של אי שוויון צ'בישב המוכלל, אולם נציג הוכחה אחרת.

$$\mathbb{E}\,f(|X|) = \int_\Omega f(|X(\omega)|)\,\mathbb{P}(d\omega) \tag{2.12}$$

$$\geq \int_{\{\omega\ |\ |X)| \geq \varepsilon\}} f(|X(\omega)|)\,\mathbb{P}(d\omega) \quad \text{כיוון ש-} f \text{ חיובית} \tag{2.13}$$

$$\geq \int_{\{\omega\ |\ |X)| \geq \varepsilon\}} f(\varepsilon)\,\mathbb{P}(d\omega) \quad \text{כיוון ש-} f \text{ לא יורדת} \tag{2.14}$$

$$= f(\varepsilon)\,\mathbb{E}\,\mathbf{1}_{\{|X| \geq \varepsilon\}} \tag{2.15}$$

$$= f(\varepsilon)\,\mathbb{P}(|X| \geq \varepsilon)\ . \tag{2.16}$$

דוגמה 2.46 הפונקציה $f(X) = |X|^r$ עבור $r \geq 1$ מקיימת תנאי המשפט.
לכן לכל $r \geq 1$, $\mathbb{P}\{|X| \geq \varepsilon\} \leq \varepsilon^{-r}\,\mathbb{E}(|X|^r)$.

משפט 2.47 אם $X_n \xrightarrow{r.m.} X$ אזי $X_n \xrightarrow{\mathbb{P}} X$.

הוכחה: מאי שוויון מרקוב

$$\mathbb{P}(|X_n - X| \geq \varepsilon) \leq \varepsilon^{-r}\,\mathbb{E}(|X_n - X|^r)$$

דוגמה להתכנסות (מ- Durret).

יהיו $\{X_n\}$ מ"א ב-L_2 שהם חסרי קורלציה, כלומר לכל $i \neq j$ מתקיים $\mathbb{E}\,X_iX_j = \mathbb{E}\,X_i\,\mathbb{E}\,X_j$. נניח שלמשתנים אלו תוחלת קבועה ווריאנס חסום, כלומר לכל n

$$\mathbb{E}\,X_n = \mu, \quad VarX_n \leq C\,. \tag{2.17}$$

אזי הסדרה מקיימת את חוק המספרים הגדולים במובן הבא: נסמן $S_n = X_1 + \ldots + X_n$. אזי כאשר $n \to \infty$

$$\frac{S_n}{n} \to \mu \tag{2.18}$$

ב-L_2 ובהסתברות.

הוכחה: מספיק להראות התכנסות ב-L_2. נחשב

$$\mathbb{E}\left(\frac{S_n}{n} - \mu\right)^2 = Var\left(\frac{S_n}{n}\right) \tag{2.19}$$

$$= \frac{1}{n^2}VarS_n \tag{2.20}$$

$$= \frac{1}{n^2}\left(VarX_1 + \ldots + VarX_n\right) \tag{2.21}$$

$$\leq \frac{Cn}{n^2} = \frac{C}{n} \tag{2.22}$$

כאשר השוויון הלפני אחרון נובע מכך שהמשתנים חסרי קורלציה.

תרגיל 2.48 הרחב את הדוגמה לעיל למקרה שהממוצע של המשתנים אינו קבוע; איזה תנאי על הממוצעים דרוש כדי להרחיב הוכחה זו?

התכנסות של חוקי הסתברות

יהיה $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, ויהיה X מ"א.

הגדרה 2.49 נאמר כי מ"א X_n מתכנסים בהתפלגות ל-X ונסמן זאת ב- $X_n \xrightarrow{L} X$ אם $F_{X_n}(\alpha) \to F_X(\alpha)$ לכל α שהיא נקודת רציפות של $F_X(\alpha)$.

2.50 הערות

א. כיוון שההגדרה היא דרך פונקציות פילוג, יש לה משמעות אפילו כאשר כל X_n (ו-X) מוגדר על מרחב הסתברות אחר!

ב. ההגדרה ניתנת להכללה למ"א עם טווח כללי; נסמן $X_n \xrightarrow{L} X$ אם לכל פונקציה רציפה וחסומה f מתקיים $\mathbb{E}\, f(X_n) \to \mathbb{E}\, f(X)$.

ג. ההגבלה ל"נקודת רציפות" מיועדת להבטיח שההכללה ב׳ תקיפה, וכן כדי שהתכנסות זו תהיה חלשה יותר מהתכנסות בהסתברות (ראה משפט).

ד. באופן <u>שקול</u> לב׳ - ניתן לדבר על <u>התכנסות</u> חלשה של פילוגים או של מידות הסתברות: נסמן $\mathbb{P}_n \Rightarrow \mathbb{P}$ אם לכל פונקציה רציפה וחסומה f, $\int_\Omega f(\alpha)\, d\,\mathbb{P}_n(\alpha) \to \int_\Omega f(\alpha)\, d\,\mathbb{P}(\alpha)$ (או, לפונקציות פילוג, $\int_\mathbb{R} f(\alpha)\, dF_n(\alpha) \to \int_\mathbb{R} f(\alpha)\, dF(\alpha)$).

כיון שבהסתברות זו חשוב רק הפילוג, אומרים לפעמים ש-X_n מתכנס חלש ל-X.

דוגמה 2.51 עבור n שלם נגדיר

$$F_X(\alpha) = \begin{cases} 0 & \alpha < 0 \\ 1 & \alpha \geq 0 \end{cases} \qquad F_{X_n}(\alpha) = \begin{cases} 0 & \alpha < 1/n \\ 1 & \alpha \geq 1/n \end{cases} \tag{2.23}$$

תרגיל 2.52 הראה שיש התכנסות חלשה בדוגמה. מתי יש גם התכנסות בהסתברות?

משפט 2.53 (לא הוכחה) אם $X_n \xrightarrow{\mathbb{P}} X$ אזי $X_n \xrightarrow{L} X$.

אם $X_n \xrightarrow{L} X$ ו-X דטרמניסטי (בהסתברות 1) אזי $X_n \xrightarrow{\mathbb{P}} X$.

להוכחת הטענה הראשונה ראה משפט 2.57.

דוגמאות להתכנסות חלשה.

נסמן התכנסות חלשה על ידי חץ כפול: $X_n \Rightarrow Y$.

1. יהיו $X_1, X_2, \ldots$ מ"א בת"ס ושווי פילוג i.i.d. עם פילוג

$$\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}. \tag{2.24}$$

נגדיר $S_n = X_1 + \cdots + X_n$. אזי ממשפט הגבול המרכזי

$$F_n(y) \doteq \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq y\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2}\, dx \tag{2.25}$$

עבור כל y. כיוון שצד ימין הוא פילוג והוא רציף ב-y נסיק מכך כי הפילוג של $S_n/\sqrt{n}$ מתכנס לפילוג נורמלי, או המ"א מתכנסים התכנסות חלשה למשתנה נורמלי.

2. מדידה ברעש. תהיה F_x פונקצית פילוג של מ"א X. נניח שמודדים את X בנוכחות רעש. את הרעש נייצג על ידי מ"א $Y_1, Y_2, \ldots$ i.i.d. כאשר המדידה ה-n תתן $X + \frac{1}{n}Y_n$. נניח שהרעש חסום כלומר $|Y_i| \leq C$. אזי פונקצית הפילוג של המדידה ה- n מקיימת

$$F_n(x) \doteq \mathbb{P}\left(X + \frac{1}{n}Y_n \leq x\right) \tag{2.26}$$

$$\leq \mathbb{P}\left(X \leq x + \frac{C}{n}\right) \tag{2.27}$$

$$F_n(x) \geq \mathbb{P}\left(X \leq x - \frac{C}{n}\right)\,. \tag{2.28}$$

אם F_x רציף בנקודה x אזי מהשוויון למעלה נובע כי

$$F_n(x) \to \mathbb{P}(X \leq x) = F_x(x) \tag{2.29}$$

ולכן יש התכנסות בהתפלגות או התכנסות חלשה של המ"א.

3. זמן עד מאורע נדיר.

נסמן ב-X_p את המ"א הסופר את מספר הנסיונות עד הצלחה של ניסוי עם הסתברות הצלחה p בסדרת נסיונות i.i.d. למשל, זריקת מטבע עם סיכוי p לעץ, כאשר X_p סופר את מספר הפלי עד העץ הראשון. אזי הפילוג הוא

$$\mathbb{P}\left(X_p > n\right) = (1-p)^n\,. \tag{2.30}$$

נתבונן במ"א pX_p. כאשר $p \to 0$ נקבל

$$\mathbb{P}\left(pX_p > x\right) = \mathbb{P}\left(X_p > x/p\right) \tag{2.31}$$

$$= (1-p)^{x/p} \tag{2.32}$$

$$\to e^{-x} \tag{2.33}$$

כאשר $p \to 0$ לכל x. מכאן שיש התכנסות חלשה והגבול הוא משתנה אקספוננציאלי.

משפט 2.54 *Skorohod representation* נניח כי $X_n \Rightarrow X$. אזי קיים מרחב הסתברות עם מ"א Y, Y_n כך ש-

$$F_X = F_Y, \quad F_{X_n} = F_{Y_n} \tag{2.34}$$

ובנוסף $Y_n \to Y$ *w.p.1*.

הערה 2.55 חשוב לזכור כי במעבר מ-X ל-Y הפילוגים הרב ממדיים משתנים: כלומר, בדרך כלל

$$\mathbb{P}(X_n \in A,\ X_{n+1} \in B) \neq \mathbb{P}(Y_n \in A,\ Y_{n+1} \in B).$$

זהו רק אחד מהמשפטים הנקראים על שם Skorohod. משפט זה תקף במרחבים כלליים (לא רק למ"א ממשיים), ונותן כלי נוח להוכחות הקשורות בהתכנסות חלשה. למשל

משפט 2.56 $X_n \Rightarrow X$ אמ"מ לכל פונקציה רציפה וחסומה $\mathbb{E}\, g(X_n) \to \mathbb{E}\, g(X)$. בנוסף לכל f רציפה מתקיים $f(X_n) \Rightarrow f(X)$.

הוכחה: נניח $X_n \Rightarrow X$. נשתמש בייצוג של סקורוחוד ונקבל משתנים Y, Y_n בעלי פילוג זהה המתכנסים בהסתברות 1. כיוון ש- g רציפה, נובע כי $g(Y_n) \to g(Y)$ בהסתברות 1, ולכן ממשפט ההתכנסות החסומה $\mathbb{E}\, g(Y_n) \to \mathbb{E}\, g(Y)$. כיוון שהפילוגים של X ושל Y זהים, נובע כי $\mathbb{E}\, g(X_n) \to \mathbb{E}\, g(X)$.

את הכיוון ההפוך לא נוכיח: הוא נעשה על ידי קירוב הפונקציה המציינת על ידי פונקציות רציפות.

להוכחת הטענה האחרונה נשים לב כי לכל פונקציה רציפה וחסומה g הפונקציה $h(x) = g[f(x)]$ היא רציפה וחסומה. מהדיון הקודם נקבל מייד כי $f(Y_n) \to f(Y)$ בהסתברות 1 ולכן $\mathbb{E}\, g[f(X_n)] \to \mathbb{E}\, g[f(X)]$ ולכן לפי חציו הראשון של המשפט $f(X_n) \Rightarrow f(X)$.

לסיכום נתאר (ללא הוכחה) כמה דרכים לבדיקת התכנסות חלשה:

משפט 2.57 הטענות הבאות שקולות:

1. $\mathbb{P}_{X_n} \Rightarrow \mathbb{P}_X$
2. לכל קבוצה פתוחה G

$$\liminf_n \mathbb{P}(X_n \in G) \geq \mathbb{P}(X \in G) \tag{2.35}$$

3. לכל קבוצה סגורה K

$$\limsup_n \mathbb{P}(X_n \in K) \leq \mathbb{P}(X \in K) \tag{2.36}$$

4. לכל קבוצה מדידה A כך ש- $\mathbb{P}(X \in \partial A) = 0$

$$\lim_n \mathbb{P}(X_n \in A) = \mathbb{P}(X \in A) \tag{2.37}$$

הערה: $\partial A \doteq A^c - A^o$ **כלומר כל הנקודות בסגור של הקבוצה שאינן בתוך קבוצה פתוחה שבתוך** A.

<u>סיכום והקשרים בין התכנסויות</u>

ראינו כי התכנסויות $a.s. \Leftarrow \mathbb{P} \Rightarrow r.m.$, **ובדוגמה נגדית** $a.s. \not=> \mathbb{P}$.

תרגיל 2.58 הראה ע"י דוגמה כי $r.m. \not=> a.s.$ וכן $a.s. \not=> r.m.$ לכל $r \geq 1$.
<u>רמז</u> לחלק הראשון, יהיו $A_n, n = 1, 2, \ldots,$ מאורעות כך ש- $\mathbb{P}(A_n) > 0$ ו- $A_i \bigcap A_j = \phi$ לכל $i \neq j$.
הגדר $X_n = \alpha_n \mathbf{1}_{\{A_n\}}$. הראה שניתן לבחור את α_n כך ש- $\mathbb{E}\,|X_n|^r = C$ קבוע, X_n מתכנס בהסתברות 1, אך לא במובן $r.m$. הכוון השני - השתמש בדוגמה $a.s. \not=> \mathbb{P}$.

מסתבר שדרוש תנאי מעט חזק מ- $\mathbb{E}\,|X_n|^r \leq C$ **כדי לקבל** $r.m. \Rightarrow a.s.$.

הגדרה 2.59 קבוצת מ"א $\{X_\alpha, \alpha \in I\}$ היא *(UI) Uniformly Integrable* אם

$$\lim_{N\to\infty} \sup_\alpha \mathbb{E}(|X_\alpha| I_{\{|X_\alpha|>N\}}) = 0$$

הערה 2.60 התנאי בא למנוע "בריחה" של מסת ההסתברות לאינסוף.

תרגיל 2.61 :

א. הראה כי X_n מהתרגיל הקודם (עבור $r = 1$) אינם מקיימים תנאי זה.

ב. הראה כי אם קיימת פונקציה $f(x) \geq 0$ כך ש- $\frac{f(x)}{|x|} \to \infty$ כאשר $|x| \to \infty$ ו- $\mathbb{E}(f(X_n)) \leq C$ אזי $\{X_n\}$ היא *uniformly integrable*.

משפט 2.62 (ללא הוכחה) אם $X_n \xrightarrow{P} X$ ו- $\{|X_n|^r\}$ היא UI אזי $X_n \xrightarrow{r.m.} X$.
אם $X_n \xrightarrow{r.m.} X$ אזי $\{|X_n|^r\}$ היא UI.

שאלה טכנית שתעלה פעמים רבות בהמשך היא - החלפת סדר תוחלת וגבול, כלומר מתי $\mathbb{E} \lim X_n = \lim \mathbb{E} X_n$. נתן מספר משפטים ללא הוכחות.

הגדרה 2.63 נסמן $\underline{\lim} X_n = X$ או $\liminf X_n = X$ אם מתקיים $\lim_{n\to\infty} \inf_{k\geq n} X_k = X$.
הגדרת $\overline{\lim}$ או $\limsup$ מקבילה.

לדוגמה, הסדרה $a_n = (-1)^n$ אינה מתכנסת, אך

$$\liminf a_n = -1, \quad \limsup a_n = 1.$$

קל להראות כי לסידרה $\{b_n\}$ יש גבול אם ורק אם $\liminf b_n = \limsup b_n$.

משפט 2.64 [הלמה של *Fatou*] נניח שקיים $X(\omega)$ עבורו $\mathbb{E}|X| < \infty$, ולכל n, $X_n(\omega) \geq X(\omega)$, אזי

$$\liminf_{n\to\infty} \mathbb{E} X_n \geq \mathbb{E} \liminf_{n\to\infty} X_n$$

מקרה פרטי אך נפוץ הוא כאשר המשתנים X_n הם לא שליליים, ואז התנאי מתקיים עבור $X \equiv 0$.

משפט 2.65 [$Dominated\ Convergence$] אם קיים Y עבורו $\mathbb{E}\, Y < \infty$ ולכל n, $|X_n| \le Y$ ואם $X_n \xrightarrow{\mathbb{P}} X$ אזי $\mathbb{E}\, X_n \to \mathbb{E}\, X$.

תרגיל 2.66 [$Bounded\ Convergence$] אם $|X_n| \le C$ ו- $X_n \xrightarrow{\mathbb{P}} X$ אזי $\mathbb{E}\, X_n \to \mathbb{E}\, X$.

משפט 2.67 [$Monotone\ Convergence$] אם $X_n \ge 0$, $X_n \uparrow X$ אזי $\mathbb{E}\, X_n \uparrow \mathbb{E}\, X$.

תרגיל 2.68 אם $X_n \xrightarrow{r.m.} X$ אזי $\mathbb{E}\,|X_n|^r \to \mathbb{E}\,|X|^r$

<u>רמז</u> $|X_n| = |X_n - X + X|$ והעזר באי שויון המשולש.

תרגיל 2.69 יהי $\Omega = [0,1]$, $\mathcal{F} = \mathbb{B}$ שדה בורל ו- $\mathbb{P}(a,b) = b - q$ כאשר $a < b$. נגדיר $X_n = \begin{cases} n & \omega \le 1/n \\ 0 & \text{אחרת} \end{cases}$. הראה כי X_n מתכנס (באיזה מובן?). אלו מהמשפטים ניתן לישם עבור $\mathbb{E}\, X_n$?

תרגיל 2.70 נתונה סדרה $\{X_n\}$. נגדיר $Y_n = \left(\frac{X_n}{1+|X_n|}\right)$. אזי $Y_n \in (-1,1)$

הוכח: $X_n \xrightarrow{\mathbb{P}} X$ אם ורק אם $Y_n \xrightarrow{r} Y$ עבור $Y \in (-1,1)$ ועבור $r \ge 1$ כלשהו.

חשיבותם של מושגי ומשפטי ההתכנסות קשורים בעיקר לקרובים. כך למשל, נראה בהמשך כיצד לתת הגדרה מדויקת לתוחלת מותנית, עבור מ"א בעלי מומנט שני. כדי להרחיב את ההגדרה למ"א בעלי תוחלת (אך לא בהכרח מומנט שני) נקרב מ"א כאלו ע"י סדרת מ"א בעלי מומנט שני. בעזרת משפטי ההתכנסות נבסס את ההגדרה המורחבת.

תרגיל 2.71 יהי X מ"א עם צפיפות f_x. הוכח כי אם X אינטגרבילי,

$$\mathbb{E}\, X = \int_{-\infty}^{\infty} y f_x(y) dy$$

<u>רמז</u> רשום את קרובי רימן לצד ימין של המשוואה והשתמש במשפטי התכנסות עבור תוחלות.

תרגיל זה מראה שהגדרת האינטגרל שבחרנו מכסה את אינטגרל רימן.

2.5 תוחלת מותנית

דוגמה 2.72 :

$$\xrightarrow[signal]{X(t)} \boxed{\text{מערכת מדידה}} \xrightarrow[measured\quad signal]{Y(t)}$$

ידועה הסטטיסטיקה של $X(t)$. מהו למשל הסכוי ש- $X(t) \geq 8$ <u>בהנתן</u> מדידה $Y(t) = 17$? מושג התוחלת המותנית בא לענות על שאלות מסוג זה. למשל, X מ"א המתאר זריקת קוביה.

$$Y = \begin{cases} 1 & X = 1, 3, 5 \\ 0 & X = 2, 4, 6 \end{cases}$$

לפי חוק בייס

$$\mathbb{P}\left[X = 3 \mid Y = 1\right] = \mathbb{P}\left[X = 3 \mid \text{אי זוגי } X\right] = \frac{\mathbb{P}(X = 3, \text{אי זוגי } X)}{\mathbb{P}(\text{אי זוגי } X)} \tag{2.38}$$

$$= \frac{\mathbb{P}(X = 3)}{\mathbb{P}(\text{אי זוגי } X)} = \frac{1/6}{1/2} = \frac{1}{3} \tag{2.39}$$

וכמובן

$$\mathbb{P}\left[X = 3 \mid Y = 0\right] = \mathbb{P}\left[X = 3 \mid \text{זוגי } X\right] = \frac{0}{1/2} = 0 \tag{2.40}$$

ברישום מקוצר

$$\mathbb{P}\left[X = 3 \mid Y\right] = \begin{cases} 1/3 & Y = 1 \\ 0 & Y = 0 \end{cases}$$

זהו מ"א! שים לב כי

$$\mathbb{P}\left[X = 3 \mid Y\right] = \mathbb{P}\left[X = 3 \mid Y = 1\right] \mathbf{1}_{\{Y=1\}} + \mathbb{P}\left[X = 3 \mid Y = 0\right] \mathbf{1}_{\{Y=0\}}$$

באנלוגיה לקשר

$$\mathbb{P}(A) = \mathbb{E}\, \mathbf{1}_{\{A\}}$$

נקבל עבור המקרה המותנה

$$\mathbb{P}\left[X = 3 \mid Y\right] = \mathbb{E}\left[\mathbf{1}_{\{X=3\}} \,\middle|\, Y\right]$$

וזאת נכליל <u>לתוחלת מותנית</u>: עבור פונקציה f כלשהיא (בורל):

$$\mathbb{E}\left[f(X) \mid Y\right] = \mathbb{E}[f(X)\mathbf{1}_{\{Y=1\}}]\left(\frac{\mathbf{1}_{\{Y=1\}}}{\mathbb{E}\,\mathbf{1}_{\{Y=1\}}}\right) + \mathbb{E}[f(X)\mathbf{1}_{\{Y=0\}}]\left(\frac{\mathbf{1}_{\{Y=0\}}}{\mathbb{E}\,\mathbf{1}_{\{Y=0\}}}\right)$$

<u>המשמעות:</u> בהנתן מאורע המוגדר ע"י Y, נקבל ממוצע של $f(X)$ על פני מאורע זה. התוצאה היא מ"א המדיד $\mathcal{F}_Y$!

ניתן להגדיר כך תוחלת מותנית בתנאי ש- $\mathbb{P}\{Y = 1\} \neq 0$.

מה קורה למשתנים בעלי צפיפות?

<u>הגדרת תוחלת מותנית</u>

יהי X **מ"א ב-** $L^2(\Omega, \mathcal{F}, \mathbb{P})$ **, ויהי** Y **מ"א.**

יהי H_Y **תת-מרחב הילברט של משתנים המדידים** $\mathcal{F}_Y$, **ונמצאים ב-** $L^2(\Omega, \mathcal{F}, \mathbb{P})$. (**<u>לא</u> דרשנו** $\mathbb{E}\,Y^2 < \infty$)! **טענו בעבר כי כל מ"א ב-** H_Y **ניתן לרשום כ-** $g(Y)$ **עבור פונקצית בורל כלשהי** g. **לכן ניתן ליצג**

$$H_Y = \{g(Y) : \mathbb{E}[g(Y)]^2 < \infty \quad \text{בורל } g\}$$

הגדרה 2.73 <u>התוחלת המותנת</u> $\mathbb{E}\left[X \mid Y\right]$ היא ההשלכה של X על H_Y.

הערה 2.74 הוכחנו שהשלכה כזו היא יחידה, ומקיימת לכל $Z \in H_Y$

$$\mathbb{E}[(X - \mathbb{E}\left[X \mid Y\right]) \cdot Z] = 0 \tag{2.41}$$

או, בצורה שקולה, לכל פונקצית בורל g,

$$\mathbb{E}[(X - \mathbb{E}\left[X \mid Y\right]) \cdot g(Y)] = 0$$

כאשר זהו <u>איפיון</u> של $\mathbb{E}\left[X \mid Y\right]$. כלומר $\mathbb{E}\left[X \mid Y\right]$ הוא מ"א מדיד $\mathcal{F}_Y$ המקיים את (2.41) , ואם גם $\widetilde{X}$ מקיים את (2.41) אזי $\mathbb{P}(\widetilde{X} \neq \mathbb{E}\left[X \mid Y\right]) = 0$.

שים לב כי $\mathbb{E}\left[X \mid Y\right]$ **תלוי ב-** Y **רק דרך** $\mathcal{F}_Y$. **לכן, אם** f **היא פונקציה חד-חד ערכית, אזי** $\mathbb{E}\left[X \mid Y\right] = \mathbb{E}\left[X \mid f(Y)\right]$.

הערה 2.75 אנו נשתמש בהתניה רק ב- σ -שדות. כמובן שאם $\mathbb{E}[X \mid Y=3]$ מוגדר היטב (כלומר ההתניה היא במאורע שהסתברותו אינה אפס) אזי $\mathbb{E}[X \mid Y=3] \neq \mathbb{E}[X \mid f(Y)=3]$ כי כמובן, $\mathbb{E}[X \mid Y=3]=\mathbb{E}[X \mid f(Y)=f(3)]$ אם f הפיכה.

כיון שיש תלות רק ב- $\mathcal{F}_Y$, ניתן להגדיר תוחלת מותנית ב- σ-field באותה צורה. בהנתן $\mathcal{F}_1$ שהוא sub σ-field של $\mathcal{F}$, יהי H_1 תת מרחב הילברט של מ"א מדידים $\mathcal{F}_1$. אזי $\mathbb{E}[X \mid \mathcal{F}_1]$ הוא המ"א היחיד המקיים (2.41) ומדיד $\mathcal{F}_1$.

טענה 2.76 (ללא הוכחה) (2.41) מתקיים אם ורק אם לכל מאורע $A \in \mathcal{F}_Y$,

$$\mathbb{E}[(X-\mathbb{E}[X \mid Y]) \cdot \mathbf{1}_{\{A\}}]=0$$

<u>**הוכחה**</u> מיידית למ"א פשוט, ואח"כ נקרב.

הגדרה 2.77 יהי X מ"א בעל תוחלת סופית $\mathbb{E}|X|<\infty$.
אזי התוחלת המותנית $\mathbb{E}[X \mid \mathcal{F}_1]$ מוגדרת כמ"א המדיד $\mathcal{F}_1$ ומקיים, לכל $A \in \mathcal{F}_1$

$$\mathbb{E}[(X-\mathbb{E}[X \mid \mathcal{F}_1]) \cdot \mathbf{1}_{\{A\}}]=0$$

בצורה שקולה,

$$\int_A [X(\omega)-\mathbb{E}[X \mid \mathcal{F}_1](\omega)]\, \mathbb{P}(d\omega)=0$$

כאן לא דרשנו $\mathbb{E}|X^2|<\infty$, אלא רק $\mathbb{E}|X|<\infty$. לכן, נשאלת השאלה מנין שקיים כזה מ"א, ומנין שהוא יחיד? כאן אין מרחב הילברט!

טענה 2.78 (הוכחה בתרגיל בהמשך) אם $\mathbb{E}|X|<\infty$ אזי $\mathbb{E}[X \mid \mathcal{F}_1]$ קיים ויחיד בהסתברות 1.

ההוכחה קשורה לאפשרות לקרב מ"א על ידי משתנים בעלי מומנט שני. כדי להוכיח קירוב זה, נוח תחילה להוכיח תכונות תוחלת מותנית עבור משתנים בעלי מומנט שני, ורק אחר כך להרחיב את ההגדרה וכן את התכונות למשתנים בעלי מומנט ראשון בלבד.

הגדרה 2.79 הסתברות מותנת

$$\mathbb{P}[A \mid \mathcal{F}_1]=\mathbb{E}\left[\mathbf{1}_{\{A\}} \mid \mathcal{F}_1\right]$$

הסימון $\mathbb{E}[X \mid Y]$ יובן מעתה כ- $\mathbb{E}[X \mid \mathcal{F}_Y]$.

תרגיל 2.80 יהי Y מ"א המקבל ערכים מ- $1,2,\ldots,N$ ו- X מ"א כך ש- $\mathbb{E}\,|X| < \infty$.

א. הוכח כי $\mathbb{E}\,[X \mid Y]$ קבוע על כל קבוצה מהצורה $\{\omega \mid Y(\omega)=j\}$.

ב. תן יצוג מפורש ל- $\mathbb{E}\,[X \mid Y]$ (ראה דוגמת הקוביה).

ג. הראה כי (2.41) אינו חייב להתקיים אם Z אינו מדיד $\mathcal{F}_Y$.

ד. יהי $\mathcal{F}_1 = \{\Omega,\phi\}$. חשב את $\mathbb{E}\,[X \mid \mathcal{F}_1]$.

לפני שנפתח את תכונות התוחלת המותנית, נציג גישה אחרת (מודרנית יותר) להגדרת התוחלת המותנית. שתי הגישות זהות במובן שההגדרה זהה - אולם הפיתוח שונה.

הגדרנו תוחלת מותנית על ידי השלכה במרחב הילברט (אם יש להם מומנט שני), וקירובים אם לא. ההגדרה המודרנית נעשית דרך נגזרת רדון ניקודים (אך כמובן מביאה לאותו איפיון).

הגדרה 2.81 יהי $(\Omega,\mathcal{F})$ מרחב הסתברות. מידה (חיובית) סופית ν היא כזו המקבלת ערכים חיובים, ו-$\nu(A)/\nu(\Omega)$ היא מידת הסתברות.

עבור שתי מידות סופיות ν,μ נסמן $\mu \ll \nu$ אם לכל קבוצה מדידה A, אם $\mu(A)=0$ אזי גם $\nu(A)=0$. לדוגמה, יהי X מ"א חיובי עם תוחלת סופית. אזי $\nu(A)=\int_A X\,d\mathbb{P}$ מגדיר מידה חיובית ו-$\nu \ll \mathbb{P}$.

משפט 2.82 (רדון-ניקודים). אם ν,μ מידות סופיות על $(\Omega,\mathcal{F})$ ו-$\mu \ll \nu$ אזי קיים מ"א X כך שלכל A ב-$\mathcal{F}$ מתקיים $\nu(A)=\int_A X\,d\mu$. לכל מ"א Y בעל תוחלת סופית מתקיים

$$\int_\Omega Y\,d\mu = \int_\Omega YX\,d\nu. \tag{2.42}$$

המ"א X לעיל נקרא נגזרת רדון-ניקודים של ν ביחס ל-μ, ומקובל לסמנו בצורה

$$X = \frac{d\mu}{d\nu}\ . \tag{2.43}$$

המשוואה (2.42) נראת אז כך:

$$\int_\Omega Y\,d\mu = \int_\Omega Y\frac{d\mu}{d\nu}\,d\nu \qquad \mathbb{E}_\mu[Y] = \mathbb{E}_\nu\left[\frac{d\mu}{d\nu}Y\right]\ . \tag{2.44}$$

משפט 2.83 (קיום תוחלת מותנית). נתון מרחב הסתברות $(\Omega, \mathcal{F})$ ותת-סיגמה-שדה $\mathcal{F}_1$. יהי X מ"א חיובי בעל תוחלת סופית. נגדיר מידה ν על $(\Omega, \mathcal{F}_1)$ על ידי

$$\nu(A) = \int_A X \, d\mathbb{P}, \quad A \in \mathcal{F}_1 \ . \tag{2.45}$$

אזי ν מידה סופית על $(\Omega, \mathcal{F}_1)$ ו-

$$\mathbb{E}\left[X \mid \mathcal{F}_1\right] = \frac{d\nu}{d\,\mathbb{P}} \ . \tag{2.46}$$

נשים לב כי ν מוגדרת על $\mathcal{F}_1$ אבל לא על $\mathcal{F}$, ולכן גם נגזרת רדון ניקודים היא ביחס למרחב $(\Omega, \mathcal{F}_1)$, וממשפט רדון ניקודים היא מדידה $\mathcal{F}_1$.

הוכחה: קל לראות כי ν חיובית וממלאת את הדרישות כדי להיות מידה (ואם ננרמל בתוחלת של X נקבל מידת הסתברות). ממשפט רדון ניקודים הנגזרת קיימת, והיא מ"א מדיד $\mathcal{F}_1$. כדי להוכיח שזו התוחלת המותנית נשאר להראות כי לכל $A \in \mathcal{F}_1$ מתקיים

$$\mathbb{E}[X\mathbf{1}_{\{A\}}] = \mathbb{E}\left[\frac{d\nu}{d\,\mathbb{P}}\mathbf{1}_{\{A\}}\right] \ . \tag{2.47}$$

אולם לפי משפט רדון ניקודים, כיון ש-$\mathbf{1}_{\{A\}}$ מדידה על $\mathcal{F}_1$ אזי

$$\int_\Omega \mathbf{1}_{\{A\}} X \, d\,\mathbb{P} = \nu(A) \tag{2.48}$$

$$= \int_A \left[\frac{d\nu}{d\,\mathbb{P}}\right] \, d\,\mathbb{P} \tag{2.49}$$

$$= \int_\Omega \left[\frac{d\nu}{d\,\mathbb{P}}\right] \mathbf{1}_{\{A\}} \, d\,\mathbb{P} \ . \tag{2.50}$$

אם X אינו חיובי אזי מפרקים $X = X^+ - X^-$ כאשר שניהם חיוביים, ומגדירים את התוחלת המותנית של X כהפרש התוחלות המותנות של שני החלקים.

את היחידות אפשר להוכיח, אך היא נובעת מתכונת ההחלקה שנוכיח בהמשך.

<u>**תכונות תוחלת מותנית**</u>

1. ליניאריות

$$\mathbb{E}\left[\alpha X_1+\beta X_2 \mid \mathcal{F}_1\right]=\alpha \,\mathbb{E}\left[X_1 \mid \mathcal{F}_1\right]+\beta \,\mathbb{E}\left[X_2 \mid \mathcal{F}_1\right]$$

בתנאי שצד ימין מוגדר היטב.

הוכחה: מההגדרה. אם צד ימין מוגדר היטב אזי $X_1,\ X_2$ אינטגרבילים, ולכן סכומם אינטגרבילי ומכאן שצד שמאל מוגדר היטב. כעת, שני הצדדים מדידים $\mathcal{F}_1$. יהי $A \in \mathcal{F}_1$; אזי מהגדרת צד שמאל:

$$\int \mathbb{E}\left[\alpha X_1+\beta X_2 \mid \mathcal{F}_1\right]\mathbf{1}_{\{A\}}\,\mathbb{P}(d\omega) \tag{2.51}$$

$$=\int(\alpha X_1+\beta X_2)\mathbf{1}_{\{A\}}\,\mathbb{P}(d\omega) \tag{2.52}$$

$$=\alpha\int X_1\mathbf{1}_{\{A\}}\,\mathbb{P}(d\omega)+\beta\int X_2\mathbf{1}_{\{A\}}\,\mathbb{P}(d\omega) \tag{2.53}$$

מלינאריות התוחלת. כעת מהגדרת צד ימין,

$$=\alpha\int \mathbb{E}\left[X_1 \mid \mathcal{F}_1\right]\mathbf{1}_{\{A\}}\,\mathbb{P}(d\omega)+\beta\int \mathbb{E}\left[X_2 \mid \mathcal{F}_1\right]\mathbf{1}_{\{A\}}\,\mathbb{P}(d\omega) \tag{2.54}$$

נשתמש שוב בלינאריות התוחלת, ונקבל שצד ימין מקיים את (2.41).

2. אם $X(\omega)=C$ בהסתברות 1 אזי $\mathbb{E}\left[X \mid \mathcal{F}_1\right]=C$ בהסתברות 1.

3. אם X מדיד על $\mathcal{F}_1$ ו- $\mathbb{E}|X|<\infty$ אזי $\mathbb{E}\left[X \mid \mathcal{F}_1\right]=X$.

4. אם X מדיד על $\mathcal{F}_1$ ושני הביטויים מוגדרים, אזי $\mathbb{E}\left[XY \mid \mathcal{F}_1\right]=X\,\mathbb{E}\left[Y \mid \mathcal{F}_1\right]$.
הוכחה: שני הצדדים מדידים $\mathcal{F}_1$. כעת עבור $A \in \mathcal{F}_1$

$$\mathbb{E}[(XY-X\,\mathbb{E}\left[Y \mid \mathcal{F}_1\right])\cdot\mathbf{1}_{\{A\}}]=\mathbb{E}[(Y-\mathbb{E}\left[Y \mid \mathcal{F}_1\right])\cdot(X\mathbf{1}_{\{A\}})]$$

אבל $X\mathbf{1}_{\{A\}}$ מדיד $\mathcal{F}_1$, ומההגדרה (2.41) הבטוי מתאפס.

תרגיל 2.84 הוכח את 3,2 לעיל.

תרגיל 2.85 הוכח כי $\mathbb{E}X=\mathbb{E}\{\mathbb{E}\left[X \mid \mathcal{F}_0\right]\}$ כאשר $\mathcal{F}_0=\{\emptyset,\Omega\}$. הוכח כי $\mathbb{E}\{\mathbb{E}\left[X \mid \mathcal{F}_1\right]\}=\mathbb{E}X$.

5. תכונת ההחלקה: יהי $\mathcal{F}_2 \subset \mathcal{F}_1 \subset \mathcal{F}$. אזי

$$\mathbb{E}\left[X \mid \mathcal{F}_2\right] = \mathbb{E}\left[\,\mathbb{E}\left[X \mid \mathcal{F}_1\right] \mid \mathcal{F}_2\right] = \mathbb{E}\left[\,\mathbb{E}\left[X \mid \mathcal{F}_2\right] \mid \mathcal{F}_1\right]$$

(הסימון $\mathcal{F}_1 \subset \mathcal{F}$ פרושו: כל קבוצה השייכת ל $\mathcal{F}_1$ בהכרח שייכת גם ל-$\mathcal{F}$).

הוכחה: כיון ש- $\mathbb{E}\left[X \mid \mathcal{F}_2\right]$ מדיד $\mathcal{F}_2$, הוא גם מדיד $\mathcal{F}_1$, ומ-3 נובע כי הבטוי הראשון והאחרון שוים. יהי $A \in \mathcal{F}_2$; אזי מהגדרת תוחלת מותנת (בייחס ל-$\mathcal{F}_2$):

$$\mathbb{E}\{\mathbb{E}\left[\,\mathbb{E}\left[X \mid \mathcal{F}_1\right] \mid \mathcal{F}_2\right] \cdot \mathbf{1}_{\{A\}}\} = \mathbb{E}\{\mathbb{E}\left[X \mid \mathcal{F}_1\right] \cdot \mathbf{1}_{\{A\}}\} \tag{2.55}$$

אבל $A \in \mathcal{F}_1$ ולכן מתכונה 4

$$= \mathbb{E}[\mathbb{E}\left[\mathbf{1}_{\{A\}}X \;\middle|\; \mathcal{F}_1\right]] \tag{2.56}$$

$$= \mathbb{E}(\mathbf{1}_{\{A\}}X) = \mathbb{E}\{(\mathbf{1}_{\{A\}})\,\mathbb{E}\left[X \mid \mathcal{F}_2\right]\} \tag{2.57}$$

לכן הבטוי האמצעי מדיד $\mathcal{F}_2$ ומקיים את (2.41) ולכן הוא שווה לביטוי השמאלי.

6. אם X בלתי תלוי ב-$\mathcal{F}_1$ (כלומר X אינו תלוי בכל Z המדיד $\mathcal{F}_1$) אזי

$$\mathbb{E}\left[X \mid \mathcal{F}_1\right] = \mathbb{E}\,X$$

הוכחה: יהי Z מ"א מדיד $\mathcal{F}_1$. מתכונות התוחלת, עבור מ"א בלתי תלויים $\mathbb{E}\,XZ = \mathbb{E}\,X \cdot \mathbb{E}\,Z$ לכן

$$\mathbb{E}[XZ] = \mathbb{E}\,X \cdot \mathbb{E}\,Z = \mathbb{E}[Z \cdot \mathbb{E}\,X]$$

והטענה נובעת מכך ש-$\mathbb{E}\,X$ בברור מדיד על $\mathcal{F}_1$.

7. **משפט 2.86** [אי שוויון ינסן - $Jensen\ inequality$]: אם $g(\cdot)$ פונקציה ממשית קונווקסית אזי מתקיים

$$\mathbb{E}\left[g(X) \mid \mathcal{F}_1\right] \geq g[\mathbb{E}\left[X \mid \mathcal{F}_1\right]]$$

בפרט

$$|\,\mathbb{E}\left[X \mid \mathcal{F}_1\right]| \leq \mathbb{E}\left[|X| \mid \mathcal{F}_1\right]$$

8. <u>התכנסות מונוטונית</u>: יהיו X_n ו-X ב-L^1. אם $X_n \uparrow X$ אזי $\mathbb{E}\left[X_n \mid \mathcal{F}_1\right] \uparrow \mathbb{E}\left[X \mid \mathcal{F}_1\right]$ אם $X_n \downarrow X$ אזי $\mathbb{E}\left[X_n \mid \mathcal{F}_1\right] \downarrow \mathbb{E}\left[X \mid \mathcal{F}_1\right]$. בפרט אם $X_1 \leq X_2$ אזי $\mathbb{E}\left[X_1 \mid \mathcal{F}_1\right] \leq \mathbb{E}\left[X_2 \mid \mathcal{F}_1\right]$.

9. <u>Dominated Convergence</u>: **אם** $|X_n| \le Y$ **ו-** $\mathbb{E}\, Y < \infty$, **ואם** $X_n \to X$ **בהסתברות 1 אזי**

$$\mathbb{E}\left[X_n \mid \mathcal{F}_1\right] \to \mathbb{E}\left[X \mid \mathcal{F}_1\right]$$

תרגיל 2.87 הוכח, בעזרת משפטי ההתכנסות, את אי השוויון של ינסן, בשיטה הבאה:
לכל פונקציה קונווקסית g, $g(x) - g(y) \ge g'(y)(x-y)$ כאשר $g'(y)$ היא הנגזרת מימין של g בנקודה y.
נסמן

$$A_N = \{w : |\mathbb{E}\left[X \mid \mathcal{F}_1\right]| \ge N\} \tag{2.58}$$

$$X_N = X\mathbf{1}_{\{A_N\}} \tag{2.59}$$

ונסמן,

$$Y_N = \mathbb{E}\left[X_N \mid \mathcal{F}_1\right] = \mathbb{E}\left[X \mid \mathcal{F}_1\right]\mathbf{1}_{\{A_N\}} \tag{2.60}$$

הצב זאת בהערה (2.41) העזר ב- $\mathbb{E}\left[g(X_N) \mid \mathcal{F}_1\right] = \mathbb{E}\left[g(X) \mid \mathcal{F}_1\right]\mathbf{1}_{\{A_N\}} + g(0)(1 - \mathbf{1}_{\{A_N\}})$.
יש צורך ב-A_N <u>רק</u> כדי לטפל במקרה בו $\mathbb{E}\,|X| < \infty$ אבל $\mathbb{E}\,|g(X)| = \infty$ ואז צריך להבהיר כיצד $\mathbb{E}\left[g(X) \mid \mathcal{F}_1\right]$ מוגדר.

תרגיל 2.88 הוכח בעזרת 7 לעיל כי אם $X_n \xrightarrow{r} X$ אזי גם $\mathbb{E}\left[X_n \mid \mathcal{F}_1\right] \xrightarrow{r} \mathbb{E}\left[X \mid \mathcal{F}_1\right]$.

בתרגיל הבא נרחיב הגדרת תוחלת מותנת למ"א X אינטגרבלים <u>שאינם</u> ב-L_2. נעשה זאת בשלושה שלבים. ראשית, נניח X חיובי, ונגדיר $\mathbb{E}\left[X \mid \mathcal{F}_1\right]$ ע"י גבול, שנית, נראה שהגבול אינו תלוי בצורת הקירוב, לבסוף נגדיר $\mathbb{E}\left[X \mid \mathcal{F}_1\right] = \mathbb{E}\left[X^+ \mid \mathcal{F}_1\right] - \mathbb{E}\left[X^- \mid \mathcal{F}_1\right]$ למשתנה X אינטגרבילי שאינו חיובי.

תרגיל 2.89 קיום תוחלת מותנית למשתנים בעלי תוחלת סופית.
יהי X מ"א חיובי אינטגרבילי (אך לא בהכרח מקיים $\mathbb{E}\, X^2 < \infty$), סידרת המ"א

$$X_n(\omega) = \begin{cases} X(\omega) & if \quad X(\omega) \le n \\ n & if \quad X(\omega) > n \end{cases}$$

היא סדרת מ"א חסומים המקיימים $X_n \uparrow X$ כמעט בכל מקום. <u>נגדיר</u>

$$\mathbb{E}\left[X \mid \mathcal{F}_1\right] = \lim_{n\to\infty} \mathbb{E}\left[X_n \mid \mathcal{F}_1\right]$$

כאשר צד ימין מוגדר כרגיל, דרך ההטלה ב-L_2.

א. הוכח כי הגבול קיים בהסתברות 1. נסמן גבול זה ב-Z_X.

כדי שההגדרה תהיה סבירה, צריך ש-Z_X לא יהיה תלוי בבחירת הסדרה X_n. תהי Y_n סדרת מ"א המקימים $\mathbb{E}\,|Y_n|^2 < \infty$ וכן $Y_n \uparrow X\ a.s.$.

ב. הוכח כי קיים הגבול $Z_Y = \lim\limits_{n\to\infty} \mathbb{E}\left[Y_n \mid \mathcal{F}_1\right]$.

ג. הוכח $\mathbb{P}(Z_X = Z_Y) = 1$. רמז: הוכח $\mathbb{E}\left|\mathbb{E}\left[Y_n \mid \mathcal{F}_1\right] - \mathbb{E}\left[X_n \mid \mathcal{F}_1\right]\right| \to 0$. העזר באי-שויון ינסן, אי שויון המשולש ומשפט ההתכנסות המונוטונית.

על סמך תרגיל זה נגדיר, כאשר $\mathbb{E}\,|X| < \infty$

$$\mathbb{E}\left[X \mid \mathcal{F}_1\right] = \mathbb{E}\left[X^+ \mid \mathcal{F}_1\right] - \mathbb{E}\left[X^- \mid \mathcal{F}_1\right] \tag{2.61}$$

$$X^+ = \max(X, 0) \tag{2.62}$$

$$X^- = \max(-X, 0) \tag{2.63}$$

וההגדרה "בסדר" בזכות התרגיל:

בגלל התנאי $\mathbb{E}\,|X| < \infty$ **נקבל** $\mathbb{E}\left[X^+ \mid \mathcal{F}_1\right] < \infty$ **וכן** $\mathbb{E}\left[X^- \mid \mathcal{F}_1\right] < \infty$ **בהסתברות 1.**

Regular Conditional Probability

תרגיל 2.90 נתונים מ"א X ו-Y עם צפיפות משותפת $\rho(x,y)$. תהי f פונקצית בורל כך ש- $\mathbb{E}\,|f(X)| < \infty$. אזי

$$\mathbb{E}\left[f(X) \mid Y\right] = \frac{\int_{-\infty}^{\infty} f(\alpha)\rho(\alpha, Y)d\alpha}{\int_{-\infty}^{\infty} \rho(\alpha, Y)d\alpha}$$

הבטוי $\rho(\cdot|Y)$ המוגדר ע"י

$$\int_A \rho(\alpha, Y)\left[\int_{-\infty}^{\infty} \rho(\beta, Y)d\beta\right]^{-1} d\alpha = \mathbb{P}\left[A \mid Y\right]$$

הוא מ"א לכל קבוצה A ב-$\mathcal{F}_Y$, והוא מידת הסתברות לכל $\omega \in \Omega$.

בגלל התכונה

$$\mathbb{E}\left[f(X) \mid Y\right] = \int_{-\infty}^{\infty} f(\alpha)\rho(d\alpha|Y)$$

הוא נקרא הסתברות מותנית רגולרית. היא קיימת בתנאים חלשים למדי.

הגדרה 2.91 יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. $\hat{\mathbb{P}}[A|\,\mathcal{F}_1]$ יקרא הסתברות מותנה רגולרית *regular conditional probability* אם

א. לכל $A \in \mathcal{F}$ $\hat{\mathbb{P}}[A|F_1] = \mathbb{P}[A|F_1]$ בהסתברות 1.

ב. לכל $\omega \in \Omega$ הגודל $\hat{\mathbb{P}}[\cdot|\,\mathcal{F}_1]$ הוא מידת הסתברות על $(\Omega, \mathcal{F})$.

יהי X מ"א על $(\Omega, \mathcal{F}, \mathbb{P})$. נקרא ל- $\hat{\mathbb{P}_X}[B|\,\mathcal{F}_1]$ פילוג מותנה רגולרי - *regular conditional distribution* של X בהנתן ה- σ-שדה $\mathcal{F}_1$ אם

ג. לכל קבוצת בורל של הממשיים $B \in \mathbb{B}$,
$\hat{\mathbb{P}}[B|\,\mathcal{F}_1] = \mathbb{P}\left[\omega : X(\omega) \in B \mid \mathcal{F}_1\right] = \mathbb{E}\left[\mathbf{1}_{\{B\}} \;\middle|\; \mathcal{F}_1\right]$ בהסתברות 1,

ד. לכל $\omega \in \Omega$, הגודל $\hat{\mathbb{P}}_X[\cdot|\,\mathcal{F}_1]$ היא מידת הסתברות על $((-\infty, \infty), \mathbb{B})$.

משפט 2.92 לכל מ"א ממשי X ולכל σ-שדה $\mathcal{F}_1$, קיים פלוג מותנה רגולרי, ומתקיים

$$\mathbb{E}\left[X \mid \mathcal{F}_1\right] = \int X(\omega)\hat{\mathbb{P}}(d\omega \mid \mathcal{F}_1) = \int_{-\infty}^{\infty} x\,\hat{\mathbb{P}}_X(dx \mid \mathcal{F}_1).$$

הערה 2.93 משפט זה נכון גם למשתנים המקבלים ערכים ב-$\mathbb{R}^N$ ואף למקרים מופשטים יותר, כגון מרחבי בורל.

החשיבות של בניות אלו היא הבאה. כפי שנראה בהמשך, בטיפול בתהליכים אקראיים, התהליך שהוא תוחלת מותנית של תהליך אחר מוגדר לכל t עד כדי מאורע שהסתברותו 1. כיוון שאוסף הזמנים אינו בן מניה, לא ברור שניתן למצוא מאורע עליו התוחלת המותנית מוגדרת לכל t. השימוש בהסתברות רגולרית מותנית או פילוג רגולרי מותנה קובעים למעשה מאורע יחיד, שהסתברותו אפס, ומחוץ לו כל התוחלות המותנות מוגדרות דרך האינטגרל לעיל.

תרגיל 2.94 הוכח את אי שויון ינסן המותנה, בשלבים הבאים. יהי X מ"א ו-ϕ פונקציה קונוקסית.

א. $\mathbb{E}\,\phi(X) \geq \phi(\mathbb{E}\,X)$. רמז: קרב את X בעזרת מ"א פשוטים והעזר בעובדה כי אם $\mathbb{P}_X$ הוא הפילוג של X, אזי $\mathbb{E}\,\phi(X) = \int_{-\infty}^{\infty} \phi(\alpha)\,\mathbb{P}_X(d\alpha)$.

ב. $\mathbb{E}\left[\phi(X) \mid \mathcal{F}_1\right] \geq \phi(\mathbb{E}\left[X \mid \mathcal{F}_1\right])$. רמז: העזר בא', ובמשפט *2.92*.

<u>אי תלות מותנית</u>

מאורעות A, B **נקראים בלתי תלוים סטטיסטית** (**בת"ס**) **אם** $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

σ-**שדות** $\mathcal{F}_1, \mathcal{F}_2$ **נקראים בת"ס אם לכל** $A \in \mathcal{F}_1$ **ו-** $B \in \mathcal{F}_2$, **מתקיים** $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

נסמן זאת ב-$\mathcal{F}_1 \perp \mathcal{F}_2$.

הגדרה שקולה היא הדרישה $\mathbb{E}\,\mathbf{1}_{\{A\}}\mathbf{1}_{\{B\}} = \mathbb{E}\,\mathbf{1}_{\{A\}}\,\mathbb{E}\,\mathbf{1}_{\{B\}}$.

מכאן נובעת התכונה לכל X **מדיד** $\mathcal{F}_1$ **ולכל** Y **מדיד** $\mathcal{F}_2$, $\mathbb{E}\,XY = \mathbb{E}\,X\,\mathbb{E}\,Y$.

משתנים בת"ס X, Y **נגדיר דרך** $\sigma(X), \sigma(Y)$.

הגדרה 2.95 $\mathcal{F}_1$ ו-$\mathcal{F}_2$ בלתי תלוים בהנתן $\mathcal{F}_3$ (סימון $\mathcal{F}_1 \underset{\mathcal{F}_3}{\perp} \mathcal{F}_2$) אם לכל $A \in \mathcal{F}_1, B \in \mathcal{F}_2$ מתקיים

$$\mathbb{E}\left[\mathbf{1}_{\{A\}}\mathbf{1}_{\{B\}} \;\middle|\; \mathcal{F}_3\right] = \mathbb{E}\left[\mathbf{1}_{\{A\}} \;\middle|\; \mathcal{F}_3\right]\mathbb{E}\left[\mathbf{1}_{\{B\}} \;\middle|\; \mathcal{F}_3\right]$$

תרגיל 2.96 נסמן ב-X_i את תוצאת ההטלה ה-i של קוביה.
יהי $Y = X_1 + X_2, Z = X_1 + X_3$ אזי $Z \underset{X_1}{\perp} Y$ אבל Y <u>אינו</u> בת"ס ב-Z.

הערה 2.97 יהיו X ו-Y מ"א כאשר $X \in L_2(\Omega, \mathcal{F}, \mathbb{P})$. אזי כזכור קיימת פונקציה g (בורל) כך ש-$\mathbb{E}\left[X \mid Y\right] = g(Y)$. לכן ניתן להגדיר $\mathbb{E}\left[X \mid Y = y\right] = g(y)$.

משפט 2.98 אם $X \perp (Y, Z)$ אזי $\mathbb{E}\left[Y \mid X, Z\right] = \mathbb{E}\left[Y \mid Z\right]$.

תרגיל 2.99 הוכח המשפט למקרה שקיימות צפיפויות.

תרגיל 2.100 <u>הפרך</u> את הטענה הבאה (בעזרת דוגמה נגדית):
אם $Y \perp X$ אזי לכל Z מתקיים $\mathbb{E}[Y \mid X, Z] = \mathbb{E}[Y \mid Z]$.

תרגיל 2.101 יהי (x, y) מ"א גאוסיים בת"ס עם תוחלת אפס ושונות 1.

א. יהא θ מספר. נגדיר

$$z_\theta = -x \sin\theta + y \cos\theta$$

$$w_\theta = x \cos\theta + y \sin\theta$$

הוכח כי (w_θ, z_θ) בלתי תלויים סטטיסטית. הוכח גם כי $N(0,1) \backsim P(z_\theta | w_\theta)$ והראה כי הנ"ל הוא *Regular Conditional Probability Distribution - r.c.p.d*

ב. יהא כעת θ מ"א המוגדר ע"י $\theta = \arctan(y/x)$, $\theta \in [0, 2\pi]$. מצא את $P(z_\theta | \theta)$ והראה כי הנ"ל הוא *r.c.p.d.*

ג. שים לב שלכל α הקבוצה $\{\omega : \theta(\omega) = \alpha\}$ והקבוצה $\{\omega : w_\alpha = 0\}$ הן אותה קבוצה. למרות זאת,

$$"P(z_\alpha > \frac{1}{2} | w_\theta = 0)" \neq "P(z_\theta > \frac{1}{2} | \theta = \alpha)"$$

<u>הסבר?</u>

3 תהליכים אקראיים

הגדרה 3.1 תהליך אקראי X_t על $(\Omega, \mathcal{F}, \mathbb{P})$ היא משפחה $\{X_t, t \in T\}$ של משתנים אקראיים X_t כאשר האינדקס t שייך לקבוצה T.

הפונקציה $t \to X_t(\omega)$ נקראת פונקצית מדגם. עבור תהליכים בזמן בדיד T יהיה בדרך כלל אוסף שלמים עוקבים (סופי או אינסופי). עבור תהליכים בזמן רציף T יהיה בדרך כלל אינטרוול (סופי או אינסופי). שדה אקראי הוא הכללה של תהליך אקראי - המתבטאת בכך ש- T הוא אוסף כללי יותר, למשל נקודות במישור.

נסמן ב-$\sigma(X_t)$ את ה-σ שדה שמשרה X_t, וב-$\mathcal{F}_X$ את ה-σ שדה הקטן ביותר המכיל את <u>כל</u> ה-$\sigma(X_t)$, לכל t.

כדי להבין את הנקודות העדינות שבהגדרת תהליך אקראי, נשים לב כי תחת הגדרה זו,

$$\{\omega : X_{t_1} \le \alpha\} \in \mathcal{F}_X$$

$$\{\omega : X_{1/k} \le \alpha, 0 < k \le N\} = \cap_{k=1}^{N} \{\omega : X_{1/k} \le \alpha\} \in \mathcal{F}_X$$

$$\left\{\omega : \sup\{X_{1/k}, 0 < k\} \le \alpha\right\} = \cap_{k=1}^{\infty} \{\omega : X_{1/k} \le \alpha\} \in \mathcal{F}_X$$

אבל לא מובטח כי

$$\left\{\omega : \sup\{X_t, 0 < t \le 1\} \le \alpha\right\} \in \mathcal{F}_X$$

כי ה-sup הוא על מספר מ"א <u>שאינו</u> בן מניה.

בצורה דומה, נתבונן בפונקציות $t \to X_t(\omega)$ ונגדיר

$$\{\omega : X_t(\omega)\, \text{measurable} \quad \text{function}\, T \to \mathbb{R}\}$$

גם זה לא בהכרח מאורע!

שאלה נוספת קשורה לתאור ההסתברותי של התהליך.

נניח <u>שלכל</u> קבוצת אינדקסים $t_1, t_2, \ldots, t_N$ ידוע הפילוג

$\mathbb{P}_{t_1, t_2, \ldots, t_N}(\alpha_1, \alpha_2, \ldots, \alpha_N) = \mathbb{P}\{X_{t_1} \le \alpha_1, \ldots, X_{t_N} \le \alpha_N\}$ **- פילוג N-ממדי של** X.

האם מידע זה מספיק כדי לחשב $\mathbb{P}$ { $\omega : [0,1]$ רציף ב- $X_t(\omega)$ }?

האם קיימת $\mathbb{P}$ המתארת את <u>התהליך</u> כולו?

דוגמה 3.2 $\Omega = [0,1]$, עם שדה בורל ומידת לבג l (כלומר $l[a,b]=b-a$, עבור $b > a$). עבור $T=[0,1]$

נגדיר שני תהליכים: $Y_t(\omega) \equiv 0$ ו-

$$X_t(\omega) = \begin{cases} 1 & t = \omega \\ 0 & t \neq \omega \end{cases}$$

אזי לכל t, $\mathbb{P}(X_t = Y_t) = 1$. יהי $A_t = \{\omega : X_t \neq Y_t\} = \{t\}$. אזי אם $A^N = \bigcup_{i=1}^{N} A_{t_i}$, ודאי יתקיים $\mathbb{P}(A^N) = 0$, ובאותה צורה $A^\infty = \bigcup_{i=1}^{\infty} A_{t_i}$ גורר $\mathbb{P}(A^\infty) = 0$. אבל $\mathbb{P}(\sup\{X_t(\omega), 0 \leq t \leq 1\} \neq 0) = \mathbb{P}(1 \neq 0) = 1$.

ולעומת זאת $\mathbb{P}(\sup\{Y_t(\omega), 0 \leq t \leq 1\} \neq 0) = 0$.

קיבלנו שכל הפילוגים הסופיים של X ושל Y זהים, אך פילוג המקסימום שונה!

בנוסף, $\mathbb{P}(\{\omega : \text{רציף } t \to Y_t(\omega)\}) = 1$, $\mathbb{P}(\{\omega : \text{רציף } t \to X_t(\omega)\}) = 0$.

3.1 שיויון מדידות ורציפות של תהליכים

הגדרה 3.3 $\{X_t\}$ ו-$\{Y_t\}$ נקראים גרסה *(version)* אחד של השני אם לשניהם אותם פילוגים סופיים.

ראינו בדוגמה ששנוי גרסה גורר שנויים במסלולי התהליך. השאלה היא כיצד לתת מסגרת תחתה ניתן לדון בשאלות כמו מדידות ב-t, הסתברות שהפונקציה רציפה וכו'.

הגדרה 3.4 תהליך Y_t יקרא *modification* של X_t אם $\mathbb{P}(X_t = Y_t) = 1$ לכל t.

<u>שים לב</u> שגרסה אינה חייבת להיות מוגדרת על אותו מרחב הסתברות בניגוד ל-modification.

תרגיל 3.5 אם Y_t הוא *modification* של X_t אזי Y_t הוא גרסה של X_t.

זוהי דרישה חזקה יותר מאשר הדרישה כי Y_t הוא גרסה של X_t.

תרגיל 3.6 נגדיר X_t ו-Y_t על $0 \leq t \leq 2$ כך: נזרוק שתי מטבעות בלתי תלויים; ואז, באינטרוול $i - 1 \leq t \leq i$, כאשר $i = 1$ או $i = 2$

$$X_t = \begin{cases} 1 & \text{מטבע } i \text{ נפלה על עץ} \\ 0 & \text{מטבע } i \text{ נפלה על פלי} \end{cases} \qquad Y_t = \begin{cases} 1 & \text{מטבע } i \text{ נפלה על פלי} \\ 0 & \text{מטבע } i \text{ נפלה על עץ} \end{cases}$$

הוכח כי Y_t הוא גרסה של X_t אך אינו *modification*.

הגדרה 3.7 ת"א Z_t נקרא רציף בהסתברות בזמן t_0 אם לכל $\varepsilon > 0$

$$\lim_{t\to t_0} \mathbb{P}\{|Z_t(\omega) - Z_{t_0}(\omega)| > \varepsilon\} = 0$$

בדוגמה 3.2, X_t **רציף בהסתברות! רציפות בהסתברות אפשר לבדוק מתוך ידיעת** $\mathbb{P}_{t_1,t_2}(\alpha_1,\alpha_2)$ **לכל** $t_1,t_2,\alpha_1,\alpha_2$!

הגדרה 3.8 תהליך X_t עם מרחב זמן T נקרא ספרבילי *(Separable)* אם קיימת קבוצה בת מניה $T_c \subset T$ וקיים מאורע C עבורו $\mathbb{P}(C)=0$, המקיימים את התכונה הבאה. בהנתן α,β ואינטרוול סגור I, נגדיר

$$A = \{\omega : \alpha \le X_t \le \beta,\, t \in T\cap I \text{ לכל}\},\quad B = \{\omega : \alpha \le X_t \le \beta,\, t \in T_c\cap I \text{ לכל}\}.$$

אזי $\{\omega : \omega \notin A,\, \omega \in B\} \subset C$.

הערה 3.9 בדרך כלל בתפקיד T_c ישמשו המספרים הרציונליים אשר ב-T, אלא אם כן נאמר במפורש אחרת.
שים לב כי הקבוצה A אינה בהכרח מאורע!

הערה 3.10 X_t בדוגמה *3.2* אינו ספרבילי.

משפט 3.11 [עמ' 57 *Doob*, ללא הוכחה]: אם $\tilde{X}_t$ ת"א רציף בהסתברות בכל t בקטע *[a,b]* אזי קיים תהליך X_t באותו אינטרוול כך ש:

א. X_t *modification* של $\tilde{X}_t$.

ב. X_t ספרבילי.

ג. $t\to X_t(\omega)$ פונקצית בורל עבור כמעט כל ω.

לגרסה X_t נקרא גרסה ספרבילית. בהמשך נדון רק בתהליכים ספרבילים. בהגדרת התנועה הבראונית, למשל, נגדיר את מידת ההסתברות על תהליכים רציפים בלבד, כך שהבעיה תפתר.

תרגיל 3.12 אם לת"א X_t כל הדגמים רציפים מימין עם גבולות משמאל אזי התהליך הוא *Separable*.

<u>רציפות של ת"א</u> - ת"א יקרא רציף ב- t_0 (a.s.), בהסתברות, בממוצע מסדר (r) אם $X_t \to X_{t_0}$ כאשר $t \to t_0$ במובן המתאים. התהליך יקרא רציף ב-[a,b] אם הוא רציף לכל $t \in [a,b]$.
תכונה חזקה יותר היא רציפות הדגמים בהסתברות 1:

משפט 3.13 [קולמוגורוב - *Kolmogorov*]: נתון תהליך $\{X_t,\ t \in T\}$. אם קיימים קבועים חיוביים α, β, c כך שלכל $h < h_0$

$$\mathbb{E}\,|X_{t_h} - X_t|^{\alpha} \le c h^{1+\beta}$$

אזי קיים *modification* ספרבילי, עבורו

$$\sup_{t,s \in T, |t-s|<h} |\tilde{X}_t - \tilde{X}_s| \overset{a.s.}{\underset{h\to 0}{\longrightarrow}} 0$$

כלומר הדגמים רציפים בהסתברות 1.
<u>הוכחה:</u> ראה *Wong & Hajek* עמ' *57*.

הגדרה 3.14 תהליך X_t יקרא סטציונרי אם פונקציות הפילוג הרב ממדיות מקיימות

$$F_{t_1,t_2,\ldots,t_N}(x_1,x_2,\ldots,x_N) = \mathbb{P}(x_{t_1} \le x_1,\ldots,x_{t_N} \le x_N) = F_{t_1+\tau,t_2+\tau,\ldots t_N+\tau}(x_1,\ldots,x_N)$$

לכל סדרה $t_1,\ldots,t_N$ לכל τ ולכל $x_1,\ldots,x_N$.

תרגיל 3.15 יהי X_t ת"א, תוחלת אפס, $\mathbb{E}\,{X_t}^2 = 1$ ו-$\lambda > 0$.

א. אם לכל $0 < \varepsilon < \varepsilon_0$ מתקיים $1 - \lambda\varepsilon^2 \le \mathbb{E}(X_t X_{t+\varepsilon}) \le 1$ אזי קיים *modification* רציף.

ב. אם X_t גאוסי ו- $1 - \varepsilon\lambda \le \mathbb{E}(X_t X_{t+\varepsilon}) \le 1$ אזי קיים *modification* רציף.

ג. מצא תנאים (דומים), תחתם קיים *modification* גזיר ברציפות j פעמים.

ד. אם X_t גאוסי, בטא את תנאי ג' בעזרת מומנטים מסדר שני בלבד.

הערה 3.16 באות הנקרא *"Random telegraph signal"* מתקיים ב', (ללא גאוסיות) אך אין רציפות הדג-מים שכן זהו "תהליך מניה" <u>שאינו</u> רציף בהסתברות 1.

תרגיל 3.17 הראה כי תהליך פואסון הוא רציף בכל נקודה בהסתברות 1, אך הדגמים אינם רציפים בהסתב-
רות 1.

הערה 3.18 לסעיף ג' של תרגיל *3.15* ניתן להעזר במשפט להלן.

משפט 3.19 [*K&S* עמ' *53*]: יהי $\{X_t,\ 0 \le t \le T\}$ תהליך על $(\Omega, \mathcal{F}, \mathbb{P})$ כך שעבור α, β ו- c חיוביים

$$\mathbb{E}\,|X_t - X_s|^{\alpha} \le C|t-s|^{1+\beta} \quad 0 \le s,t \le T$$

אזי קיים תהליך Y_t שהוא רציף, והוא *modification* של X_t, וכך ש-Y_t רציף לוקלית במובן הולדר *(Holder continuous)* עם אקספוננט $\gamma < 1$, לכל $\gamma \in (0, \beta/\alpha)$, כלומר

$$\mathbb{P}\left\{\omega : \sup_{\substack{0<t-s<h(\omega)\\ 0\le s,t\le T}} \frac{|Y_t(\omega) - Y_s(\omega)|}{|t-s|^{\gamma}} \le \delta\right\} = 1$$

כאשר $h(\omega)$ מ"א חיובי ו-$\delta > 0$ קבוע.

<u>מושגי מדידות וצרות של מדידות:</u>

הגדרה 3.20 ת"א X_t יקרא תהליך מדיד אם הפונקציה $X_t(\omega) : \mathbb{R} \times \Omega \to \mathbb{R}$, כפונקציה של שני משתנים, היא פונקציה מדידה מ-$(\mathbb{R} \times \Omega, \mathbb{B} \times \mathcal{F})$ ל-$(\mathbb{R}, \mathbb{B})$.

הערה 3.21 ההגדרה לת"א ווקטוריים אנלוגית.

שים לב שהגדרת ת"א לא דרשה מדידות!

משפט 3.22 [ללא הוכחה]: אם לת"א X_t דגמים רציפים בהסתברות 1, אזי X_t מדיד.

משפט 3.23 [פוביני, ללא הוכחה]: אם X_t מדיד אזי, כמעט לכל ω, הפונקציה $t \to X_t(\omega)$ מדידה (מ-$(\mathbb{R}, \mathbb{B})$ ל-$(\mathbb{R}, \mathbb{B})$). אם I הוא אינטרוול כך ש- $\int_I \mathbb{E}\,|X_t| dt < \infty$ אזי

$$\int_I \mathbb{E}\, X_t dt = \mathbb{E} \int_I X_t dt \quad \text{וכן} \quad \int_I X_t dt < \infty \quad a.s.$$

והאנטגרל האחרון מגדיר מ"א (כלומר הוא מדיד). אם $\mathbb{E}\,|X_t| < \infty$ באינטרוול I אזי הפונקציה $t \to \mathbb{E}\,X_t$ מדידה על אינטרוול זה.

נשים לב כי עבור תהליך שאינו מדיד, יתכן ולא ניתן לבצע אינטגרל במישור הזמן, שכן פונקציות המדגם אינן מדידות! בהמשך הקורס נדון רק בתהליכים עם פונקציות מדגם רציפות, או לפחות רציפות מימין עם גבולות משמאל. במקרה זה התהליכים תמיד מדידים.

3.2 תהליכי מרטינגל

דוגמה 3.24 יהי X מ"א. נניח שמודדים את X, והתוצאה במדידה ה-i היא Y_i.
נעריך את X, בשלב i ע"י משערך אופטמלי

$$\hat{X}_1 = \mathbb{E}\left[X \mid Y_1\right] \tag{3.1}$$

$$\hat{X}_2 = \mathbb{E}\left[X \mid Y_1, Y_2\right] \tag{3.2}$$

$$\hat{X}_i = \mathbb{E}\left[X \mid Y_1, Y_2, \ldots, Y_i\right] = \mathbb{E}\left[X \mid \mathbb{B}_i\right] \tag{3.3}$$

כאשר $\mathbb{B}_i = \sigma(Y_1, \ldots, Y_i)$. אזי $\mathbb{B}_i \subset \mathbb{B}_{i+1}$ ומתקיים

$$\mathbb{E}\left[\hat{X}_{i+1} \;\middle|\; \mathbb{B}_i\right] = \mathbb{E}\left[\,\mathbb{E}\left[X \mid \mathbb{B}_{i+1}\right] \mid \mathbb{B}_i\right] = \mathbb{E}\left[X \mid \mathbb{B}_i\right] = \hat{X}_i$$

נסמן $\mathbb{D}_i = \sigma(\hat{X}_1, \ldots, \hat{X}_i)$. אזי $\mathbb{D}_i \subset \mathbb{B}_i$, X_i מדיד על $\mathbb{D}_i$ וכן (שוב בעזרת תכונת ההחלקה) - $\mathbb{E}\left[\hat{X}_{i+1} \;\middle|\; \mathbb{D}_i\right] = \hat{X}_i$.

קל לראות (ע"י החלקה) **כי שגיאת השערוך $X - \hat{X}_i$ ניצבת לכל מ"א Z המדיד $\mathbb{B}_i$, כלומר** $\mathbb{E}[(X - \hat{X}_i)Z] = 0$.

הגדרה 3.25 סדרת σ-שדות $\mathbb{B}_n$ המקיימת $\mathbb{B}_n \subset \mathbb{B}_{n+1}$ לכל n נקראת *filtration*. סדרה של מ"א $\{X_n\}$ נקראת מתואמת *adapted* לסדרת σ-שדות $\mathbb{B}_n$ כזו אם לכל n, X_n מדיד על $\mathbb{B}_n$. תהי $\{X_n\}$ סדרה מתואמת כך ש- $\mathbb{E}\,|X_n| < \infty$. $\{X_n, \mathbb{B}_n\}$ נקרא:
מרטינגל *Martingale* אם מתקיים, לכל n, $\mathbb{E}\left[X_{n+1} \mid \mathbb{B}_n\right] = X_n \quad a.s.$
סופרמרטינגל *Supermartingale* אם מתקיים, לכל n, $\mathbb{E}\left[X_{n+1} \mid \mathbb{B}_n\right] \le X_n \quad a.s.$
סבמרטינגל *Submartingale* אם מתקיים, לכל n, $\mathbb{E}\left[X_{n+1} \mid \mathbb{B}_n\right] \ge X_n \quad a.s.$.

שים לב שמההגדרה נובע כי עבור מרטינגל מתקיים $\mathbb{E}\left[X_n \mid \mathbb{B}_m\right] = X_m$ בהסתברות 1 עבור כל $m < n$, ובצורה דומה ניתן להחליף את $n+1, n$ ב- n, m כנ"ל בהגדרות של סופר וסבמרטינגל.

למשל, אם X_n מתאר את ההון של מהמר לאחר n משחקים ו-$\mathbb{B}_n$ המידע שברשותו ברגע זה, אזי מרטינגל מתאר משחק הוגן (אין רווח/הפסד בממוצע), וסופר מרטינגל משחק בו צפוי הפסד למהמר.

תרגיל 3.26 אם $\{X_n, \mathbb{B}_n\}$ הוא מרטינגל ו-ϕ קמורה (קונווקסית) כך ש- $\mathbb{E}\left|\phi(X_n)\right| < \infty$ אזי $\{\phi(X_n), \mathbb{B}_n\}$ סב מרטינגל (רמז - השתמשו בינסן). בפרט $\{X_n^2, \mathbb{B}_n\}$ יהיה סב מרטינגל. אם $(-\phi)$ קונווקסית, אזי $\{\phi(X_n), \mathbb{B}_n\}$ הוא סופרמרטינגל.

תרגיל 3.27 יהי $\{X_n, \mathbb{B}_n\}$ מרטינגל. האם $\{X_n, \mathbb{D}_n\}$ מרטינגל אם נתון:

א. $\mathbb{B}_n \subset \mathbb{D}_n$

ב. $\mathbb{D}_n \subset \mathbb{B}_n$ ו- $\{X_n\}$ מתואם ל- $\{\mathbb{D}_n\}$.

לשם קיצור, נהוג לומר כי X_n הוא מרטינגל כאשר $\mathbb{B}_n = \sigma(X_1, \ldots, X_n)$.

הערה 3.28 ניתן להגדיר מרטינגל ע"י הדרישה כי לכל מ"א Z_n המדיד $\mathbb{B}_n$, $\mathbb{E}[(X_{n+1} - X_n)Z_n] = 0$. דהינו תהליך הוא מרטינגל אם ורק אם הוא בעל תוספות אורטוגונליות. לכן, בפרט כל תהליך בעל תוספות בת"ס עם ממוצע 0 הוא מרטינגל (בתנאי אינטגרביליות כמובן).

הגדרה 3.29 תהליך $\{X_t, \mathbb{B}_t, t \geq 0\}$ הוא מרטינגל בזמן רציף אם $\mathbb{E}\left|X_t\right| < \infty$ לכל t, ולכל $h > 0$ מתקיים $\mathbb{B}_t \subset \mathbb{B}_{t+h}$ וכן X_t מדיד $\mathbb{B}_t$ ו-

$$\mathbb{E}\left[X_{t+h} \mid \mathbb{B}_t\right] = X_t \quad a.s.$$

התהליך יקרא *submartingale* אם במקום שיוויון מופיע $\geq$, *supermartingale* אם $\leq$.

תרגיל 3.30 יהי $\{X_t, \mathbb{B}_t, t \geq 0\}$ מרטינגל (סב מרטינגל) ו-$\mathcal{F}_t = \sigma(X_s, s \leq t)$. אזי $\{X_t, \mathcal{F}_t, t \geq 0\}$ הוא מרטינגל (סב מרטינגל).

תרגיל 3.31 יהי $\{X_n, \mathbb{B}_n\}$ מרטינגל בזמן בדיד. נגדיר עבור $n \leq t < n+1$, $X_t = X_n,\ \mathbb{B}_t = \mathbb{B}_n$. האם $\{X_t, \mathbb{B}_t, t \geq 0\}$ מרטינגל?

תרגיל 3.32 יהי $\{X_t, \mathbb{B}_t, t \geq 0\}$ סב מרטינגל. אזי הוא מרטינגל אם ורק אם $\mathbb{E}\, X_t = \mathbb{E}\, X_0$

הערה 3.33 ה- σ-שדה $\mathbb{B}_t$ יקרא רציף מימין אם $\mathbb{B}_t = \bigcap\limits_{h>0} \mathbb{B}_{t+h}$.

אם $\mathbb{B}_t$ רציף מימין, אזי למרטינגל X_t יש מודיפיקציה רציפה מימין, כלומר כזו שבהסתברות 1 הדגמים רציפים מימין ($\mathbb{P}\{\omega : t \in [0,T]$ רציף מימין לכל $X_t(\omega)\} = 1$) , ובהסתברות 1 לדגמים יש גבולות משמאל.

אם X_t הוא סב-מרטינגל, ו- $\mathbb{B}_t$ רציף מימין, אזי ל- X_t יש מודיפיקציה רציפה מימין ובעלת גבולות משמאל אם ורק אם הפונקציה $\mathbb{E}\, X_t$ רציפה מימין.

אנו נניח תמיד שאנו עוסקים במודיפיקציה רציפה מימין ובעלת גבולות משמאל. תהליכים כאלו מתוארים בספרות בעזרת קיצורים שונים:

RCLL: Right Continuous, Left Limits

CORLOL: Continuous on Right, Limits on Left

והקיצור הצרפתי (בעל זכות ראשונים) CADLAG: Continue À Droite, Limite À Gauche

בהמשך נניח גם (ללא ציון מפורש) שה- σ-שדה $\mathcal{F}_t$ רציף מימין.

<u>תוספתן:</u> מדידות (שוב?!)

הגדרה 3.34 תהליך X_t יקרא מתואם *Adapted* ל- $\mathcal{F}_t$ אם לכל t, X_t מדיד על $\mathcal{F}_t$. התהליך יקרא *Progressively measurable* אם לכל t הפונקציה $X_s(\omega)$ (כפונקציה של (s,ω)) על הקטע $[0,t]$ מדידה כפונקציה מ- $([0,t] \times \Omega,\ \mathbb{B}([0,t]) \times \mathcal{F}_t)$ ל- $(\mathbb{R}, \mathbb{B})$.

אם X_t רציף ומתואם אזי הוא Progressively measurable**. אנו לא נדקדק במושגי מדידות, ונניח "מדידות כדרוש", למרות שיש כאן נקודות עדינות.**

תרגיל 3.35 יהיו $\{Y_i,\ i = 1, 2, \ldots\}$ מ"א בת"ס, ונניח $\mathbb{E}\,|Y_i| < \infty$.

א. $X_n = \sum\limits_{i=1}^{n} (Y_i - \mathbb{E}\, Y_i)$ הוא מרטינגל.

ב. נניח Y_i שווי פילוג (ובת"ס) נגדיר

$$X_1 =\frac{1}{4}(Y_1+Y_2+Y_3+Y_4) \tag{3.4}$$

$$X_2 =\frac{1}{3}(Y_1+Y_2+Y_3) \tag{3.5}$$

$$X_3 =\frac{1}{2}(Y_1+Y_2) \tag{3.6}$$

$$X_4 =Y_1 \tag{3.7}$$

אזי X_1, X_2, X_3, X_4 מרטינגל. רמז: בגלל סימטריה $\mathbb{E}\left[Y_1 \mid Y_1+Y_2\right] = \mathbb{E}\left[Y_2 \mid Y_1+Y_2\right]$. כדי להוכיח הרמז, הראה כי אם $A \in \mathcal{F}_{Y_1+Y_2}$ אזי

$$\int_A (Y_1-Y_2)dP(\omega) = \int_A (Y_2-Y_1)dP(\omega) \quad (=0)$$

הגדרה 3.36 V_n הוא *Predictable* ביחס ל- B_n אם V_n מדיד B_{n-1} לכל n.

יש מושג מקביל בזמן רציף, שהגדרתו כמובן מורכבת יותר. למזלנו לא נזדקק לו בקורס זה.

הגדרה 3.37 יהי $(X_n, \mathbb{B}_n)$ מרטינגל ו-$\{V_n\}$ סדרת מ"א כך ש- V_n מדיד $\mathbb{B}_{n-1}$ ו- $|V_n| \le C$ בהסתברות 1 (C קבוע). נגדיר

$$Y_n = V_1(X_1-X_0)+V_2(X_2-X_1)+\cdots+V_n(X_n-X_{n-1})$$

Y_n נקרא ה- *Martingale transform* של X ע"י V.

Y_n הוא מרטינגל, וניתן להחליש את דרישת החסימות על $\{v_n\}$---ראה תרגיל. שים לב כי Y_n הוא מעין "אינטגרל" של V_i לפי X_i: בכל צעד נוספת "הקפיצה ב-X_n" כפול ערך V_n. הקשר של מושג המרטינגל ובעית הסינון יובהר בדוגמה הבאה.

תרגיל 3.38 הוכח כי Y_n הוא מרטינגל. הרחב למקרה $\frac{1}{p}+\frac{1}{q}=1$, $\mathbb{E}\,|V_n|^q < \infty$, $\mathbb{E}\,|X_n|^p < \infty$.

דוגמה 3.39 יהי $n(t)$ "רעש לבן", כלומר תהליך עם ערכים בת"ס בזמנים שונים. אזי $Y_t = \int_0^t n(s)\,ds$ הוא מרטינגל (בכל רגע מוסיפים חלק בלתי תלוי - ראה תרגיל לעיל). באופן כללי יותר, אם Y_t מקיים $\mathbb{E}\,Y_t = 0$, ו-$Y_{t+s}-Y_t$ בת"ס ב-Y_ν לכל $\nu \le t$, אזי Y_t הוא מרטינגל בזמן רציף.

על תהליכי מרטינגל ידוע הרבה מאד - ולכן מבנה זה יעזור בחקירת בעיות שיערוך שונות.

דוגמה 3.40 יהי X_n מרטינגל שהוא מספר הניחושים הנכונים (פחות השגויים) בין 0 ל-n בהימור מטבע הוגנת. נניח שברגע n מהמרים V_n התלוי בתוצאות ההטלות <u>עד</u> $n-1$. אזי $(V\cdot X)_n=\sum_1^n V_k(X_k-X_{k-1})$ הוא הסכום המצטבר.

טענה 3.41 אם X_n סופרמרטינגל ו- V_n הוא $Predictable$ חיובי וחסום לכל n, אזי $V\cdot X$ סופרמרטיגל. (כלומר במובן <u>ממוצע</u> לא ניתן להרוויח)!

הוכחה: $(V\cdot X)_n$ **מדיד** $\mathbb{B}_n$. **כיוון ש-** V_n **חסום לכל** n **אזי** $(V\cdot X)_n$ **אינטגרבילי. כעת**

$$\begin{aligned}\mathbb{E}\left[(V\cdot X)_{n+1}\;\middle|\;\mathbb{B}_n\right]&=\mathbb{E}\left[(V\cdot X)_n+V_{n+1}(X_{n+1}-X_n)\mid\mathbb{B}_n\right]\\&=(V\cdot X)_n+V_{n+1}\,\mathbb{E}\left[X_{n+1}-X_n\mid\mathbb{B}_n\right]\le(V\cdot X)_n\end{aligned}$$

כי V_{n+1} **חיובי.**

מסקנה: במובן זה אין אסטרטגיה מנצחת בהימורים!

הערה 3.42 כנ"ל לסב-מרטינגל ולמרטינגל (אם x_n הוא מרטינגל אין צורך להניח ש- V חיובי!).

<u>הפרוק של Doob</u>

תהי $\mathbb{B}_n$ **סדרת** σ **שדות עולה** (**כלומר** $\mathbb{B}_n\subset\mathbb{B}_{n+1}$) **ו-**$X_n$ **סדרת מ"א מדידים על** $\mathbb{B}_n$, **המקיימים** $\mathbb{E}\left|X_n\right|<\infty$.

משפט 3.43 $[Doob]$ X_n ניתן לפרוק $X_n=A_n+Y_n$ כאשר $(Y_n,\mathbb{B}_n)$ מרטינגל, ו- A_{n+1} מדיד על $\mathbb{B}_n$. הפרוק יחיד עד כדי קבוע (הקבוע אקראי אך מדיד $\mathbb{B}_0$).

הוכחה: נגדיר

$$\begin{aligned}
A_0 =& 0\\
A_1 =& A_0 + \mathbb{E}\left[\,-X_0 + X_1 \mid \mathbb{B}_0\right]\\
&\vdots\\
A_n =& A_{n-1} + \mathbb{E}\left[\,-X_{n-1} + X_n \mid \mathbb{B}_{n-1}\right]
\end{aligned}$$

$$\begin{aligned}
Y_0 =& X_0\\
Y_1 =& Y_0 + X_1 - \mathbb{E}\left[X_1 \mid \mathbb{B}_0\right]\\
&\vdots\\
Y_n =& Y_{n-1} + X_n - \mathbb{E}\left[X_n \mid \mathbb{B}_{n-1}\right]
\end{aligned}$$

אזי $Y_n, \mathbb{B}_n$ מרטינגל, A_n מדיד $\mathbb{B}_{n-1}$, ו-

$$Y_n + A_n - X_n = Y_{n-1} + A_{n-1} - X_{n-1} = \cdots = Y_0 + A_0 - X_0 = 0$$

יחידות: אם $\tilde{A}_n, \tilde{Y}_n$ הוא פרוק נוסף, אזי

$$\begin{aligned}
0 = X_n - X_n =& A_n - \tilde{A}_n + Y_n - \tilde{Y}_n\\
A_n - \tilde{A}_n =& \tilde{Y}_n - Y_n
\end{aligned}$$

לכן

$$A_n - \tilde{A}_n = \mathbb{E}\left[A_n - \tilde{A}_n \;\middle|\; \mathbb{B}_{n-1}\right] = \mathbb{E}\left[\tilde{Y}_n - Y_n \;\middle|\; \mathbb{B}_{n-1}\right] = \tilde{Y}_{n-1} - Y_{n-1} = A_{n-1} - \tilde{A}_{n-1}$$

בגלל תכונות המרטינגל. לכן $A_n - \tilde{A}_n = A_0 - \tilde{A}_0 = \tilde{Y}_n - Y_n$.

הערה 3.44 אם $\mathbb{B}_n = \sigma(X_1, \cdots, X_n)$ אזי Y_n נקרא תהליך החידוש *(innovation)* של X_n, כיוון שאם נרשום $X_{n+1} = A_{n+1} + Y_n + (Y_{n+1} - Y_n)$ אזי $Y_{n+1} - Y_n$ "ניצב" ל- $\mathbb{B}_n$ ולכן מתאר את החלק ה"חדש" של X_{n+1}.

תרגיל 3.45 אם במשפט $Doob$ למעלה $\{X_n, \mathbb{B}_n\}$ הוא סב-מרטינגל אזי הסדרה $\{A_n\}$ סדרה עולה.

לפרוק של Doob **יש הקבלה לזמן רציף, כאשר X_t מרטינגל.**

משפט 3.46 (ללא הוכחה: *Doob-Meyer decomposition*). יהי $(X_t, \mathcal{F}_t)$ מרטינגל רציף בהסתברות 1 וכך ש- $\mathbb{E}\, X_t^2 < \infty$. אזי קיים תהליך יחיד A המקיים את התכונות הבאות:

א. $A(0) = 0$

ב. $A(t)$ רציף מימין.

ג. $A(t)$ הוא מתואם (יותר מכך, הוא *Progressively measurable*).

ד. $A(t)$ תהליך עולה בהסתברות 1.

ה. התהליך $X^2(t) - A(t)$ הוא מרטינגל.

הערה 3.47 זהו בעצם פרוק של הסב-מרטינגל $X^2(t)$; השווה לתרגיל בעמוד הקודם. $A(t)$ נקרא התהליך העולה של המרטינגל $X(t)$. בנוסף,

א. לתנועת בראון (אודותיה נלמד בהמשך) $A(t) = t$

ב. הפירוק תקף עבור סב-מרטינגל כללי (לאו דווקא $X^2(t)$), בתנאים טכניים, עם המסקנות א׳-ה׳.

3.3 תכונות של תהליכי מרטינגל

משפט 3.48 [אי שוויון של *Doob*]: יהי $C > 0$.

א. אם $\{X_n, \mathbb{B}_n, n \geq 0\}$ סב-מרטינגל, אזי לכל N,

$$\mathbb{P}\left\{\max_{n \leq N} X_n \geq C\right\} \leq C^{-1}\,\mathbb{E}\,|X_N|$$

ב. אם $\{X_1, X_2, \ldots, X_n, \ldots, X\}$ סב-מרטינגל ביחס ל- $\{\mathbb{B}_1, \mathbb{B}_2, \ldots, \mathbb{B}_n, \ldots, \mathbb{B}\}$, כלומר $\mathbb{E}\left[X \mid \mathbb{B}_n\right] \geq X_n$, אזי

$$\mathbb{P}\left\{\sup_n X_n \geq C\right\} \leq C^{-1}\,\mathbb{E}\,|X|$$

ג. אם $\{X_t, \mathbb{B}_t, t \geq 0\}$ סב-מרטינגל ספרבילי, ו- $0 \leq t \leq T$, אזי

$$\mathbb{P}\left\{\sup_{0\leq t\leq T} X_t \geq C\right\} \leq C^{-1}\,\mathbb{E}\,|X_T|$$

הערה 3.49 ראה דוגמה *3.24* לתהליך המופיע בסעיף ב. השווה לאי שויון מרקוב.
שים לב כי ג $\Leftarrow$ ב $\Leftarrow$ א.

כפי שנראה מתוך ההוכחות, אפשר לקבל חסם מעט הדוק יותר בכל הסעיפים---להחליף את $\mathbb{E}\,|X|$ ב- $\mathbb{E}\,X^+$ בכל הסעיפים---ראה בהמשך.

הוכחה:

א. נגדיר מאורעות

$$A_1 = \{\omega : X_1(\omega) \geq C\}$$

$$A_k = \{\omega : X_i(\omega) < C,\ i < k,\ X_k(\omega) \geq C\} \quad k = 2, 3, \ldots$$

דהיינו A_k **הוא המאורע כי** k **הוא האינדקס** <u>**הראשון**</u> **כך ש-** $X_k \geq C$. **אזי** $A_k \in \mathbb{B}_k$ **ו-** $A_i \bigcap A_j = \emptyset$ **עבור** $i \neq j$ **ו-**

$$A = \bigcup_{k=1}^{N} A_k = \left\{\omega : \max_{n\leq N} X_n \geq C\right\}$$

כעת

$$C \cdot \mathbb{P}\{A\} = C \cdot \sum_{k=1}^{N} \mathbb{P}\{A_k\} = \sum_k \int_{A_k} C dP$$

$$\leq \sum_k \int_{A_k} X_k dP$$ מהגדרת A_k

$$\leq \sum_k \int_{A_k} \mathbb{E}\left[X_N \mid \mathbb{B}_k\right] dP$$ מהגדרת הסב-מרטינגל:

שים לב כי A_k מדיד על $\mathbb{B}_k$,

ולכן מהגדרת תוחלת מותנית $$= \sum_k \int_{A_k} X_N dP \leq \sum_k \int_{A_k} |X_N| dP$$

$$= \int_A |X_N| dP \leq \int_\Omega |X_N| dP = \mathbb{E}\,|X_n|$$

ו-א' הוכח. בשורה השנייה מהסוף אפשר להחליף את $\mathbb{E}\,|X|$ ב- $\mathbb{E}\,X^+$ ולקבל חסם הדוק יותר.

ב. נגדיר $\{B_N = \{\omega : \max(X_1, X_2, \ldots, X_N) \geq C)\}$. כיוון ש- $\{X_1, X_2, \ldots, X_N, X\}$ הוא סב-מרטינגל, נובע מחלק א' כי

$$\mathbb{P}\{B_N\} \leq \mathbb{P}\{\max(X_1, X_2, \ldots, X_N, X) \geq C\} \leq C^{-1}\,\mathbb{E}\,|X|$$

כיוון ש- $B_N \subset B_{N+1}$, נובע מהנחות היסוד של הסתברות כי

$$\mathbb{P}\left\{\bigcup_1^\infty B_N\right\} = \lim_{N\to\infty} \mathbb{P}\{B_N\} \leq C^{-1}\,\mathbb{E}\,|X|$$

אך נשים לב כי $\bigcup_1^\infty B_N$ הוא המאורע שאחד מה- X_n גדול או שווה ל- C, ויתכן שה- $\sup$ שווה ל- C גם אם המאורע לא קרה! לכן נחשב עבור $\varepsilon > 0$

$$\mathbb{P}\left\{\sup_n X_n \geq C\right\} \leq \mathbb{P}\{X_n \geq C - \varepsilon,\ \text{some n}\} \leq \frac{1}{C-\varepsilon}\,\mathbb{E}\,|X|$$

ובכך הוכחנו את ב' כי צד ימין נכון לכל ε חיובי.

ג. תהי I קבוצה צפופה ובת מניה ב- [0,T], למשל הרציונלים.
יהיו I_N קבוצות אינדקסים כך ש- $I_N \subset I_{N+1}$ ו-

$$I_N = \{t_i,\ i = 0, 1, \ldots, N, t_0 = 0, t_i < t_{i+1}, t_N < T\}$$

(כמובן ש- t_i תלוי ב- N) וכך ש- $\bigcup_{N=2}^{\infty} I_N = I - \{T\}$. נגדיר $B_N = \left\{\omega : \max\limits_{i \in I_N} X_i \geq C\right\}$. כיוון ש- $\{X_i,\ i \in I_N\}$ סב-מרטינגל, נפעיל שוב את הטיעון של סעיף ב' ונקבל:

$$\mathbb{P}\left\{\sup_{i \in I} X_i \geq C\right\} \leq C^{-1}\,\mathbb{E}\,|X_T|$$

אבל, כיוון שהתהליך ספרבילי, ו- I צפופה,

$$\mathbb{P}\left\{\sup_{i \in I} X_i \geq C\right\} = \mathbb{P}\left\{\sup_{0 \leq t \leq T} X_t \geq C\right\}$$

תרגיל 3.50 אם $\{X_t, \mathbb{B}_t\}$ מרטינגל ו- ϕ קונוווקסית, אזי $\{\phi(X_t), \mathbb{B}_t\}$ סב-מרטינגל.
אם $\{X_t, \mathbb{B}_t\}$ סב-מרטינגל ו- ϕ קונוווקסית ולא יורדת, אזי $\{\phi(X_t), \mathbb{B}_t\}$ סב-מרטינגל.
(הנח $\mathbb{E}\,|\phi(X_t)| < \infty$).

תרגיל 3.51 יהי $\{|X_t|, \mathbb{B}_t\}$ סב-מרטינגל, $p > 1$ ונניח $\mathbb{E}\,|X_t|^p < \infty$. אזי לכל $C > 0,\ T > 0$

$$\mathbb{P}\left\{\sup_{0 \leq t \leq T} |X_t| \geq C\right\} \leq C^{-p} \int_{\Omega} |X_T|^p \mathbf{1}_{\left\{\sup_{0 \leq t \leq T} |X_t| \geq C\right\}}\, dP \leq \frac{\mathbb{E}\,|X_T|^p}{C^p}$$

<u>רמז:</u> השתמש בתרגיל הקודם ועקוב אחרי הוכחת משפט אי שוויון $Doob$.

משפט 3.52 [אי שוויון $Doob$, שוב]: יהי $p > 1$, $1/p + 1/q = 1$. אם X_n סב-מרטינגל לא שלילי, אזי

$$\mathbb{E}\left[\left(\max_{n \leq N} X_n\right)^p\right] \leq q^p\,\mathbb{E}\left[X_N{}^p\right]$$

ראינו כי האיבר האחרון של מרטינגל (או סב-מרטינגל) נותן מידע רב על המקסימום של התהליך.
נשאלת השאלה אם $\{X_t, \mathbb{B}_t, t \geq 0\}$ מרטינגל, האם קיים X כך ש- $\{X_t, \ldots, X\}$ גם הוא מרטינגל?

משפט 3.53 [משפט ההתכנסות של תהליכי מרטינגל] יהי $\{X_t, \mathbb{B}_t, t \geq 0\}$ סב-מרטינגל ספרבילי. נגדיר $X_t^+ = \max\{X_t, 0\}$, ונניח $C = \sup\limits_t \mathbb{E}\,X_t^+ < \infty$. אזי

א. $X_\infty = \lim\limits_t X_t$ קיים בהסתברות 1 ו- $\mathbb{E}\,|X_\infty| \le \varliminf\limits_t \mathbb{E}\,|X_t|$.

ב. $\{X_t,\ 0 \le t \le \infty\}$ הוא סב-מרטינגל בתנאי ש $X_t \to X_\infty$ ב- L^1.

ג. אם $\sup \mathbb{E}\,|X_t|^p < \infty$ $p > 1,$ אזי X_t מתכנס *w.p.1* וב- L_p.

הערה 3.54 תחת הנחות המשפט, תנאי מספיק להתכנסות ב- L_1 הוא *Uniform Integrability* של $\{X_t\}$.
אם X_t חיובי, טענה ב׳ היא אם ורק אם. חלק ב׳ נותן תנאי לכך ש- $X_t \le \mathbb{E}\left[X_\infty \mid \mathbb{B}_t\right]$.

תרגיל 3.55 יהי $\{X_t, \mathbb{B}_t, t \ge 0\}$ מרטינגל ספרבילי, ונניח $\sup\limits_t |X_t| < \infty$. אזי $X_\infty = \lim\limits_t X_t$ קיים בהסתברות 1, ואם ההתכנסות היא גם ב- L_1 אזי $\mathbb{E}\left[X_\infty \mid \mathbb{B}_t\right] = X_t$.

הערה 3.56 במשפטים אלו, ניתן להחליף את דרישת הספרביליות בדרישה שהתהליך רציף מימין. (כי כמובן רציפות מימין היא דרישה חזקה יותר). עבור תהליכים רבים זוהי דרישה טבעית יותר.

הוכחות ההתכנסות בדרך כלל מסובכת למדי אולם, במקרה מוגבל יותר, ההוכחה פשוטה יחסית.

משפט 3.57 יהי X_n מרטינגל, כך ש- $\mathbb{E}\,X_n^2 \le M < \infty$, אזי קיים מ"א X כך ש- $X_n \underset{a.s.}{\to} X$ ו- $\mathbb{E}\,X^2 < \infty$.

הוכחה: כיוון ש- X_n^2 הוא סב-מרטינגל, הסדרה $\mathbb{E}\,X_n^2$ לא יורדת. בגלל החסימות $\mathbb{E}\,X_n^2 \le M < \infty$, נקבל $\mathbb{E}\,X_n^2 \to \mu \le M < \infty$.
כעת נראה כי X_n היא סדרת קושי בהסתברות 1. לשם כך, מספיק להוכיח כי לכל $0 < p < \infty$ שלם קיים $m = m(p)$ (שלם דטרמניסטי) כך ש- $\left|X_{m(p)+k} - X_{m(p)}\right| \le 1/p$ עבור כל $k = 1, 2, \ldots$ <u>פרט</u> למספר <u>סופי</u> של p (או במילים אחרות, לכל p גדול מספיק. כמובן ש"גדול מספיק" יהיה תלוי ב-ω). כי אז, לכל p גדול מספיק ולכל $l \ge m(p),\ n \ge m(p)$ יתקיים

$$|X_n - X_l| \le |X_n - X_{m(p)}| + |X_l - X_{m(p)}| \le 2/p$$

קל לראות כי לכל m קבוע, $\{X_{m+k} - X_m, \mathbb{B}_{m+k}, k = 1, 2, \ldots\}$ הוא מרטינגל, ולכן $|X_{m+k} - X_m|^2$ הוא סב-מרטינגל. מאי שוויון Doob (חלק א'),

$$\mathbb{P}\left\{\max_{k \le N} |X_{m+k} - X_m|^2 \ge C^2\right\} \le C^{-2}\,\mathbb{E}(X_{m+N} - X_m)^2$$

לכל C ו- N. כעת

$$\mathbb{E}(X_{m+N} - X_m)^2 = \mathbb{E}\, X_{m+N}^2 + \mathbb{E}\, X_m^2 - 2\,\mathbb{E}(X_m\, \mathbb{E}\,[X_{m+N} \mid \mathcal{F}_m]) = \mathbb{E}\, X_{m+N}^2 - \mathbb{E}\, X_m^2$$

כיוון ש- $\mathbb{E}\, X_m^2$ סדרה מתכנסת, ניתן לבחור $C = 1/p$ ו- $m = m(p)$ כך ש-

$$\sup_{N \geq 0} \mathbb{E}(X_{m(p)+N} - X_{m(p)})^2 \leq 2^{-p}$$

ואז

$$\mathbb{P}\left\{\max_{k \leq N} |X_{m(p)+k} - X_{m(p)}|^2 \geq 1/p^2\right\} \leq p^2 2^{-p}$$

וכיוון שזה נכון <u>לכל</u> N,

$$\mathbb{P}\left\{\sup_{k} |X_{m(p)+k} - X_{m(p)}| \geq 1/p\right\} \leq p^2 2^{-p}$$

נפעיל כעת את הלמה של בורל-קנטלי, כיוון ש- $\sum_{p=1}^{\infty} p^2 2^{-p} < \infty$, נובע כי

$$\mathbb{P}\left\{\sup_{k} |X_{m(p)+k} - X_{m(p)}| \geq 1/p\right\} = 1$$

פרט למספר סופי של ערכי p. לכן הסדרה היא קושי, והיא מתכנסת.

דוגמה 3.58 ייצוג של תהליך ברעש.

1. יהי h_t תהליך אקראי על $0 \leq t \leq T$ ונניח שהפונקציה $t \to h_t(\omega)$ "טובה" מספיק כדי למנוע בעיות טכניות (למעשה, נדרוש שהתהליך הוא מדיד, כלומר מדידות של $(\omega, t) \to h_t(\omega)$).
נניח $\mathbb{E} \int_0^T |h_s|\, ds < \infty$. אזי $\int_0^T |h_s|\, ds$ קיים, ולכן גם $\int_0^T h_s\, ds$ קיים, ומתקיים (משפט *Fubini*) $\mathbb{E} \int_0^T h_s\, ds = \int_0^T (\mathbb{E}\, h_s)\, ds$.
נגדיר $y_s = \mathbb{E}\,[h_s \mid \mathbb{B}_1]$ ונניח שוב הנחות טכניות על $s \to y_s(\omega)$.
אזי מהגדרת תוחלת מותנה נובע כי

$$\mathbb{E}\left\{\int_0^T h_s\, ds \middle| \mathbb{B}_1\right\} = \int_0^T \mathbb{E}\,[h_s \mid \mathbb{B}_1]\, ds$$

2. כעת יהי $X_t=\int_0^t h_s\,ds+M_t$ כאשר M_t מרטינגל ביחס ל- $\mathbb{B}_t$ ו- h_t מדיד $\mathbb{B}_t$; זהו מעין "תהליך מדידה ברעש". נסמן $\mathbb{D}_t=\sigma\left\{X_s,\ 0\le s\le t\right\}$, ונסמן את השערוך של h ב-$\hat{h}_t=\mathbb{E}\left[h_t\mid\mathbb{D}_t\right]$.
נבחר "ורסיה טובה" של $\hat{h}$. נציג את X_t ע"י

$$X_t=\int_0^t\hat{h}_s\,ds+\left[\int_0^t(h_s-\hat{h}_s)\,ds+M_t\right]:=\int_0^t\hat{h}_s\,ds+N_t$$

תרגיל 3.59 N_t הוא מרטינגל ביחס ל- $\mathbb{D}_t$.

אם נתייחס ל- M_t כ"רעש מדידה", אזי המשמעות היא שהאות X_t הוא בעל מבנה (אינטגרל + (רעש = מרטינגל)) גם מנקודת ראות של $\mathbb{B}_t$, כלומר כשהכל ידוע, אך גם מנקודת ראות של המדידות $\mathbb{D}_t$.

הערה 3.60 לא בטוח כי $\mathbb{D}_t=\sigma(N_s,\ S\le t)$, למרות שבמקרים רבים אכן מתקיים שוויון.

3.4 זמני עצירה

זמן עצירה stopping time **הוא משתנה אקראי המתאר החלטה המתקבלת על סמך מידע הקיים בזמן העצירה: הוא "מדיד" במובן מתאים. לפני ההגדרה, כמה דוגמאות.**

דוגמה 3.61 שני אנשים משחקים בהטלת מטבע: כאשר יוצא "פלי" זוכה הראשון, וכשיוצא "עץ" זוכה השני. זכיה או הפסד הם של שקל אחד.
נסמן ב-x_t את מספר השקלים שיש לראשון בזמן t, וב-$\mathcal{F}_t$ את המידע שבידיו בזמן t: $\mathcal{F}_t=\sigma\{x_s:s\le t\}$.
נניח שהשחקן הראשון רשאי להחליט מתי להפסיק את המשחק, וברצונו להגיע לרווח ממוצע מרבי בזמן הפסקת המשחק. ברור שההחלטה ברגע t יכולה להסתמך רק על מידע שב-$\mathcal{F}_t$.

ראינו כי $\{x_t,\mathcal{F}_t\}$ **הוא מרטינגל, ולכן** $\mathbb{E}\,x_t=\mathbb{E}\,x_0=x_0$ **לכל זמן** t. **נסמן ב-**$\tau=\tau(\omega)$ **את הזמן (האקראי!) בו השחקן הראשון הפסיק את המשחק. משפט** Doob 3.73 **להלן טוען כי לכל** t, $\{x_{t\wedge\tau},\mathcal{F}_{t\wedge\tau}\}$ **גם הוא מרטינגל (ראה הגדרות בהמשך). לכן** $\mathbb{E}\,x_{t\wedge\tau}=x_0$, **ואם נקח את הגבול כאשר** $t\to\infty$ **נקבל כי** $\mathbb{E}\,x_\tau=x_0$. **כלומר, אין שיטה לעצור את המשחק כך שנרוויח בממוצע!**

דוגמה 3.62 יהי x_t תהליך אקראי המתאר את טמפרטורת האוכל בסיר, כפונקציה של הזמן. החלטות אפייניות שמטרתן למנוע הקדחת התבשיל הן לכבות את האש בזמן τ, כאשר

הטמפרטורה הגיעה ל-90 מעלות: $\tau_1(\omega) = \inf\{t : x_t = 90^o c\}$

האוכל התבשל 10 דקות בטמפרטורה של 90 מעלות: $\tau_2(\omega) = \inf\{t : x_{t-10} = 90^o c\}$

האש כובתה 10 דקות לפני שהטמפרטורה הגיעה ל-90 מעלות: $\tau_3(\omega) = \inf\{t : x_{t+10} = 90^o c\}$.

ברור כי τ_3 אינו מתאר החלטה ברת ביצוע, כיוון שעלינו לנחש את העתיד.

דוגמה 3.63 נגדיר עבור תהליך אקראי $\{x_t\}$

$$\tau_a(\omega) = \inf\left\{t : \int_0^t x_s^2\, ds = 3\right\}$$

אזי ניתן בכל רגע t לענות על השאלה: האם $\tau_a \le t$? אולם אם נגדיר למשל

$$\tau_b = \inf\left\{t : \int_0^{2t} x_s^2\, ds = 3\right\}$$

אזי לא תמיד נוכל לענות על שאלה דומה על סמך התבוננות במסלול $\{x_s,\ s \le t\}$.

הגדרה 3.64 יהיה $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, ו-$\mathcal{F}_t$ סידרה עולה של תת-סיגמה-שדות של $\mathcal{F}$. משתנה אקראי τ נקרא זמן אקראי, או זמן עצירה, אם לכל t מתקיים

$$\{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t \ . \tag{3.8}$$

ה-*stopped sigma-field*, אשר יסומן ב-$\mathcal{F}_\tau$, מוגדר להיות אוסף המאורעות $A \in \mathcal{F}$ המקיימים את התנאי

$$A \cap \{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t \quad for\ all\ t. \tag{3.9}$$

אנו נניח, כפי שמקובל, כי זמני העצירה הם לא שליליים (אלא אם נאמר במפורש אחרת). נשים לב כי בהתאם להגדרה (ובהתייחס לדוגמאות קודמות), τ_1, τ_2, τ_a הם זמני עצירה אך τ_3, τ_b אינם זמני עצירה.

תרגיל 3.65 עבור זמן עצירה τ הראה כי $\mathcal{F}_\tau$ הוא סיגמה-שדה, τ מדיד על $\mathcal{F}_\tau$ ואם $t \equiv \tau$ אזי $\mathcal{F}_\tau = \mathcal{F}_t$.

דוגמה 3.66 יהי A המאורע

$$A = \{\omega : x_s \ge 80^o, some\ 0 \le s \le T\} \tag{3.10}$$

אזי

$$A \cap \{\tau_1 \le t\} = \{\omega : x_s \ge 80^o, some\ 0 \le s \le T,\ x_u \ge 90^o\ some\ 0 \le u \le t\}$$

כעת אם $t \ge T$ אזי התהליך הגיע ל-90^o לפני t ולכן גם הגיע ל-80^o לפני T, כלומר $A \subset \{\tau_1 \le t\}$ ולכן החיתוך הוא ב-$\mathcal{F}_t$. מצד שני אם $t < T$ אזי $A \in F_t$. לכן קיבלנו ש-$A \in \mathcal{F}_{\tau_1}$.

תרגיל 3.67 לכל מספר $\alpha \ge 0$ (דטרמיניסטי!), α הוא זמן עצירה. אם τ הוא זמן עצירה אזי $\tau + \alpha$ הוא זמן עצירה.

משפט 3.68 אם τ_i זמני עצירה אזי גם $\tau_1 + \tau_2$, $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$ הם זמני עצירה.

תרגיל 3.69 הוכח את המשפט. רמז: העזר בקבוצות מהצורה $\{\tau_2 > a\} \cap \{\tau_1 + a > t\}$ עבור a רציונליים.

דוגמה 3.70 הילוך אקראי על השלמים. יהי $x_t = \sum_{i=1}^{t} y_i$ כאשר $\{y_i\}$ משתנים בלתי תלויים ושווי פילוג עם ערכים שלמים. נסמן $\mathcal{F}_t = \sigma\{x_s,\ s \le t\}$ ו-$\tau_a = \min\{t : x_t = a\}$. אזי לכל t,

$$\{\tau_a \le t\} = \{\omega : x_s = a\ some\ s \le t\} = \cup_{s=0}^{t}\{\omega : x_s = a\} \tag{3.11}$$

משפט 3.71 (ללא הוכחה): יהי $\{x_t\}$ תהליך רציף ו-B קבוצת בורל סגורה. אזי

$$\tau_B(\omega) = \inf\{t \ge 0 : x_t(\omega) \in B\} \tag{3.12}$$

הוא זמן עצירה.

הערה 3.72 זמן עצירה יכול להיות שווה ∞ בהסתברות חיובית!

הפונקציה המציינת של זמן עצירה מוגדרת כך:

$$1_\tau(t,\omega) = \begin{cases} 1 & t \le \tau \\ 0 & \text{otherwise} \end{cases} \tag{3.13}$$

זהו תהליך אקראי, והוא שייך ל-$L^2[0,T]$ כפי שנראה מייד. ראשית, מדידות:

$$\{1_\tau(t,\omega)=1\}=\{t\le\tau\}=\{t>\tau\}^c\ .$$

אולם את המשלים של צד ימין ניתן לרשום כך:

$$\{t>\tau\}=\cup_{n=1}^{\infty}\{\tau\le t-1/n\}\in\mathcal{F}_t$$

כאשר השיוויון האחרון נובע מהגדרת זמן עצירה. בנוסף, כמובן

$$\int_0^T 1_\tau^2(t,\omega)\,dt\le T\ .$$

נשים לב שאם $\mathbb{E}\,\tau<\infty$ אזי

$$\mathbb{E}\int_0^\infty 1_\tau^2(t,\omega)\,dt=\mathbb{E}\,\tau<\infty$$

ולכן במקרה זה $1_\tau(t,\omega)\in L^2[0,\infty)$.

משפט 3.73 *(Optional Sampling)*. אם τ הוא זמן עצירה ו-$\{x_t\}$ הוא מרטינגל עם מסלולים רציפים מימין אזי $\{x_{t\wedge\tau}\}$ הוא מרטינגל. אם $\tau\le s$ הם שני זמני עצירה ו-$\{x_t\}$ הוא סבמרטינגל עם מסלולים רציפים מימין על $0\le t\le\infty$ (כלומר כולל "איבר אחרון x_∞) אזי $\mathbb{E}(x_\tau\mid\mathcal{F}_s)\ge x_s$ בהסתברות 1.
את הדרישה שיש איבר אחרון ניתן להחליף בדרישה ש-τ חסום, כלומר קיים מספר a כך ש-$a<\tau$ בהסתברות 1.

תרגיל 3.74 יהיו τ_i זמני עצירה כך ש-$\tau_i\le\tau_{i+1}$ בהסתברות 1, ו-$\{x_t\}$ סבמרטינגל רציף מימין. אם ל-x יש איבר אחרון או לחילופין אם כל זמני העצירה חסומים, אזי $\{x_{\tau_i},\mathcal{F}_{\tau_i}\}$ הוא מרטינגל (בזמן בדיד).

3.5 התנועה הבראונית, או תהליך וינר

הגדרה 3.75 התהליך $(W_t,\ t\in[0,T])$ הוא תנועה בראונית אם

א. W_t ת"א גאוסי.

ב. $\mathbb{E}\,W_tW_s=\min(t,s),\ \mathbb{E}\,W_t=0$.

ג. עבור כמעט כל ω, פונקצית המדגם $t\to W_t(\omega)$ רציפה ב- $[0,T]$.

אם לא נאמר אחרת, נניח תמיד בנוסף כי $W_0 = 0$ (בהסתברות 1).

תרגיל 3.76 לכל $h > 0$ ו- $t \in [0, T-h]$, המ"א W_t והמ"א $W_{t+h} - W_t$ בת"ס.
מכאן נובע כי W_t הוא תהליך "תוספות בלתי תלויות", וכיוון ש- $\mathbb{E}\, W_t = 0$ $W_t,$ מרטינגל.

<u>קיום התנועה</u> הבראונית - בניה

ניתן להראות בצורה מפורשת את קיום התנועה הבראונית - וזאת ע"י בניתה כגבול של תהליכים. ראה למשל בעמ' 56 של Karatzas - Shreve. שיטה אחת היא לבנות תהליך על ידי אינטרפולציה של סכום של משתנים בת"ס ושווי פילוג. ממשפט הגבול המרכזי אוסף זה מתכנס למ"א גאוסי, ואפשר להראות כי כל הפילוגים הרב ממדיים של האינטרפולציה מתכנסים לפילוג גאוסי רב ממדי, עם הפרמטרים הרצויים. מכאן ניתן להראות התכנסות חלשה לתהליך גאוסי שהוא תנועת בראון. אנו נבחר בשיטת קירוב אחרת. תהי $\{\phi_i(t)\}$ משפחה אורתונורמלית שלמה ב- $L_2[0,T]$ ויהיו X_i מ"א גאוסיים בת"ס ומנורמלים ($\mathbb{E}\, X_i = 0,\ var X_i = 1$). נגדיר

$$V_t^N = \sum_{i=1}^{N} X_i \int_0^t \phi_i(s)\, ds$$

ברור כי V_t^N ת"א גאוסי. נראה כעת כי V_t^N סדרת קושי ב- $L_2(\Omega \times [0,T])$ (כלומר $\int_0^T \mathbb{E}\left(V_t^n - V_t^m\right)^2 dt \to 0$), ולכן יש לה גבול שיסומן V_t. נניח ש- $m > n$.

$$\begin{aligned}
\mathbb{E}\left(V_t^n - V_t^m\right)^2 &= \sum_{i=n+1}^{m} \sum_{j=n+1}^{m} \mathbb{E}\, X_i X_j \int_0^t \phi_i(s)\, ds \int_0^t \phi_j(s)\, ds \\
&= \sum_{i=n+1}^{m} \left(\int_0^t \phi_i(s)\, ds \right)^2
\end{aligned}$$

בגלל ש- $\mathbb{E}\, X_i X_j = \delta_{ij}$. כעת נשים לב כי את האנטגרל ניתן לייצג כמכפלה פנימית

$$\int_0^t \phi_i(s)\, ds = \int_0^T \mathbf{1}_{\{s \le t\}} \phi_i(s)\, ds = \left(\mathbf{1}_{\{s \le t\}},\ \phi(s)\right)$$

וממשפט פרסוול ב- $L_2[0,T]$,

$$\int_0^T \left(\mathbf{1}_{\{s \le t\}}\right)^2 ds = t = \sum_{i=1}^{\infty} \left(\mathbf{1}_{\{s \le t\}},\ \phi_i(s)\right)^2 \le T$$

לכל $t \le T$. לכן הסכום מ- $n+1$ שואף לאפס, וממשפט ההתכנסות החסומה

$$\int_0^T \mathbb{E}\left(V_t^n - V_t^m\right)^2 dt \le \int_0^T \sum_{i=n+1}^{\infty} \left(\mathbf{1}_{\{s\le t\}},\ \phi_i(s)\right)^2 dt \underset{n\to\infty}{\to} 0$$

לכן ניתן לרשום $V_t = \sum_{i=1}^{\infty} X_i \int_0^t \phi_i(s)\, ds$.

נבדוק כעת את תכונות התהליך הגבולי - V_t. מתוצאות כלליות על פילוגים גאוסיים נובע כי V_t, שהוא גבול של ת"א גאוסיים, גם הוא גאוסי. כיוון ש- $\mathbb{E}\, V_t^N = 0$, ולכל N, $\sup_{0\le t\le T} \mathbb{E}\left(V_t^N\right)^2 < \infty$, גם $\mathbb{E}\, V_t = 0$.

נחשב את האוטוקורלציה. נשתמש שוב ב- $\mathbb{E}\, X_i X_j = \delta_{ij}$ ובמשפט פרסוול:

$$\begin{aligned}
\mathbb{E}(V_t V_s) &= \mathbb{E} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} X_i X_j \int_0^t \phi_i(\alpha)\, d\alpha \int_0^s \phi_j(\beta)\, d\beta \\
&= \sum_{i=1}^{\infty} \int_0^t \phi_i(\alpha)\, d\alpha \int_0^s \phi_i(\beta)\, d\beta = \sum_{i=1}^{\infty} \left(\mathbf{1}_{\{\alpha\le t\}}, \phi_i(\alpha)\right)\left(\mathbf{1}_{\{\beta\le s\}}, \phi_i(\beta)\right) \\
&= \left(\mathbf{1}_{\{\alpha\le t\}}, \mathbf{1}_{\{\alpha\le s\}}\right) = \int_0^T \mathbf{1}_{\{\alpha\le t\}}\, \mathbf{1}_{\{\alpha\le s\}}\, d\alpha = \min(t,s)
\end{aligned}$$

כאשר את החלפת התוחלת והסכום ניתן להצדיק משיקולי קיום מומנטים. כדי לקבל מודיפיקציה עם פונקציות מדגם רציפות, נשתמש במשפט קולמוגורוב. ממשפט Doob נוכל להראות שהתהליך רציף בהסתברות. נבחר ורסיה ספרבילית כדי להפעיל את משפט קולמוגורוב. נזכר כי למ"א גאוסי עם $\mathbb{E}\, Y = 0$, $\mathbb{E}\, Y^2 = \sigma^2$, מתקיים

$$\mathbb{E}\, e^{uY} = e^{u^2\sigma^2/2}$$

$$\mathbb{E}\, Y^{2n} = \frac{(2n)!}{2^n n!}(\sigma^2)^n$$

נבחר $\alpha = 4$. $V_{t+h} - V_T$ גאוסי, ממוצע אפס.

$$\mathbb{E}(V_{t+h} - V_t)^2 = \mathbb{E}(V_{t+h} - V_t)V_{t+h} - \mathbb{E}(V_{t+h} - V_t)V_t = t + h - t + t - t = h$$

כאשר השתמשנו ב- $\mathbb{E}(V_t V_s) = \min(t,s)$.

קיבלנו כי $V_{t+h} - V_t$ גאוסי, ממוצע אפס, ווריאנס h, ולכן

$$\mathbb{E}(V_{t+h} - V_t)^4 = \frac{4!}{4\cdot 2}\cdot h^2 = 3h^2$$

ממשפט קולמוגורוב ($\alpha = 4, \beta = 1$) נובע כעת קיום ורסיה רציפה.
עוד על שאלת הרציפות של התנועה הבראונית - בהמשך.

משפט 3.77

א. אם W_t היא תנועה בראונית, אזי $\left\{{W_t}^2 - t,\ \sigma(W_s, s \leq t)\right\}$ הוא מרטינגל. דהיינו, התהליך העולה של התנועה הבראונית הוא t.

ב. *(P. Levy)* אם $\{X_t,\ \sigma(X_s, s \leq t)\}$ מרטינגל בעל דגמים רציפים ו- $\left\{{X_t}^2 - t,\ \sigma(X_s, s \leq t)\right\}$ הוא מרטינגל, אזי X_t תנועה בראונית.

תרגיל 3.78 הוכח את א.

לתנועה הבראונית (וכן לכל מרטינגל בעל דגמים רציפים) מסלולים רציפים, אך לא חלקים במיוחד. מדד לכך היא התנודה הריבועית.

<u>התנודה הריבועית</u>

נקבע קטע $[a, b]$, $0 \leq a < b < \infty$. תהי π^n חלוקה סופית של הקטע:
$\pi^n = \left[a = t_0^{(n)} < t_1^{(n)} < \cdots < t_k^{(n)} = b\right]$ ונסמן $\delta^n = \max_i \left(t_{i+1}^{(n)} - t_i^{(n)}\right)$.
עבור תהליך כלשהו X_t, נסמן

$$S_n = \sum_i \left(X_{t_{i+1}}^{(n)} - X_{t_i}^{(n)}\right)^2$$

שים לב כי אם ל- X_t דגמים רציפים Lipschitz (כלומר $|X_t - X_s| \leq L(t-s)$, כאשר L יכול להיות אקראי), אז

$$|S_n| = \sum_i \left|X_{t_{i+1}}^{(n)} - X_{t_i}^{(n)}\right|^2 \leq \sum_i L^2 \left|t_{i+1}^{(n)} - t_i^{(n)}\right|^2 \leq$$
$$\leq L^2 \delta^{(n)} \sum \left|t_{i+1}^{(n)} - t_i^{(n)}\right| = L^2 \delta^{(n)}(b-a)$$

ולכן אם $\delta^{(n)} \to 0$, אזי בהכרח $S_n \overset{a.s.}{\to} 0$.
התנודה הריבועית היא הגבול של S_n כאשר $\delta^{(n)} \to 0$, אם הגבול קיים. באופן כללי ניתן להראות שהגבול קיים עבור תהליכי מרטינגל ב-$L^2[0,T]$. למעשה, לתהליכים כאלו ההגדרה המודרנית של התנודה הריבועית היא התהליך העולה (ראה Doob-Meyer decomposition). נראה זאת עבור תנועת בראון.

משפט 3.79 עבור התנועה הבראונית, אם $\delta^{(n)}\to 0$ כאשר $n\to\infty$ אזי $S_n \xrightarrow{q.m.} (b-a)$.

המשפט טוען כי עבור התנועה הבראונית, התנודה הריבועית היא בדיוק t - הזמן, שהוא גם (ראה משפט Levy לעיל) התהליך העולה. מכאן נובע כי לתנועה הבראונית <u>אין</u> דגמים רציפים Lipschitz.

הוכחה: $\mathbb{E}\,S_n = b-a$ לכל n. נחשב וריאנס של $S_n-(b-a)$:

$$\mathbb{E}[S_n-(b-a)]^2 = \mathbb{E}\,{S_n}^2-(b-a)^2$$

לחישוב האיבר הראשון, נשתמש בעובדה כי המומנט הרביעי של משתנה גאוסי שווה ל-3 כפול המומנט השני בריבוע:

$$\begin{aligned}
\mathbb{E}\,{S_n}^2 &= \mathbb{E}\sum_i\sum_j \left(W^{(n)}_{t_{i+1}}-W^{(n)}_{t_i}\right)^2\left(W^{(n)}_{t_{j+1}}-W^{(n)}_{t_j}\right)^2 = \\
&= \mathbb{E}\sum_i \left(W^{(n)}_{t_{i+1}}-W^{(n)}_{t_i}\right)^4 + \mathbb{E}\sum_{i\neq j}\left(W^{(n)}_{t_{i+1}}-W^{(n)}_{t_i}\right)^2\left(W^{(n)}_{t_{j+1}}-W^{(n)}_{t_j}\right)^2 = \\
&= 3\sum_i\left(t^{(n)}_{i+1}-t^{(n)}_i\right)^2+\sum_{i\neq j}\left(t^{(n)}_{i+1}-t^{(n)}_i\right)\left(t^{(n)}_{j+1}-t^{(n)}_j\right) = \\
&= 2\sum_i\left(t^{(n)}_{i+1}-t^{(n)}_i\right)^2+\left[\sum_i\left(t^{(n)}_{i+1}-t^{(n)}_i\right)\right]^2 = \\
&= 2\sum_i\left(t^{(n)}_{i+1}-t^{(n)}_i\right)^2+(b-a)^2
\end{aligned}$$

ולכן

$$\mathbb{E}\left[S_n-(b-a)\right]^2 \le 2\delta^{(n)}(b-a)$$

הגדרה 3.80 <u>התנודה</u> של פונקציה f, שתסומן $V(f)$ מוגדרת

$$V(f)=\sup\sum_i\left|f\left(t^{(n)}_{i+1}\right)-f\left(t^{(n)}_i\right)\right|$$

כאשר $\left\{t^{(n)}_i\right\}$ היא חלוקה כלשהי, וה- sup נלקח על פני כל החלוקות π^n של האנטרוול $[a,\ b]$.

תרגיל 3.81

א. חשב את התנודה של פונקציה מונוטונית.

ב. מצא חסם עבור התנודה של פונקציה, אשר ניתן לבטאה כהפרש של שתי פונקציות מונוטוניות.

ג. הראה עבור התנועה הבראונית כי אם בוחרים סדרת חלוקות π^n כך ש- $\delta^{(n)} = 2^{-n}$, אזי $S_n \overset{a.s.}{\to} t$.
רמז - העזר במשפט בורל-קנטלי ובצ'בישב.

טענה 3.82 <u>התנודה</u> של התנועה הבראונית אין-סופית.

הוכחה: נסמן $\alpha(h) = \sup_t |W_{t+h} - W_t|$ כאשר $a \le t \le b-h$. כיוון של- W_t דגמים רציפים, הדגמים רציפים במידה שווה בקטע $[a,b]$ ולכן $\alpha(h) < \infty$ בהסתברות 1 ($\alpha(h)$ הוא מ"א!), ובנוסף $\alpha(h) \overset{a.s.}{\to} 0$ כאשר $h \to 0$.
נחלק את $[a,b]$ ל- 2^n חלקים שוים, אזי $\delta^{(n)} = (b-a)2^{-n}$, ו-

$$S_n = \sum_{0}^{2^n-1} \left(W_{(i+1)\delta^{(n)}} - W_{i\delta^{(n)}}\right)^2 \le \alpha\left(\delta^{(n)}\right) \sum_{0}^{2^n-1} \left|W_{(i+1)\delta^{(n)}} - W_{i\delta^{(n)}}\right|$$

אבל מהתרגיל $S_n \overset{a.s.}{\to} t$ וכן $\alpha\left(\delta^{(n)}\right) \overset{a.s.}{\to} 0$, לכן בהכרח

$$V(W) \ge \sum_{0}^{2^n-1} \left|W_{(i+1)\delta^{(n)}} - W_{i\delta^{(n)}}\right| \to \infty$$

ניתן גם להראות (ראה למשל Breiman) כי התנועה הבראונית אינה גזירה בשום נקודה באינטר-וול, בהסתברות 1.

תרגיל 3.83

א. תהי W_t תנועה בראונית. אזי $\sqrt{C}W(t/C)$ ו- $tW(1/t)$ תנועות בראוניות.

ב. אם $W_t^{(i)}$ תנועות בראוניות בלתי תלויות, אזי $\alpha \sum_i W_t^{(i)}$ תנועה בראונית (חשב את α)!

ג. $\sum_i \left(W^{(1)}_{t^{(n)}_{i+1}} - W^{(1)}_{t_i{(n)}}\right)\left(W^{(2)}_{t^{(n)}_{i+1}} - W^{(2)}_{t_i{(n)}}\right) \overset{q.m.}{\to} 0$ כאשר $\delta(n) \to 0$.

הערה 3.84 עד כמה התנועה הבראונית רציפה אפשר להסיק ממשפט הבא.

משפט 3.85 [ראה *Karatzas & Shreve* עמ' 114 *Levy Modulus*]
נגדיר עבור $\delta > 0$, $g(\delta) = \sqrt{2\delta \log(1/\delta)}$ אזי

$$\mathbb{P}\left\{\overline{\lim_{\delta\downarrow 0}}\frac{1}{g(\delta)} \quad \max_{\substack{0\le s\le r\le 1\\ t-s\le\delta}} |W_t - W_s| = 1\right\} = 1$$

<u>מסקנה:</u> $|W_t - W_s| \le Cg(\delta)$ **לכל** $C > 1$ **ולכל** δ **קטן מספיק (התלוי ב-** ω **!) ו-** $\delta > |t-s|$, **זהו "מודולוס הרציפות" של התנועה הבראונית.**

משפט 3.86 [*Law of Iterated log* עמ' 112 ב- *K&S*]:
כמעט לכל ω מתקיים

$$\overline{\lim_{t\downarrow 0}}\frac{W_t(\omega)}{\sqrt{2t\log\log 1/t}} = \overline{\lim_{t\to\infty}}\frac{W_t(\omega)}{\sqrt{2t\log\log t}} = 1$$

$$\underline{\lim_{t\downarrow 0}}\frac{W_t(\omega)}{\sqrt{2t\log\log 1/t}} = \underline{\lim_{t\to\infty}}\frac{W_t(\omega)}{\sqrt{2t\log\log t}} = -1$$

תכונות מפורטות של התנועה הבראונית ניתן למצוא ב- K&S במיוחד, נציין בקצרה: (ללא הוכחות) כמעט לכל ω

1. **התנועה הבראונית אינה מונוטונית על שום אינטרוול (כי אחרת התנודה היתה סופית ואז התנודה הריבועית היתה אפס).**
2. **כל נקודות המקסימום הן מקסימום ממש, ואוסף נקודות המקסימום בן מניה וצפוף ב-** $[0,\infty)$.
3. **אוסף הנקודות ב- t בהן** $W_t(\omega) = 0$ **סגור, בעל מידת לבג אפס, בעל נקודת הצטברות באפס, וללא נקודות מבודדות ולכן צפוף בעצמו.**

המסקנה, כפי שכבר ראינו מתכונות התנודה הריבועית, היא שהתנועה הבראונית פרועה מאוד במובן שהיא עולה ויורדת ("משנה כיוון") במהירות רבה.

וכמובן חוק המספרים הגדולים נובע מיד מ- LIL.

משפט 3.87 $\lim_{t\to\infty} \frac{W_t}{t} = 0$ בהסתברות 1.

תרגיל 3.88 תהי $W(t)$ תנועה בראונית, אזי $\exp\left(W(t) - \frac{t}{2}\right)$ מרטינגל.

תרגיל 3.89 יהי U_t תהליך גאוסי סטציונרי, בעל דגמים רציפים בהסתברות 1, ו- $\mathbb{E}\, U_t = 0$. אם $\mathbb{E}(U_{t+h}U_t) = e^{-\beta|h|}$, אזי $V_t = t^{1/2} U\left(\frac{1}{2\beta}\log t\right)$ תנועה בראונית.

4 אנליזה סטוכסטית

בפרק זה נגדיר מושגים של אינטגרל ונגזרת: כאלו המתאימים לתהליכים סטוכסטיים, ואשר יאפשרו לנו להרחיב את המושג של משוואות דיפרנציאליות רגילות למקרה בו יש רעש המתואר על ידי תנועת בראון.

4.1 האינטגרל הסטוכסטי

לאינטגרל מהצורה $\int_0^t X_s\,dW_s$ אין משמעות במסגרת האינטגרציה הקלאסית. זאת משום שכפי שראינו, ל- W_s <u>אין</u> השתנות חסומה, ולכן לא ניתן להגדיר אינטגרל כזה דרך הגישה של רימן, לבג, או סטילצ'ס-לבג. גם בנסיונות קירוב נתקל בקשיים: אם נקרב את הפונקציה $f(t,x)$ לטור טיילור סביב אפס, הקרוב הראשון יהיה

$$f(t,x)\approx f(0,0)+f'_t(0,0)\cdot t+f'_x(0,0)\cdot x$$

אולם אם ננסה לקרב את $f(t,W_t)$, הקרוב הראשון ישצטרך להיות

$$f(r,W_t)\approx f(0,0)+f'_t(0,0)\cdot t+f'_x(0,0)\cdot W_t+\frac{1}{2}f''_x(0,0)\cdot {W_t}^2$$

כי ראינו ש- ${W_t}^2$ "מתנהג בקירוב" כמו t!

בהמשך ננסה לבסס חוקי גזירה ואינטגרציה המתאימים לתנועה בראונית.

יהי $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$ מרחב הסתברות עם "התפתחות של מאורעות" $\mathcal{F}_t$. הסדרה $\mathcal{F}_t$ נקראת Filtration, והיא נדרשת לקיים $s\le t,\mathcal{F}_s\subset\mathcal{F}_t\subset\mathcal{F}$, וכן רציפות מימין: $\mathcal{F}_t=\bigcap_{\epsilon>0}\mathcal{F}_{t+\epsilon}$. תהי W_t תנועה בראונית סטנדרטית על $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$, ובפרט לכל $t<t_1<t_2<\ldots$, ההפרשים $\{W_{t_{i+1}}-W_{t_i}\}$ בת"ס ב- $\mathcal{F}_t$.

את האינטגרל הסטכסטי נבנה בשלבים: תחילה עבור אינטגרנדים פשוטים ביותר, ובהמשך נסלק חלק מהמגבלות. האינטגרל בו נדון נקרא האינטגרל של Itô.

כדי לקבל אינטגרלים המוגדרים על $[0,\infty)$, אפשר להתחיל בסידרת קירובים - שהם אינטגרלים המוגדרים בתחום $[0,T]$, ואז להגדיל את T.

הגדרה 4.1 תהליך אקראי X_t $(0 \le t \le T)$ יקרא <u>פשוט</u> אם קיימת חלוקה דטרמיניסטית $0 = s_0 < s_1 < s_2 < \cdots < s_n = T$ של הקטע $[0,T]$ ומ"א $\{\xi_i\}$ כך ש-

א. $X_t = \xi_i$ על הקטע $s_i \le t < s_{i+1}$

ב. ξ_i מדיד ביחס ל- $\mathcal{F}_{s_i}$ (ולכן X_t מדיד $\mathcal{F}_t$ לכל t),

ג. $\mathbb{E}\, {X_t}^2 \le K$, לכל t.

לדוגמה, אם Ψ פונקציה דטרמיניסטית חסומה, ואם $[t]$ הוא השלם הגדול ביותר החסום ע"י t ($[t] \le t$), אזי $\Psi\left(W_{[t]}\right)$ הוא ת"א פשוט.

תרגיל 4.2 תהי Ψ פונקציה דטרמיניסטית חסומה. האם $\Psi\left(W_{[kt]/k}\right)$ ת"א פשוט? האם $\Psi\left(W_{[t]+1}\right)$ ת"א פשוט?

אוסף התהליכים הפשוטים יסומן ב- E. נסמן ב- L^2 את אוסף התהליכים X_t כך ש-

א. X_t מדיד $\mathcal{F}_t$ לכל t (תהליך מתואם---adapted), ולכן בת"ס בהפרשי W_t אחרי t.

ב. $\mathbb{E} \int_0^T {X_t}^2 dt < \infty$.

נשים לב כי מההגדרה, $E \subset L^2$. המשפחה E ודאי אינה מעניינת לכשעצמה. אולם נגדיר עבורה אינטגרל סטוכסטי, ואח"כ נכלילו ל- L^2 דרך המשפט הבא (ראה למשל Lipster-Shiryaev, עמ' 92-95).

משפט 4.3 לכל $X \in L^2$ קיימת סדרת ת"א $X_n \in E$ כך ש-

$$\mathbb{E} \int_0^T \left(X_n(t) - X(t)\right)^2 dt \underset{n \to \infty}{\longrightarrow} 0$$

הגדרה 4.4 עבור ת"א $X \in E$, נגדיר את אינטגרל איטו על $[0,t]$ ע"י *martingale transform* להלן, כאשר W_t היא תנועה בראונית ביחס ל- $\mathcal{F}_t$. אם $0 = s_0 < s_1 < \cdots < s_k \le t \le s_{k+1}$ הן נקודות אי הרציפות של X,

$$\int_0^t X(s,\omega)\, dW(s,\omega) = \sum_{i=0}^{k-1} X(s_i,\omega)(W(s_{i+1},\omega) - W(s_i,\omega)) + X(s_k,\omega)(W(t,\omega) - W(s_k,\omega))$$

שים לב כי כל איבר בסכום הוא מכפלה של מ"א בת"ס. לחילופין, ניתן היה גם להשתמש למשל בערך $X\left(\frac{1}{2}(s_{i+1}+s_i)\right)$. אולם במקרה זה תכונת אי התלות היתה אובדת, ובנגוד לתורת האינט-גרציה הקלאסית (רימן), כאן ההבדל בהגדרות ניכר, כפי שנראה בהמשך.

תרגיל 4.5 הוכח כי עבור ת"א פשוט נתון, הגדלת מספר נקודות החלוקה בהגדרת האינטגרל הסטוכסטי, מעבר למספר נקודות "הקפיצה" שלו, לא תשנה את ערך האינטגרל.

תכונות האינטגרל הסטוכסטי על E: לכל $0 \le a \le b \le T$ מתקיים

1. התוחלת של אינטגרל סטוכסטי היא אפס: $\mathbb{E}\int_a^b X_t\,dW_t = 0$
2. האינטגרל הסטוכסטי הוא לינארי: עבור α ו- β ממשיים,

$$\int_a^b (\alpha X_1(t) + \beta X_2(t))\,dW_t = \alpha\int_a^b X_1(t)\,dW_t + \beta\int_a^b X_2(t)\,dW_t$$

3. המומנט השני של איטגרל סטוכסטי שווה לתוחלת של אינטרגל רגיל של התהליך בריבוע:

(א) $$\mathbb{E}\left(\int_a^b X_t\,dW_t\right)^2 = \mathbb{E}\int_a^b X_s^2\,ds < \infty$$

כאשר האינטגרל הימני הוא לבג.

(ב) $$\mathbb{E}\left\{\int_a^b X_1(t)\,dW_t \cdot \int_a^b X_2(t)\,dW_t\right\} = \mathbb{E}\int_a^b X_1(t)X_2(t)\,dt$$

נסמן $I_t = \int_a^t X(s)\,dW_s$

4. לת"א I_t דגמים רציפים בהסתברות 1.
5. התנודה הריבועית של I_t בתחום $[a,b]$ היא $\int_a^b X_s^2\,ds$.
6. I_t הוא $\mathcal{F}_t$ מרטינגל, כלומר

$$\mathbb{E}\left\{\int_t^{t+h} X_s\,dWs \mid \mathcal{F}_t\right\} = 0, \quad \mathbb{E}\left\{\int_0^{t+h} X_s\,dWs \mid \mathcal{F}_t\right\} = \int_0^t X_s\,dW_s$$

תרגיל 4.6 הוכח תכונות $1-6$. הראה עבור $0 < u < t \leq T$

$$\int_0^t X_s\,dW_s = \int_0^u X_s\,dW_s + \int_u^t X_s\,dW_s = \int_0^t X_s\mathbf{1}_{\{s\leq u\}}\,dW_s + \int_u^t X_s\,dW_s$$

אם X_s פונקציה דטרמיניסטית ב- E, אזי I_t ת"א גאוסי עם פונקצית אוטוקורלציה $R_{s,t} = \int_0^{\min(s,t)} X_u^2\,du$

(רמז: הראה $\int_0^t X_s\,dW_s = \int_0^T \mathbf{1}_{\{s\leq t\}} X_s\,dW_s$ והפעל את (3-ב') על $X_n\mathbf{1}_{\{u\leq s\}}$ ו- $X_u\mathbf{1}_{\{u\leq t\}}$).

הרחבת האינטגרל ל- L^2**. יהי** $X \in L^2$**. אזי קיימת סדרה** $X_n \in E$ **כך ש-**

$$\|X_n - X\|_2^2 = \mathbb{E}\int_0^T (X_n(t) - X(t))^2\,dt \longrightarrow 0.$$

הערה 4.7 הקירוב נעשה בד"כ בשני שלבים.

1. נרחיב לפונקציות בעלות מסלולים רציפים בממוצע ריבועי, באינטרוול הנתון, ואז הקירוב נעשה ע"י דגימה.
2. נקרב תהליך כלשהוא ב- L^2 על ידי החלקה - או מעבר במסנן מעביר נמוכים:

$$X_m(t) = m\int_{t-1/m}^t X(t)\,dt$$

הגדרה 4.8 יהי $X \in L_2[0,T]$. תהי $\{X_n\}$ סדרה ב-E כך ש-$\|X_n - X\|_2 \to 0$. האינטרגל של X על האינטרוול $[a,b]$ מוגדר על ידי הגבולב - L_2 של המ"א

$$\int_a^b X(t)\,dW_t = \lim_{n\to\infty}\int_a^b X_n(t)\,dW_t\ .$$

כדי שההגדרה תהיה בעלת משמעות, צריך להראות שיש התכנסות (בממוצע ריבועי) לגבול כלשהוא, וכן שהגבול יחיד ואינו תלוי בבחירת הסדרה המקרבת $\{X_n\}$.

כוון ש - $\|X_n - X\|_2 \to 0$ נובע ש- $\{X_n\}$ סדרת קושי, כלומר $\|X_n - X_m\|_2 \to 0$. לכן

$$\begin{aligned}
\mathbb{E}\left(\int_a^b X_n(s)\,dW_s - \int_a^b X_m(s)\,dW_s\right)^2 &= \mathbb{E}\left(\int_a^b (X_n(s) - X_m(s))\,dW_s\right)^2 \\
&= \mathbb{E}\int_a^b (X_n(s) - X_m(s))^2\,ds \\
&\le \mathbb{E}\int_0^T (X_n(s) - X_m(s))^2\,ds \to 0
\end{aligned}$$

ולכן האינטגרלים מהווים סדרת קושי של משתנים אקראיים, וקיים הגבול לכל $0 \le a \le b \le T$. הגבול אינו תלוי בסדרה המקרבת, שכן אם X_n ו- Y_n הן שתי סדרות מקרבות, אזי בדומה לחישוב האחרון,

$$\begin{aligned}
&\mathbb{E}\left(\int_a^b X_m(s)\,dWs - \int_a^b Y_n(s)\,dWs\right)^2 \\
&\le \mathbb{E}\int_0^T (X_m(s) - Y_n(s))^2\,ds \\
&= \mathbb{E}\int_0^T ((X_m(s) - X(s)) - (Y_n(s) - X(s)))^2\,ds \\
&\le 2\,\mathbb{E}\int_0^T (X_m(s) - X(s))^2\,ds + 2\,\mathbb{E}\int_0^T (Y_n(s) - X(s))^2\,ds
\end{aligned}$$

ושני הבטויים האחרונים שואפים לאפס שכן הסדרות X_m ו- Y_n מקרבות את X.

<u>תכונות האינטגרל הסטוכסטי</u> על L^2.

משפט 4.9 האינטגרל הסטוכסטי מקיים תכונות 1-6 לעיל.

את ההוכחות ניתן בתרגילים להלן, אך לא את כולן.

תרגיל 4.10 אם $Y_n \stackrel{q.m}{\to} Y$ אזי $Y_n \to Y$ גם במומנט ראשון, ו- $\mathbb{E}\,Y_n \to \mathbb{E}\,Y$. כעת הוכח את 1.

תרגיל 4.11 פעולת החיבור היא פעולה רציפה ב- L^2 במובן הבא: אם $X_n \stackrel{q.m}{\to} X$ ו- $Y_n \stackrel{q.m}{\to} Y$, אזי $X_n + Y_n \stackrel{q.m}{\to} X + Y$. כעת הוכח את 2.

סעיף 3 נובע מהוכחת הקיום לעיל, כי $\mathbb{E}\left(\int X\,dW\right)^2 = \lim \mathbb{E}\left(\int X_n\,dW\right)^2$.

תרגיל 4.12 הוכחת 4. נקבע סדרה $X^{(n)} \to X$, $X^{(n)} \in E$, ונגדיר: $M_t{}^{n,m} = \int_0^t \left(X_s{}^{(n)} - X_s{}^{(m)}\right) dW_s$ אזי $M_t{}^{n,m}$ הוא מרטינגל רציף. ידוע כי למרטינגל $CADLAG$ מתקיים $\mathbb{E}\,|\sup\limits_{0\le t\le T} M_t|^p \le q^p \sup\limits_{0\le t\le T} \mathbb{E}\,|M_t|^p$ כאשר $\frac{1}{p}+\frac{1}{q}=1$, $q,p>0$. בחר תת סדרה $X_t{}^{(n_i)}$ והעזר בבורל קנטלי כדי לקבל התכנסות יוניפורמית על $[0,T]$ של $\int_0^t X_s{}^{(n)}\,dW_s$ לגבול $\int_0^t X_s\,dW_s$. מההתכנסות היוניפורמית נובעת הרציפות.

את תכונה 5 לא נוכיח, נעיר רק כי התורה המודרנית של אינטגרציה סטוכסטית מתבססת על תכונות תנודה ריבועית, ומגדירה אינטגרל סטוכסטי בצורה מופשטת יותר דרך תכונות התנודה הריבועית.

תרגיל 4.13 הוכח תכונה 6. רמז: חשב

$$\mathbb{E}\left\{\left[\mathbb{E}\left(\int_t^{t+h} X_s\,dWs - \int_t^{t+h} X_s^n\,dW_s \mid \mathcal{F}_t\right)\right]^2\right\}$$

כאשר X_t^n סדרת קרובים.

דוגמה 4.14 אינטגרל איטו מול אינטגרל דטרמניסטי.
יהי $X_t = W_t$ תנועה בראונית. אזי $X = W \in L^2$. נחשב את האינטגרל $\int_0^T W_s\,dW_s$ בעזרת סדרת קירובים. נשווה למקרה הקלאסי: אם Y_t היא פונקציה "חלקה" (למשל עם נגזרת רציפה) אזי אינטגרציה בחלקים תתן,

$$\int_0^T Y_s\,dY_s = Y_s \cdot Y_s \Big|_0^T - \int_0^T Y_s\,dY_s$$

או

$$\int_0^T Y_s\,dY_s = \frac{1}{2}\left(Y^2(T) - Y^2(0)\right) \tag{4.1}$$

כידוע לנו, לתנועת בראון אין נגזרת, ולכן נלך בדרך הקשה של קרובים: נחלק את $[0,T]$ ל- 2^n קטעים, ונסמן $t_i^{(n)} = i2^{-n}T$. נגדיר

$$X_t^n(\omega) = W_{t_i^{(n)}}(\omega),\ t_i^{(n)} \le t < t_{i+1}^{(n)}$$

אזי

$$\mathbb{E}\int_0^T \left(W_t - X_t^n\right)^2 \, dt = \int_0^T \mathbb{E}\left(W_t - X_t^n\right)^2 \, dt$$
$$\leq \int_0^T T2^{-n} \, dt = T^2 2^{-n} \rightarrow 0$$

ולכן $\int_0^T W_t \, dW_t$ הוא הגבול (בממוצע ריבועי) של $I_n = \sum\limits_{i \leq 2^n} W_{t_i^{(n)}} \left(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}}\right)$

נחשב

$$\begin{aligned} {W_T}^2 &= \sum_i \left({W_{t_{i+1}(n)}}^2 - {W_{t_i^{(n)}}}^2\right) \\ &= \sum_i \left(W_{t_{i+1}^{(n)}} + W_{t_i^{(n)}}\right) \cdot \left(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}}\right) \\ &= \sum_i 2W_{t_i^{(n)}} \left(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}}\right) + \sum_i \left(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}}\right)^2 \\ &= 2I_n + S_n \rightarrow 2\int_0^T W_s \, dW_s + T \end{aligned}$$

$$\int_0^T W_s dW_s = \frac{1}{2}\left({W_T}^2 - T\right) \tag{4.2}$$

מסקנת החישוב, בהשוואת (4.1) ל-(4.2): לאינטגרל הסטוכסטי חוקים שונים.

הערה 4.15 ניתן להרחיב הגדרת האינטגרל הסטוכסטי בכמה כוונים. ראשית נסמן ב- L^2_{loc} את התהליכים $\{X_t, \ 0 \leq t \leq T\}$ כך ש- X_t מתואם ואינו תלוי בהפרשים עתידיים של W_t, וכן $\int_0^T X_s^2 \, ds < \infty \ a.s.$. ניתן להרחיב את האינטגרל למשפחה זו של תהליכים. במקרה זה, האינטגרל מקיים את תכונות הלינאריות (2), רציפות (4) ואפיון התנודה הריבועית (5). כוון שלא הנחנו קיום תוחלת, אין משמעות לתכונות 6, 3, 1. הרחבה נוספת בה לא נעסוק, מאפשרת להחליף את W_s בתהליך כללי יותר M_t כאשר M_t הוא, למשל, סכום של מרטינגל ופונקציה עם תנודה חסומה. תורת האינטגרציה הסטוכסטית המודרנית עוסקת באינטגרלים כאלו.

תרגיל 4.16 רשום את W_t^k ($k > 2$ שלם) כסכום של אינטגרל איטו ואינטגרל רגיל (הרחבה של הדוגמה הקודמת).

תרגיל 4.17 יהי τ זמן עצירה ו-X תהליך ב-L^2. הראה כי בהסתברות 1,

$$\int_0^\infty X_s \mathbf{1}_{\{s\leq\tau\}}\, dW_s = \int_0^\tau X_s\, dW_s \tag{4.3}$$

רמז: בדוק תחילה עבור $X\in E$.

נגדיר זמן עצירה $\tau_\alpha \doteq \inf\{t : W_t = \alpha\}$. זהו זמן עצירה שכן זהו הזמן הראשון להגיע לקבוצה סגורה.

דוגמה 4.18 ניתן להראות כי $\tau_a < \infty$ בהסתברות 1. לעומת זאת, התוחלת היא אין סופית, שכן

$$W_\tau = \int_0^{\tau_a} dW_s = \int_0^\infty \mathbf{1}_{\{s\leq\tau_a\}}\, dW_s. \tag{4.4}$$

אם $\mathbb{E}\,\tau_a < \infty$ אזי מתכונות האנטגרל הסטוכסטי (על אינטרוול זמן אין-סופי!), $\mathbf{1}_{\{s\leq\tau_a\}} \in L^2[0,\infty)$ ולכן האנטגרל הוא מרטינגל. כיוון שהאנטגרל מתאפס ב- $t=0$ קיבלנו $\mathbb{E}\,W_{\tau_a} = 0$. אולם $W_{\tau_a} = a$ בהסתברות 1 וקיבלנו סתירה.

דוגמה 4.19 אנו יודעים כי לכל s התהליך $W_{t+s} - W_s$ הוא תנועת בראון. יהי τ זמן עצירה. האם $X_t \doteq W_{t+\tau} - W_\tau$ היא תנועת בראון ביחס ל-σ-שדה $G_t \doteq \mathcal{F}_{\tau+t}$? התשובה היא חיובית, ונשתמש במשפט לוי כדי להוכיח זאת: כלומר, די להראות כי X_t וכן $X_t^2 - t$ הם תהליכי מרטינגל ביחס ל- G_t. נחשב

$$\mathbb{E}\left[X_{t+s} \mid G_t\right] = \mathbb{E}\left[W_{t+s+\tau} - W_\tau \mid \mathcal{F}_{\tau+t}\right] \tag{4.5}$$

$$= \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{r\leq t+s+\tau\}}\, dW_r - \int_0^\infty \mathbf{1}_{\{r\leq\tau\}}\, dW_r \;\middle|\; \mathcal{F}_{\tau+t}\right] \tag{4.6}$$

$$= \mathbb{E}\left[\int_0^{t+s+\tau} dW_r - \int_0^\tau dW_r \;\middle|\; \mathcal{F}_{\tau+t}\right] \tag{4.7}$$

$$= \mathbb{E}\left[\int_\tau^{t+\tau} dW_r + \int_{t+\tau}^{t+s+\tau} dW_r \;\middle|\; \mathcal{F}_{\tau+t}\right] \tag{4.8}$$

$$= X_t + 0 \tag{4.9}$$

בגלל מדידות וה- *Optional Sampling theorem* אשר נפעיל על זמני העצירה $t+s+\tau \geq t+\tau$. בצורה דומה על ידי שימוש בפירוק $W_t^2 = 2\int_0^t W_s\, dW_s - t$ נראה כי $X_t^2 - t$ הוא מרטינגל ביחס ל- G_t.

דוגמה 4.20 נחשב את הצפיפות של המ"א τ_α בעזרת תכסיס מיוחד, הנקרא עקרון השיקוף. ראשית נשים לב כי

$$\mathbb{P}(\tau_\alpha < T) = \mathbb{P}\left(\sup_{0\le s\le T} W_s > \alpha\right) \tag{4.10}$$

כאשר אין צורך לרשום $\geq$ כיוון שלתנועה הבראונית מסלולים רציפים. כעת, בגלל הסימטריה של התנועה הבראונית ואי התלות המותנית בעבר (לפני הפעם הראשונה שעברנו את α) אשר הראנו בדוגמה הקודמת,

$$\mathbb{P}\left(\sup_{0\le s\le T} W_s > \alpha, W_T \geq \alpha\right) = \mathbb{P}\left(\sup_{0\le s\le T} W_s > \alpha, W_T < \alpha\right) \tag{4.11}$$

שוויון זה נקרא "עקרון השיקוף". אולם בביטוי השמאלי, אם $W_T \geq \alpha$ אזי בפרט $\sup_{0\le s\le T} W_s \geq \alpha$ ולכן

$$\mathbb{P}\left(\sup_{0\le s\le T} W_s \geq \alpha\right) = \mathbb{P}\left(\sup_{0\le s\le T} W_s > \alpha, W_T > \alpha\right) + \mathbb{P}\left(\sup_{0\le s\le T} W_s > \alpha, W_T < \alpha\right) \tag{4.12}$$

$$= 2\,\mathbb{P}\left(\sup_{0\le s\le T} W_s > \alpha, W_T > \alpha\right) \tag{4.13}$$

$$= 2\,\mathbb{P}\left(W_T > \alpha\right)\ . \tag{4.14}$$

אולם W_T הוא מ"א גאוסי עם תוחלת אפס ווריאנס T! לכן ניתן לרשום בצורה מפורשת

$$\mathbb{P}\left(\sup_{0\le s\le T} W_s \geq \alpha\right) = 2\frac{1}{\sqrt{2\pi T}}\int_\alpha^\infty e^{-x^2/2T}\,dx = \sqrt{\frac{2}{\pi}}\int_{\alpha/\sqrt{T}}^\infty e^{-x^2/2}\,dx \tag{4.15}$$

הצפיפות של τ_a נתונה על ידי הנגזרת של $\mathbb{P}(\tau_\alpha \leq T)$ והיא

$$p_{\tau_\alpha}(x) = \sqrt{\frac{2}{\pi}}\,e^{-\alpha^2/2x}\cdot\frac{1}{2}x^{-3/2}\ . \tag{4.16}$$

לצפיפות זו אין מומנט ראשון, כפי שכבר ראינו.

4.2 גזירה ונוסחת איטו

כוון שאינטגרל סטוכסטי אינו מקיים נוסחת אינטגרציה בסיסית, אין לצפות שיקיים חוקים כמו גזירה של הרכבת פונקציות. במקרה דטרמיניסטי, אם ל- $f(t)$ נגזרת רציפה ול- $u(x)$ נגזרת רציפה, אזי

$$\frac{d}{dt}u(f(t)) = \frac{du(f(t))}{df}\cdot\frac{df(t)}{dt}$$

נרשום זאת כך:

$$du(f(t)) = u'_f(f(t))\,df(t)$$

כאשר "הכפלנו ב- dt": זהו רישום, שהמוטיבציה לו היא הנוסחה האינטגרלית השקולה:

$$u(f(T)) = u(f(0)) + \int_0^T u'_f(f(s))f'(s)\,ds$$

כדי להוכיח נוסחה זו במקרה הדטרמיניסטי, נשתמש בקירוב רימן לאינטגרל, ובקירוב טילור:

$$\begin{aligned} u(f(T)) &= u(f(0)) + \sum_i \{u\left(f(t_{i+1})\right) - u\left(f(t_i)\right)\} \\ &= u(f(0)) + \sum_i \{{u_f}'(f(t_i))\cdot(f(t_{i+1}) - f(t_i)) + error\} \end{aligned}$$

כאשר השגיאה חסומה ע"י

$$\max_{0\le s\le T} u''_f(f(s)) \max_i |f(t_{i+1}) - f(t_i)|\, V(f)$$

אבל מרציפות f וגזירות u, השגיאה שואפת לאפס אם $V(f) < \infty$.

כדי לקבל קירוב כזה עבור תנועה בראונית, יש לכלול אברים מסדר שני: אם ל- u שתי נגזרות רציפות,

$$u(W(t)) \approx u(W(0)) + \sum_i u'(W_{t_i})(W_{t_{i+1}} - W_{t_i}) + \frac{1}{2}\sum_i u''(W_{t_i})(W_{t_{i+1}} - W_{t_i})^2 + \dots$$

כאשר (ללא הוכחה) שאר האברים זניחים. האיבר השני, בהגדרתו, מתכנס ל- $\int_0^t u'(W_s)\,dW_s$. אם נזכר בתנודה הריבועית, הרי יש לצפות ש- $t_{i+1} - t_i \approx (W_{t_{i+1}} - W_{t_i})^2$, ולכן האיבר השלישי מתכנס ל- $\frac{1}{2}\int_0^t u''(W_s)\,ds$. זהו "גורם תיקון" שלא הופיע במקרה הדטרמניסטי, והוא נובע מכך שהתנודה של W אינה חסומה.

הערה 4.21 ת"א מהצורה $X_t = X_0 + \int_0^t A_s\,ds + \int_0^t B_s\,dW_s$ נקרא *Semimartingale*. רישום מקוצר

$$dX_t = A_t\,dt + B_t\,dW_t$$

<u>נוסחת איטו</u> (Itô):

יהי X_t תהליך המקיים $dX_t = A_t\,dt + B_t\,dW_t$ כאשר X_0 מדיד $\mathcal{F}_0$ ו- A ו- B שייכים ל- L^2_{loc}. תהי $u(t,x)$ פונקציה ממשית, מוגדרת על $0 \le t < \infty,\ -\infty < x < \infty$, וכך ש- u ונגזרותיה $u'_t,\ u'_x,\ u''_{xx}$ רציפות.

<u>נוסחת איטו</u> (ללא הוכחה). התהליך $Y_t = u(t, X_t)$ מקיים

$$Y_t = Y_0 + \int_0^t u'_s(s,X_s)\,ds + \int_0^t u'_x(s,X_s)\,dX_s + \frac{1}{2}\int_0^t u''_{xx}(s,X_s)B_s^2\,ds$$

כאשר האיבר האחרון הוא איבר "התיקון", והאיבר שלפניו הוא, לפי ההגדרה

$$\int_0^t u'_x(s,X_s)A_s\,ds + \int_0^t u'_x(s,X_s)B_s\,dWs$$

בכתיב מקוצר:

$$dY_t = u'_t(t,X_t)\,dt + u'_x(t,X_t)\,dX_t + \frac{1}{2}u''_{xx}(t,X_t)B_t^2\,dt$$

דוגמה 4.22

א. נבחר $X_t = W_t$ ו- $u(t,x) = x^n$ עבור $n \ge 2$. אזי

$$u'_x(t,x) = nx^{n-1},\ u''_{xx}(t,x) = n(n-1)x^{n-2},\ u'_t \equiv 0,\ B_t \equiv 1,\ A_t \equiv 0.$$

לכן, בכתיב מקוצר, $Y_t = W_t^n$ מקיים

$$dY_t = nW_t^{n-1}\,dW_t + \frac{1}{2}n(n-1)W_t^{n-2}\,dt$$

או, ברישום אינטגרלי (שקול)

$$W_t^n = Y_t = Y_0 + \int_0^t nW_s^{n-1}\,dW_s + \frac{1}{2}n(n-1)\int_0^t W_s^{n-2}\,ds$$

עבור $n=2$ מתקיים $W_s^{n-2} \equiv 1$, ומקבלים

$$W_t^2 = 2\int_0^t W_s\,dW_s + \frac{1}{2}\cdot 2t$$

כפי שקיבלנו קודם בחישוב ישיר.

ב. תהי $u(x)$ פונקציה קונווקסית בעלת שתי נגזרות רציפות. אזי $u''_{xx}(x) \geq 0$. יהי B_t תהליך ב- L^2, אזי $X_t = \int_0^t B_s\, dW_s$ הוא מרטינגל ו- $u(X_t)$ הוא *Submartingale*. לפי נוסחת איטו, אם $Y_t = u(X_t)$

$$dY_t = u'(X_t)\, dX_t + \frac{1}{2} u''(X_t) {B_t}^2\, dt = u'(X_t) B_t\, dW_t + \frac{1}{2} u''(X_t) {B_t}^2\, dt$$

נניח ש- $u'(X_t)B_t$ הוא תהליך ב- L^2. אזי (מתכונות האינטגרל הסטוכסטי) $\int_0^t u'(X_s) B_s\, dW_s$ הוא מרטינגל. בנוסף $u''(x) \geq 0$, ולכן $\int_0^t u''(X_s) {B_s}^2\, ds$ פונקציה (תהליך) <u>עולה</u>, המתחיל ב-0. קיבלנו <u>פירוק</u> של הסב-מרטינגל $u(X_t)$ למרטינגל + תהליך עולה.
בפרט, הפונקציה $u(x) = x^2$ קונווקסית, ולכן

$$X_t^2 = 2 \int_0^t X_s B_s\, dW_s + \int_0^t {B_s}^2\, ds$$

ברישום דיפרנציאלי,

$$d(X_t^2) = 2 X_t B_t\, dW_t + B_t^2\, dt$$

ג. נחפש פתרון למשוואה הסטוכסטית $dY_t = Y_t B_t\, dW_t$ (המקביל הדטרמיניסטי הוא $df/dt = f(t)B(t)$ וקיים פתרון סגור: $f(t) = C_1 e^{\int_0^t B_s\, ds}$). דיון כללי בנושא קיום ויחידות פתרונות כאלה נקיים בפרק *5*.
נניח $B_t \in L^2_{loc}$.

טענה *4.23* הפתרון הוא (עד כדי כפל בקבוע):

$$Y_t = \exp\left\{ \int_0^t B_s\, dW_s - \frac{1}{2} \int_0^t B_s^2\, ds \right\} \tag{4.17}$$

הוכחה: ע"י הצבה. כמובן שביטוי זה נותן תנאי התחלה $Y_0 = 1$. נגדיר תהליך X_t ע"י $X_0 = 0$ ו-

$$dX_t = B_t\, dW_t - \frac{1}{2} B_t^2\, dt \tag{4.18}$$

אזי הניחוש שלנו הוא כי $Y_t = e^{X_t}$. נפעיל נוסחת איטו על הפונקציה $u(x) = e^x$, ונקבל

$$dY_t = e^{X_t}\, dX_t + \frac{1}{2} e^{X_t} {B_t}^2\, dt =$$

$$= Y_t B_t\, dW_t - \frac{1}{2} Y_t {B_t}^2\, dt + \frac{1}{2} Y_t {B_t}^2\, dt = Y_t B_t\, dW_t$$

אם נציב במשוואה המקורית ל- Y_t, נקבל

$$Y_t = e^{x_t} = Y_0 + \int_0^t Y_s B_s \, dW_s = 1 + \int_0^t e^{X_s} B_s \, dW_s$$

בנוסף לפתרון המשוואה הסטוכסטית, קיבלנו את המסקנה החשובה הבאה. אם X_t מקיים את המשוואה (4.18) ו-$e^{X_t} B_t \in L^2$, אזי התהליך e^{X_t} הוא מרטינגל לא שלילי!

ד. יהי A_t תהליך ב- L_2 כך ש- $|A_t| \equiv 1$.

טענה *4.24* $V_t = \int_0^t A_s \, dWs$ הוא תנועה בראונית.

הוכחה: נביא שלוש הוכחות, המדגימות שימוש בכלים שונים שלמדנו. מתכונות האינטגרל הסטוכסטי, ל-V_t דגמים רציפים. לצורך שתי ההוכחות הראשונות נראה כי V_t תהליך גאוסי עם הפרשים בת"ס---וכדי לסכם את ההוכחה נשתמש בעובדה (אותה קל לראות) כי $\mathbb{E} \, V_t V_s = \min(t,s)$.

נראה זאת ע"י חישוב פונקציה אופיינית של ווקטורים גאוסיים. תהי $f(s)$ פונקציה דטרמניסטית וחסומה.

תרגיל *4.25* V_t גאוסי עם הפרשים בת"ס אם לכל f פשוט ודטרמניסטי,

$$\mathbb{E} \exp \left\{ i \int_0^T f(s) A_s \, dV_s \right\} = \exp \left\{ -\frac{1}{2} \int_0^T f^2(s) \, ds \right\} \tag{4.19}$$

אם בנוסף $W_0 = 0$ אזי W_t תנועת בראון סטנדרטית.

רמז: בדיקת גאוסיות של תהליך נעשית ע"י בדיקת פילוגים סופיים. פילוגים אלו ניתן לבדוק על ידי חישוב הפונקציה האפיינית של הווקטור המתקבל. בנוסף, אי תלות סטטיסטית ניתן לבדוק על ידי בדיקה שהפונקציה האפיינית מתפרקת למכפלה.

הוכחה ראשונה: נגדיר תהליך X_t על ידי

$$dX_t = i f(t) A_t \, dW_t - \frac{1}{2}(i f(t) A_t)^2 \, dt \ .$$

זוהי בדיוק משוואה מהטיפוס של (4.18), ולכן מהדוגמה הקודמת נסיק כי e_t^X הוא מרטינגל. בפרט, כיוון ש- $A_s^2 = 1$,

$$\mathbb{E} \, e^{X_t} = \mathbb{E} \exp \left(i \int_0^t f(s) A_s \, dW_s + \frac{1}{2} \int_0^t f^2(s) \, ds \right) = \mathbb{E} \, e^{X_0} = 1$$

ולכן (4.19) מתקיים.

הוכחה שנייה: כעת נגדיר $Y_t = \int_0^t f(s)A_s\,dW_s$ ו- $u(x) = e^{ix}$.
ממשפט איטו, התהליך $C_t = u(Y_t),\ 0 \le t \le T$ מקיים

$$C_t = 1 + \int_0^t iC_s f(s)A_s\,dW_s + \frac{1}{2}(i)^2\int_0^t C_s f_s^2 A_s^2\,ds$$

כיוון ש C_s חסום, ו- $f(s)$ חסום, התהליך $C_s f(s)A_s \in L^2$, והאינטגרל הראשון הוא מרטינגל, ותוחלתו אפס. כוון ש- ${A_s}^2 = 1$,

$$\mathbb{E}\,C_t = 1 - \frac{1}{2}\int_0^t \mathbb{E}\,C_s f^2(s)\,ds$$

זוהי משוואה דיפרנציאלית דטרמיניסטית עבור $\mathbb{E}\,C_t$:

$$\frac{d\,\mathbb{E}\,C_t}{dt} = -\frac{1}{2}f^2(t)\,\mathbb{E}\,C_t, \quad \mathbb{E}\,C_0 = 1$$

ופתרונה

$$\mathbb{E}\,C_t = \exp\left\{-\frac{1}{2}\int_0^t f^2(s)\,ds\right\}$$

הוכחה שלישית: לא ישירה.
ל- V_t ורסיה עם דגמים רציפים והיא מרטינגל. לפי משפט $P.\ Levy$, V_t תנועה בראונית אם ${V_t}^2 - t$ הוא מרטינגל. מנוסחת איטו,

$$V_t^2 = V_0^2 + 2\int_0^t V_s A_s\,dW_s + 2\cdot\frac{1}{2}\int_0^t A_s^2\,ds = 2\int_0^t V_s A_s\,dW_s + t$$

ומתכונות האינטגרל הסטוכסטי, ${V_t}^2 - t$ מרטינגל.

ה. יהי $X_t = \int_0^t h_s\,ds + W_t$ כאשר h_t ב- L^2 (ובפרט h_t מדיד F_t לכל t). נסמן $D_t = \sigma\{X_\theta,\ \theta \le t\}$ ו- $\hat{h}_t = \mathbb{E}\left[h_t \mid D_t\right]$. נפרש את X_t כמדידה רועשת של h_t אחרי "מסנן" אינטגרלי ו- $\hat{h}_t$ הוא השערוך של h. נניח לצורך כך קיום "ורסיה טובה" של $\hat{h}$ (כפונקציה של t). אזי

$$X_t = \int_0^t \hat{h}_s\,ds + \int_0^t (h_s - \hat{h}_s)\,ds + W_t$$

הוכחנו בדוגמה 3.58 כי אם W_t מרטינגל, אזי $N_t = \int_0^t (h_s - \hat{h}_s)\,ds + W_t$ גם הוא מרטינגל.

טענה 4.26 אם W_t תנועה בראונית אזי גם N_t תנועה בראונית.

הוכחה: נסמן

$$Y_t = \int_0^t if(s)(h_s - \hat{h}_s)\,ds + \int_0^t if(s)\,dW_s = \int_0^t if(s)\,dN_s$$

כאשר $f(t)$ היא פונקציה דטרמניסטית חסומה שרירותית (מספיק כמובן שתהיה פשוטה). מהדוגמה הקודמת (דוגמה ד' והתרגיל), מספיק להוכיח

$$\mathbb{E}\exp\left\{i\int_0^T f(s)\,dN_s\right\} = \exp\left\{-\frac{1}{2}\int_0^T f^2(s)\,ds\right\}$$

נפעיל נוסחת איטו עבור $\alpha_t = e^{Y_t}$. אזי

$$\begin{aligned}
d\alpha_t &= \alpha_t\,dY_t + \frac{1}{2}\alpha_t(if(t))^2\,dt \\
&= i\alpha_t f(t)(h_t - \hat{h}_t)\,dt + \alpha_t if(t)\,dW_t - \frac{1}{2}\alpha_t f^2(t)\,dt
\end{aligned}$$

ובצורה אינטגרלית, כיוון ש- $Y_0 = 0$, ומתכונות האינטגרל הסטוכסטי,

$$\mathbb{E}\,\alpha_t = 1 + i\,\mathbb{E}\int_0^t \alpha_s f(s)(h_s - \hat{h}_s)\,ds - \frac{1}{2}\,\mathbb{E}\int_0^t \alpha_s f^2(s)\,ds$$

שים לב כי Y_t ולכן גם α_t מדידה על D_t (מדוע?) לכן:

$$\begin{aligned}
\mathbb{E}\int_0^t \alpha_s f(s)(h_s - \hat{h}_s)\,ds &= \int_0^t \mathbb{E}\,\alpha_s f(s)(h_s - \hat{h}_s)\,ds \\
&= \int_0^t \mathbb{E}\left\{\mathbb{E}\left[\alpha_s f(s)(h_s - \hat{h}_s)\;\Big|\;D_s\right]\right\}\,ds \\
&= \int_0^t \mathbb{E}\,\alpha_s f(s)\,\mathbb{E}\left[h_s - \hat{h}_s\;\Big|\;D_s\right]\,ds = 0
\end{aligned}$$

כיוון שמהגדרת $\hat{h}_t$, $\mathbb{E}\left[\hat{h}_t - h_t \;\Big|\; D_t\right] = 0$ קבלנו כי $\mathbb{E}\,\alpha_t$ פותר את המשוואה הדיפרנציאלית של דוגמה ד' ולכן

$$\mathbb{E}\,\alpha_t = 1 - \frac{1}{2}\int_0^t \mathbb{E}\,\alpha_s f^2(s)\,ds = \exp\left\{-\frac{1}{2}\int_0^t f^2(s)\,ds\right\}$$

תרגיל 4.27 נגדיר $X_t = W_t\,,\ Y_t = e^{W_t^2}$ כאשר W_t תנועת בראון.
הראה כי הווקטור $(X_t Y_t)$ פותר את המשוואה הסטוכסטית

$$\begin{cases} dX_t = dW_t\,, & X_0 = 0 \\ \\ dY_t = 2X_tY_t\,dW_t + (Y_t + 2X_t^2Y_t)\,dt\,, & Y_0 = 1 \end{cases}$$

תרגיל 4.28 תהיינה W_t ו- $\tilde{W}_t$ תנועות בראוניות בלתי תלויות סטטיסטית.

א. הסבר והוכח את הטבלה

	dW_t	$d\tilde{W}_t$	dt
dWt	dt	0	0
$d\tilde{W}_t$	0	dt	0
dt	0	0	0

תהי u פונקציה "חלקה" $u = u(t,x,y)$.
נגדיר

$$X_t = X_0 + \int_0^t A_1(s)\,ds + \int_0^t B_{11}(s)\,dWs + \int_0^t B_{12}(s)\,d\tilde{W}_s$$

$$Y_t = Y_0 + \int_0^t A_2(s)\,ds + \int_0^t B_{21}(s)\,dWs + \int_0^t B_{22}(s)\,d\tilde{W}_s$$

ב. נחש את נוסחת איטו (עבור $u(t,X_t,Y_t)$), מתוך הטבלה ב-א'.

ג. הראה על סמך ב' את נוסחת האינטגרציה בחלקים:

$$X_tY_t = X_0Y_0 + \int_0^t X_s\,dY_s + \int_0^t Y_s\,dX_s + \int_0^t B_{11}B_{21}\,ds + \int_0^t B_{22}B_{21}\,ds$$

<u>נוסחת איטו רב-מימדית</u>

יהי $\underline{W}_t = \left\{W_t^1, W_t^2, \cdots, W_t^K\right\}$**, כאשר** W_t^i **תנועת בראון סטנדרטית (כלומר** $t = \mathbb{E}(W_t^i)^2$, $\mathbb{E}\,W_t^i = 0$**) ו-** ${W_t}^i$ **לא תלויה סטטיסטית ב-** $\{{W_t}^j,\ j \neq i\}$.

תהי $u(t,x)$ **פונקציה ממשית, כאשר** x **וקטור** N- **מימדי, וכך ש-** $u(t,x)$ $u'_t(t,x),\ u'_{x_i}(t,x)$ **ו-**

$u''_{x_ix_j}(t,x)$ **רציפים ב-** (t,x) **לכל** i,j.

יהי X^i_t **פתרון של**

$$dX^i_t = A^i_t\,dt + \sum_{j=1}^{K} b^{ij}_t\,dW^j_t$$

כאשר A^i_t , b^{ij}_t **ב-** L^2_{loc}. **אזי** $Y_t = u(t,\underline{X}_t)$ **מקיים את המ.ד.ס.**

$$dY_t = u'_t(t,\underline{X}_t)\,dt + \sum_{i=1}^{N} u'_{x_i}(t,\underline{X}_t)\,dX^i_t + \frac{1}{2}\sum_{i,j=1}^{N}\sum_{l=1}^{K} u''_{x_ix_j}(t,\underline{X}_t) b^{il}_t b^{jl}_t\,dt$$

הערה 4.29 הרחבה כאשר $u(t,X)$ הוא וקטורי, כלומר

$$\begin{pmatrix} u^1(t,\underline{X}) \\ u^2(t,\underline{X}) \\ \vdots \\ u^L(t,\underline{X}) \end{pmatrix}$$

היא מיידית - יש לטפל בנפרד בכל קואורדינטה.

תרגיל 4.30 משוואות סטוכסטיות לינאריות. בבעיות להלן, העזר בנוסחת איטו.

א. הפתרון למשוואה הדפרנציאלית $\frac{dx(t)}{dt} = -\alpha x(t) + u(t)$ היא, כידוע,

$$x(t) = e^{-\alpha t}X_0 + \int_0^t e^{\alpha(t-\tau)}u(\tau)\,d\tau$$

הראה כי

$$x(t) = e^{-\alpha t}X_0 + \int_0^t e^{\alpha(t-\tau)}\,dW(\tau)$$

הוא פתרון למ.ד.ס. $dX_t = -\alpha X_t dt + dW_t$.

רמז: הגדר $dY_t = e^{\alpha t}\,dW_t$, ו- $u(t,y) = e^{-\alpha t}X_0 + e^{-\alpha t}y$.

ב. פתור את $dZ_t = -\alpha Z_t\,dt + b\,dW_t$.

ג. פתור את המ.ד.ס. הווקטורית $d\underline{X}_t = -\underline{\underline{A}}\ \underline{X}_t\,dt + \underline{\underline{B}}\ d\underline{W}_t$ כאשר $\underline{X}$, $\underline{W}$, וקטורי תנועות בראוניות בת"ס, $\underline{\underline{A}}$ ו- $\underline{\underline{B}}$ מטריצות בממדים מתאימים.

ד. הפתרון של המשוואה הסקלרית $\frac{dx(t)}{dt} = -\alpha(t)x(t) + u(t)$ הוא

$$x(t) = \frac{1}{\mu(t)} \int_0^t \mu(s)u(s)\,ds$$

כאשר

$$\mu(t) = \exp\left\{\int_0^t \alpha(\tau)\,d\tau\right\}$$

פתור את המד"ס $dX_t = -\alpha(t)X_t\,dt + b\,dW_t$ כאשר $\alpha(t)$ פונקציה דטרמיניסטית.

תרגיל 4.31 נתונה משוואה דטרמניסטית $\ddot{Y}_t + b(t, Y_t, \dot{Y}_t) = f(t)$. נניח כי $f(t)$ הוא "רעש לבן" גאוסי, $\mathbb{E}\,f(t) = 0$, כלומר $\mathbb{E}\left\{f(t_1)f(t_2)\right\} = \beta^2\delta(t_1 - t_2)$. רשום מד"ס למערכת.
רמז: $\int_0^t f(s)\,ds = \beta W_t$.

חסר שרטוט

תרגיל 4.32

א. רשום מ.ד.ס. ל- Y_t (מ.ד.ס. ווקטורי).

ב. רשום מ.ד.ס. עבור המטען בקבל.

ג. חזור על ב' כאשר הקבל אינו לינארי, וקיבולו $C(y)$.

ד. חזור על א' עבור המערכת.

ה. חזור על ד' כאשר R_2 לא לינארי: $i_2 = g(X_t - Y_t)$.

5 משוואות דיפרנציאליות סטוכסטיות

תהיינה $m(t,x)$ ו-$\sigma(t,x)$ פונקציות דטרמיניסטיות, מוגדרות עבור $t \geq 0, -\infty < x < \infty$. תהי W_t תנועה בראונית. בפרק זה $\mathcal{F}_t$ הינו $\mathcal{F}_t = \sigma(W_s, s \leq t)$.

הגדרה 5.1 תהליך X_t הוא פתרון של המשוואה הדיפרנציאלית הסטוכסטית (מ.ד.ס)

$$dX_t = m(t, X_t)\,dt + \sigma(t, X_t)\,dW_t \tag{5.1}$$

אם X_t, $m(t, X_t)$ ו-$\sigma(t, X_t)$ שייכים ל-L^2_{loc}, וכן עבור כל t מתקיים, בהסתברות 1,

$$X_t = X_0 + \int_0^t m(s, X_s)\,ds + \int_0^t \sigma(s, X_s)\,dW_s \tag{5.2}$$

הערה 5.2 ניתן להרחיב ללא קשיים טכניים למקרה בו X_0 הוא מ"א בת"ס ב-W_t.

תרגיל 5.3 רשום הגדרה מקבילה למשוואה וקטורית

$$X_t = \left[X_t^{(1)}, X_t^{(2)}, \cdots, X_t^{(N)}\right]^T$$

תרגיל 5.4 מצא מ.ד.ס וקטורית שפתרונה הוא

$$X_t = R_0 \cos W_t$$

$$Y_t = R_0 \sin W_t$$

<u>קיום פתרונות למ.ד.ס</u> בעיית קיום הפתרונות קיימת גם במשוואות רגילות.

דוגמה 5.5 נחפש פתרון $X_t, 0 \leq t < \infty$ למשוואה

$$\frac{dx}{dt} = x^2, \quad x_0 = 1$$

אם $x_0 = 1$ הפתרון עבור $0 \le t < 1$ הוא $x_t = (1-t)^{-1}$. פתרון זה "מתפוצץ" כאשר $t \to 1$. שימו לב ש-x^2 היא פונקציה חלקה של x! הבעיה היא שקצב הגידול שלה מהיר מדי.

הגדרה 5.6 המשוואה הכללית למ.ד.ס וקטורי

$$d\underline{X}_t = m(t, \underline{X}_t)\, dt + \sigma(t, \underline{X}_t)\, d\underline{W}_t \tag{5.3}$$

כאשר:

$\underline{X}_t$ וקטור N ממדי, ולכל t, x

$m(t,x)$ וקטור N ממדי

$\sigma(t,x)$ מטריצה $N \times K$

$\underline{W}_t$ וקטור K ממדי:

$$\underline{W}_t = [W_t^1, W_t^2, \cdots, W_t^K]^T$$

ו-$\{W_t^i\}$ תנועות בראוניות סטנדרטיות בת"ס. בסימון סקלרי:

$$d\underline{X}_t^i = m_i(t, \underline{X}_t)\, dt + \sum_{j=1}^{K} \sigma_{ij}(t, \underline{X}_t)\, dW_t^j \tag{5.4}$$

הערה 5.7 על ידי הוספת משוואה סקלרית נוספת: $dX_t^{N+1} = dt$ שפתרונה $X_t^{N+1} = t$, וע"י עדכון מתאים של m ו-σ ניתן לעבור למשוואה שאינה תלויה מפורשות ב-t . לכן, נדון במשוואה

$$d\underline{X}_t = m(\underline{X}_t)\, dt + \sigma(x_t)\, d\underline{W}_t \tag{5.5}$$

משפט 5.8 אם $\{W_t\}$ $\mathbb{E}\, X_0^2 < \infty, X_0 \perp$ ואם $m(x)$ ו-$\sigma(x)$ רציפות ליפשיץ, כלומר

$$\|m(x) - m(y)\| \le c\|x - y\| \tag{5.6}$$

$$\|\sigma(x) - \sigma(y)| \le c\|x - y\| \tag{5.7}$$

כאשר c קבוע ו-$|\cdot|$ מסמן נורמה אוקלידית כלשהיא, אזי יש פתרון יחיד ב- $\boldsymbol{L}^2$ למ.ד.ס (5.1).

ההכללה למקרה הוקטורית פשוטה.

הערה 5.9 כלומר, יש תהליך יחיד שהוא מדיד, מתואם ל-$\mathcal{F}_t$ ומקיים

$$\mathbb{E}\int_0^T X_s^2\,ds < \infty$$

וכן ש-X_t פותר את (5.1).

הוכחה: (בהמשך ניתן הוכחה חלופית, מתוחכמת יותר). נגדיר $X_t^0 = X_0$ ו-

$$X_t^{N+1} = X_0 + \int_0^t m(X_t^N)\,dt + \int_0^t \sigma(X_t^N)\,dW_t \qquad N = 0, 1, ... \tag{5.8}$$

$$X_t^{N+1} - X_t^N = \int_0^t (m(X_t^N) - m(X_t^{N-1}))\,dt + \int_0^t (\sigma(X_t^N) - \sigma(X_t^{N-1}))\,dW_t \tag{5.9}$$

בפיתוח שנבצע עכשיו, נשתמש בעובדות הבאות, לפי הסדר:

א. $\|\sum_1^N a_i\|^2 \le N\sum(a_i^2)$ בעובדה זו, הנובעת מיידית מאי שוויון ינסן, נשתמש רבות גם בהוכחות הבאות.

ב. אפשר להתייחס ל-$\frac{1}{t}\int_0^t f(s)\,ds$ כאל תוחלת של $f(s)$ על צפיפות אחידה $U[0,t]$, ולכן אפשר להתשתמש באי שיוויון ינסן.

ג. הליפשציות של הפונקציות $m(x), \sigma(x)$

ובכן מעובדות אלו ותכונות האנטגרל הסטוכסטי,

$$\begin{aligned}
&\mathbb{E}\|X_t^{N+1} - X_t^N\|^2 \\
&\quad\le 2\,\mathbb{E}\,t^2\left(\frac{1}{t}\int_0^t (m(X_s^N) - m(X_s^{N-1}))\,ds\right)^2 + 2\,\mathbb{E}\left(\int_0^t (\sigma(X_s^N) - \sigma(X_s^{N-1}))\,dW_s\right)^2 \\
&\quad\le 2t^2\,\mathbb{E}\,\frac{1}{t}\int_0^t (m(X_s^N) - m(X_s^{N-1}))^2\,ds + 2\,\mathbb{E}\int_0^t (\sigma(X_s^N) - \sigma(X_s^{N-1}))^2\,ds \\
&\quad\le 2(t+1)c^2\int_0^t \mathbb{E}\,\|X_s^N - X_s^{N-1}\|^2\,ds
\end{aligned}$$

כאשר c הוא קבוע הליפשיץ של הפונקציות m, σ. יהא כעת T קבוע (נבחר אותו בהמשך), אזי

$$\mathbb{E}\int_0^T \|X_t^{N+1} - X_t^N\|^2\, dt \leq 2T(T+1)c^2\, \mathbb{E}\int_0^T \|X_t^N - X_t^{N-1}\|^2\, dt$$

נסמן $a_{N+1} = \mathbb{E}\int_0^T \|X_t^{N+1} - X_t^N\|^2\, dt$, קיבלנו מקודם כי $a_{N+1} \leq 2T(T+1)c^2 a_N$. נבחר T_1 כך ש- $2T_1(T_1+1)c^2 = \alpha < 1$. אזי $a_{N+1} \leq \alpha a_N$ ולכן $a_N \leq \alpha^N a_0$. מכאן נובע כי a_N היא סדרת קושי המתכנסת ל-0: מהפעלה חוזרת של אי שויון המשולש,

$$\|a_{N+M} - a_N\| \leq \alpha^N a_0(\alpha^M - 1) \overset{N\to\infty}{\to} 0$$

<u>מסקנה</u>: באינטרוול $[0, T_1]$, יש התכנסות של האיטרציות: $X_t^N \overset{L^2}{\to} X_t$ ונותר לבדוק כי X_t מקיים את המשוואה. אבל

$$X_t^{N+1} = X_0 + \int_0^t m(X_s^N) + \int_0^t \sigma(X_s^N)\, dW_s$$

ולכן, כמו בחישוב קודם,

$$\begin{aligned}
&\mathbb{E}\left(X_t - \left[X_0 + \int_0^t m(X_s)\, ds + \int_0^t \sigma(X_s)\, dW_s\right]\right)^2 \\
&\leq 3\,\mathbb{E}\left(X_t - X_t^{N+1}\right)^2 \\
&\qquad + 3\,\mathbb{E}\left(\int_0^t \left(m(X_s) - m(X_s^N)\right)\, ds\right)^2 + 3\,\mathbb{E}\left(\int_0^t \left(\sigma(X_s) - \sigma(X_s^N)\right)\, dW_s\right)^2 \\
&\leq 3\,\mathbb{E}(X_t - X_t^{N+1})^2 + 3(t+1)c^2 \int_0^t \mathbb{E}\left(X_s - X_s^N\right)^2\, ds \overset{N\to\infty}{\to} 0
\end{aligned}$$

הוכחנו לכן כי X מקיים את המשוואה, דהיינו הוכחנו קיום הפתרון למשוואה באינטרוול $[0, T_1]$. שימו לב כי אורך האינטרוול בו הוכחנו קיום אינו תלוי בתנאי ההתחלה. נמשיך כעת את הפתרון לאינטרוול $[T_1, 2T_1]$ עם תנאי התחלה X_{T_1} באותה הצורה, ובדרך איטרטיבית נסיק כי קיים פתרון לכל אינטרוול זמן סופי. נשאר להוכיח יחידות, ולשם כך נעזר במשפט נקודת השבת (Fixed Point) 8.4.

<u>הוכחת קיום (חלופית) ויחידות הפתרון</u>

תהי L^2 משפחת התהליכים המדידים, מתואמים ל-$\mathcal{F}_t$, $0 \leq t \leq \tau$. נשתמש במטריקה (המבוססת

על הנורמה) הרגילה עבור משפחה זו

$$d^2(X,Y) = \mathbb{E}\int_0^\tau (Y(t,\omega) - X(t,\omega))^2\, dt = \|Y - X\|^2 \tag{5.10}$$

אזי, כפי שכבר נטען בעבר, זהו מרחב מטרי שלם (למעשה---מרחב לינארי עם נורמה, ושלם - כלומר מרחב בנך). נגדיר אופרטור $T : L^2 \to L^2$

$$(TX)(t,\omega) = X_0 + \int_0^t m(X_s)\, ds + \int_0^t \sigma(X_s)\, dW_s \tag{5.11}$$

אופרטור זה מעביר אלמנטים ב-L^2 לאלמנטים ב-L^2 בגלל תכונת רציפות ליפשיץ של המקדמים, ונקודת השבת שלו (fixed point-ראה משוואה (8.2)) היא פתרון למ.ד.ס. כעת, לכל $X, Y \in L^2$

$$\mathbb{E}\left((TX)(t,\omega) - (TY)(t,\omega)\right)^2 = \mathbb{E}\left(\int_0^t \left(m(X_s) - m(Y_s)\right)\, ds + \int_0^t \left(\sigma(X_s) - \sigma(Y_s)\right)\, dW_s\right)^2$$

נשתמש בעובדות הבאות:

$$\mathbb{E}\left(\int_0^t V_s\, dW_s\right)^2 = \mathbb{E}\int_0^t V_s^2\, ds\ , \qquad (a+b)^2 \le 2(a^2 + b^2)$$

לכל $V_s \in L^2$. לכן

$$\mathbb{E}\left((TX)(t,\omega) - (TY)(t,\omega)\right)^2$$

$$\le 2\,\mathbb{E}\left(\int_0^t \left(m(X_s) - m(Y_s)\right)\, ds\right)^2 + 2\,\mathbb{E}\int_0^t \left(\sigma(X_s) - \sigma(Y_s)\right)^2\, ds$$

נשתמש באי שיוויון ינסן בצורה הבאה:

$$\left(\frac{1}{t}\int_0^t f(s)\, ds\right)^2 \le \frac{1}{t}\int_0^t f^2(s)\, ds$$

ובעובדה שלפונקציות m, σ יש מקדם ליפשיץ c ונקבל

$$\le 2c^2 t\,\mathbb{E}\int_0^t \left(X_s - Y_s\right)^2\, ds + 2c^2\,\mathbb{E}\int_0^t \left(X_s - Y_s\right)^2\, ds = 2c^2(t+1)\,\mathbb{E}\int_0^t \left(X_s - Y_s\right)^2\, ds$$

לכן, אם נסמן $k = 2c^2(\tau+1)$,

$$\begin{aligned}
\|T^nX - T^nY\|^2 &= \mathbb{E}\int_0^\tau dt\,[(T^nX)(t,\omega) - (T^nY)(t,\omega)]^2 \\
&\le k\,\mathbb{E}\int_0^\tau dt\int_0^t ds\,\left[(T^{n-1}X)(t,\omega) - (T^{n-1}Y)(t,\omega)\right]^2 \\
&\le k^2\,\mathbb{E}\int_0^\tau dt_1\int_0^{t_1} dt_2\int_0^{t_2}\left[(T^{n-2}X)(t,\omega) - (T^{n-2}Y)(t,\omega)\right]^2 \\
&\le k^n\int_0^\tau dt_1\int_0^{t_1} dt_2\cdots\int_0^{t_n} ds\,\mathbb{E}\,(X_s - Y_s)^2 \\
&\le k^n\|X-Y\|^2\frac{\tau^n}{n!}
\end{aligned}$$

כאשר המעבר האחרון נובע מזהות אינטגרלית ידועה.

מצד אחד, אם נבחר $n=1$ ו-$Y = K = constant$ (לכן גם $Y \in L^2$), אז ברור כי $TY \in L^2$, (מדוע?), ומכאן

$$\begin{aligned}
\|TX\| &= \|TX - TK + TK\| \le \|TX - TK\| + \|TK\| \le k\tau\|X-K\| + \|TK\| \\
&\le k\tau\|X\| + k\tau\|K\| + \|TK\| < \infty
\end{aligned}$$

ולכן $TX \in L^2, \forall X \in L^2$.

ומצד שני, נבחר n_0 כך ש- $\frac{(k\tau)^n}{n!} = q^2 < 1$ ואז

$$\|T^{n_0}X - T^{n_0}Y\| \le q\|X-Y\|$$

ולכן T^{n_0} העתקה מכווצת.

מכאן, ממשפט נקודת השבת 8.4 והתרגיל 8.5 שאחריו, ל-T יש נקודת שבת יחידה ב-L^2, כלומר, קיים X יחיד כך ש-

$$X_t = (TX)(t,\omega) = X_0 + \int_0^t m(X_s)\,ds + \int_0^t \sigma(X_s)\,dW_s$$

וזהו כמובן פתרון גלובלי, לכל τ!

תרגיל 5.10 נניח כי עבור כל $-\infty < \beta < \infty$, מתקיים

$$F(t,\beta) = F_0(t,\beta) + \int_0^t G(s,\beta)\,ds + \int_0^t H(s,\beta)\,dW_s^{(1)}$$

כאשר לכל β , F_0, G, H שייכים ל-L^2_{loc}. יהא X_t ת"א המקיים

$$dX_t = A_t\, dt + B_t\, dW_t^{(2)}$$

כאשר A_t ו-B_t תהליכים ב-L^2_{loc}. מצא משיקולים אינטואיטיביים ביטוי עבור $F(t, X_t)$ כאשר

א. $W_t^{(1)} = W_t^{(2)}$ תנועה בראונית.

ב. $W_t^{(i)}$ תנועות בראוניות בת"ס.

הנח קיום נגזרות של H, G במידה הנדרשת. התרגיל לקוח מ-

Vestnik,Moscow Univ. #1, 1973, pp. 26-32

רמז: פתח לטור טיילר סביב t_i, X_{t_i}, כאשר האינטגרלים רשומים כסכומים.

דוגמה 5.11 מודל שוק ע"י משוואה סטוכסטית.

ברגע t בידנו הון X_t. כיצד לחלקו בין השקעות בטוחות וספקולציות? המודל להון הוא

$$dX_t = [p_t\, (\beta X_t\, dt + \sigma X_t\, dW_t) + (1 - p_t)\alpha X_t\, dt]$$

כאשר האיבר השמאלי מייצג מניות (תנודתיות) והשני אגרות חוב (יציבות), ו- $0 \le p_t \le 1$, מדיד על $\sigma(X_s, 0 \le s \le t)$: נתון לשליטתנו ומתאר את חלוקת ההשקעה. α, β הם מקדמי ה"סחיפה" (*drift*) של ההשקעה. נחפש p_t כך שנשיג $\sup_{0 \le p_\cdot \le 1} \mathbb{E}\, X_t$, כאשר ה-sup הוא על כל ערכי p_t שבשליטתנו, דהיינו p_t מדיד על $\sigma(X_s, s \le t)$. לשם כך, נגדיר משתנה $\log X_t$. מנוסחת איטו,

$$\begin{aligned} d \log X_t &= \frac{1}{X_t}\, dX_t - \frac{1}{2X_t^2}\sigma^2 X_t^2 p_t^2\, dt \\ &= [\beta p_t + \alpha(1 - p_t)]\; dt + \sigma p_t\, dW_t - \frac{1}{2}\sigma^2 p_t^2\, dt \end{aligned}$$

קיבלנו משוואה מפורשת, כלומר ניתן לרשום ביטוי מפורש (אינטגרל) עבור $\log X$ (שכן המשתנה X_t אינו מופיע בצד ימין!). בצורה אינטגרלית,

$$\begin{aligned} \log X_t &= \log X_0 + \int_0^t \sigma p_s\, dW_s - \frac{1}{2}\int_0^t \sigma^2 p_s^2\, ds + \int_0^t [\beta p_s + (1 - p_s)\alpha]\; ds \\ X_t &= X_0 \exp\left(\int_0^t \sigma p_s\, dW_s - \frac{1}{2}\int_0^t \sigma^2 p_s^2\, ds\right) \cdot \exp \int_0^t ((1 - p_s)\alpha + p_s\beta)\; ds \\ &\le X_0 \exp\left(\int_0^t \sigma p_s\, dW_s - \frac{1}{2}\int_0^t \sigma^2 p_s^2\, ds\right) \cdot \sup_{0 \le p_\cdot \le 1} \exp \int_0^t ((1 - p_s)\alpha + p_s\beta)\; ds \end{aligned}$$

המקסימום בביטוי האחרון <u>מושג</u> עבור $p_t \equiv 1$ אם $\alpha > \beta$ ובמקרה בו $\alpha \leq \beta$ המקסימום יתקבל עבור $p_t \equiv 0$, כלומר האקספוננט השני בביטוי הוא דטרמיניסטי. האקספוננט הראשון הוא, כמו שראינו כבר קודם, מרטינגל עם תוחלת 1, ללא תלות ב- p_t, ולכן (עבור תנאי התחלה דטרמיניסטי):

$$\mathbb{E}\, X_t \leq X_0 \cdot 1 \cdot e^{t \cdot \max\{\alpha,\beta\}}$$

מסקנה: מצאנו p_t עבורו אי השיוויון מושג בשיוויון, ולכן המקסימום <u>מושג</u> עבור p_t זה.
דוגמא זו לקוחה מ-*R.C. Metron Rev. Economics and Statistics, 51, 1969*.

6 תהליכים מרקוביים ותהליכי דיפוזיה

הגדרה 6.1 תהליך $\{x_t\}$ נקרא מרקובי אם לכל אוסף זמנים $t_1 < t_2 < \ldots$, לכל n ולכל קבוצת בורל B,

$$\mathbb{P}\left[x_{t_{n+1}} \in B \;\middle|\; x_{t_n}, \ldots, x_{t_1}\right] = \mathbb{P}\left[x_{t_{n+1}} \in B \;\middle|\; x_{t_n}\right] \ . \tag{6.1}$$

תחת תנאי ספרביליות ההגדרה שקולה לתנאי כי לכל t, T ו-$\varepsilon > 0$,

$$\sigma\{x_s,\ 0 \le s \le t\} \underset{\sigma\{x_t\}}{\perp} \sigma\{x_s,\ t+\varepsilon \le s \le T\} \ .$$

שם לב שההגדרה כמעט סימטרית ביחס להיפוך בזמן - פרט לדרישה הטכנית כי $\varepsilon > 0$ (ולא $= 0$).

הגדרה 6.2 תהליך $\{x_t\}$ נקרא מרקובי חזק אם לכל זמן עצירה τ ו- $\varepsilon > 0$ מתקיים

$$\sigma\{x_s,\ 0 \le s \le \tau\} \underset{\sigma\{x_\tau\}}{\perp} \sigma\{x_s,\ \tau+\varepsilon \le s \le T\} \ .$$

תהליך מרקובי חזק בעל פונקציות מדגם רציפות נקרא תהליך דיפוזיה diffusion process. נבדוק בצורה אינטואיטיבית מרקוביות של פתרון של משוואה סטוכסטית. נניח ש-m, σ מקיימים תנאי ליפשיץ, והתהליך מוגדר על ידי

$$dx_t = m(x_t)\,dt + \sigma(x_t)\,dw_t \ . \tag{6.2}$$

באופן אינטואיטיבי נקבל את x_t על ידי

$$x_{t+\varepsilon} = x_t + m(x_t)\cdot\varepsilon + \sigma(x_t)\cdot(w_{t+\varepsilon} - w_t) \ . \tag{6.3}$$

בהנתן x_t המ"א $x_{t+\varepsilon}$ הוא פונקציה של $w_{t+\varepsilon} - w_t$, ומתכונות התנועה הבראונית גודל זה בלתי תלוי סטטיסטית בעבר, כלומר התהליך מרקובי.

ההוכחה המדויקת של מרקוביות (וכן של מרקוביות חזקה) נעשית על ידי בניית פתרון המקיים את תכונת המרקוביות ב- t מסויים. על ידי שימוש בתכונת יחידות הפתרון מסיקים מרקוביות ב- t שרירותי של כל פתרון.

המשוואה האחורית של קולמוגורוב נתון תהליך $\{x_t\}$ הפותר את (6.2). נניח שנתונות פונקציות $f(x), u(t,x)$ המקיימות $u(0,x) = f(x)$ ובנוסף $u(t,x), u'_t(t,x), u'_x(t,x), u''_{xx}(t,x)$ רציפות וחסומות עבור $0 \le t \le T$ לכל x.

משפט 6.3 אם u מקיימת את המשוואה הדיפרנציאלית החלקית

$$\frac{\partial u(t,x)}{\partial t} = m(x)\frac{\partial u(t,x)}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2 u(t,x)}{\partial^2 x} \tag{6.4}$$

עם תנאי ההתחלה $u(0,x) = f(x)$, אזי (עבור x_t הפותר את (6.2))

$$u(t,x) = \mathbb{E}\left[\, f(x_t) \;|\; x_0 = x\right] \doteq \mathbb{E}_x[f(x_t)]\ . \tag{6.5}$$

הערה 6.4 את המשוואה הדיפרנציאלית אפשר לרשום בקיצור בצורה הבאה:

$$\frac{\partial u(t,x)}{\partial t} = Lu \qquad L = m(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial^2 x} \tag{6.6}$$

כאן L הוא אופרטור הפועל על פונקציה ומחזיר פונקציה. הוא נקרא "הגנרטור האינפיניטסימלי" infinitesimal generator של המשוואה החלקית (6.4) וכן של המשוואה הדיפרנציאלית הסטוכסטית (6.2). שים לב כי L מתאר את שתי המשוואות.

הערה 6.5 נוסחה (6.5) מציעה שיטה לחשב פתרונות של משוואה דיפרנציאלית חלקית על ידי סימולצית מונטה-קרלו. כלומר, עבור כל תנאי התחלה נגריל מספר רב של פתרונות של המשוואה הסטוכסטית, ונקרב את התוחלת המותנית על ידי ממוצע אמפירי של התוצאות.

הוכחה: נקבע זמן t ונקודה כלשהי x_0. נגדיר פונקציה $\Psi(s,x) = u(t-s,x)$ (שים לב שכאן t הוא פרמטר קבוע, ומשתנה הזמן הוא s). מנוסחת איטו,

$$\begin{aligned}\Psi(t,x_t) = \Psi(0,x_0) &+ \int_0^t \frac{\partial \Psi}{\partial s}(s,x_s)\,ds + \int_0^t \frac{\partial \Psi}{\partial x}(s,x_s)\cdot m(x_s)\,ds \\ &+ \frac{1}{2}\int_0^t \frac{\partial^2 \Psi}{\partial^2 x}(s,x_s)\cdot \sigma^2(x_s)\,ds + \int_0^t \frac{\partial \Psi}{\partial x}(s,x_s)\cdot\sigma(x_s)\,dW_s\ .\end{aligned} \tag{6.7}$$

לפי ההגדרה (6.6) של האופרטור L ניתן לרשום משוואה זו כך:

$$\Psi(t,x_t)=\Psi(0,x_0)+\int_0^t\left(\frac{\partial}{\partial s}+L\right)\Psi(s,x_s)\,ds+\int_0^t\frac{\partial\Psi}{\partial x}(s,x_s)\cdot\sigma(x_s)\,dW_s\ . \tag{6.8}$$

נשים לב כי

$$\frac{\partial\Psi(s,x)}{\partial s}=-\frac{\partial u(t-s,x)}{\partial t}\ . \tag{6.9}$$

כיוון ש- u מקיימת את (6.4), האנטגרל הראשון מתאפס שכן $\left(\frac{\partial}{\partial s}+L\right)u\equiv 0$. ניקח תוחלת: כיוון שלאנטגרל הסטוכסטי תוחלת אפס, גם האיבר האחרון מתאפס. לכן

$$\mathbb{E}_{x_0}\,\Psi(t,x_t)=\Psi(0,x_0) \tag{6.10}$$

מהגדרת Ψ נקבל

$$\mathbb{E}_{x_0}\,\Psi(t,x_t)=\mathbb{E}_{x_0}\,u(0,x_t)=\mathbb{E}_{x_0}\,f(x_t) \tag{6.11}$$

$$\Psi(0,x_0)=u(t,x_0) \tag{6.12}$$

וכיוון ש- t ו- x_0 שרירותיים, הוכחנו את המשפט.

מכאן ניתן, תחת תנאים מסויימים, לקבל נוסחה (כלומר תאור על ידי משוואה דיפרנציאלית חלקית) עבור הצפיפות של הפתרון של (6.2). . לפני הפיתוח, נציג טענת-עזר. יהי C_0^2 אוסף הפונקציות f אשר מתאפסות מחוץ לאינטרוול סופי (גודל האנטרוול תלוי בפונקציה) ובעלות שתי נגזרות רציפות (נובע מהדרישות כי f ושתי נגזרותיה חסומים).

טענה 6.6 אם הפונקציות α_1,α_2 מקיימות $\int_{-\infty}^{\infty}\alpha_2(x)f(x)\,dx=\int_{-\infty}^{\infty}\alpha_1(x)f(x)\,dx$ עבור כל $f\in C_0^2$, אזי $\alpha_1=\alpha_2$ כמעט בכל x (לבג). אם בנוסף α_1,α_2 רציפות אזי הן שוות.

נניח שקיימת צפיפות p עבור ערכי התהליך בכל רגע t (התלויה בהתניה - שהיא המצב x ברגע $t=0$), שהיא רציפה ב- y כך שלכל פונקציה $f\in C_0^2$ מתקיים

$$\mathbb{E}_x(f(x_t))=\int_{-\infty}^{\infty}f(y)p(y,t\mid x)\,dy\ . \tag{6.13}$$

על ידי החלפת סדר גזירה ואינטגרציה ומהגדרת הצפיפות,

$$\frac{\partial}{\partial t}\,\mathbb{E}_x[f(x_t)] = \int_{-\infty}^{\infty} f(y)\frac{\partial}{\partial t}\,p(y,t \mid x)\,dy \tag{6.14}$$

$$\frac{\partial}{\partial x}\,\mathbb{E}_x[f(x_t)] = \int_{-\infty}^{\infty} f(y)\frac{\partial}{\partial x}\,p(y,t \mid x)\,dy \tag{6.15}$$

$$L\,\mathbb{E}_x[f(x_t)] = \int_{-\infty}^{\infty} f(y) Lp(y,t \mid x)\,dy \tag{6.16}$$

$$\int_{-\infty}^{\infty} f(y)\frac{\partial}{\partial t}\,p(y,t \mid x)\,dy = \int_{-\infty}^{\infty} f(y) Lp(y,t \mid x)\,dy\ . \tag{6.17}$$

כאשר השוויון האחרון נובע מהתוצאה שקבלנו $\partial/\partial t\,\mathbb{E}_x[f(x_t)] = L\,\mathbb{E}_x[f(x_t)]$. כיוון שהאנטגרלים שווים לכל f, אזי מהנחת הרציפות ומטענה 6.6 האנטגרנדים שווים כלומר

$$\frac{\partial}{\partial t}\,p(y,t \mid x) = Lp(y,t \mid x) = m(x)\frac{\partial}{\partial x}p(y,t \mid x) + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial^2 x}p(y,t \mid x)\ . \tag{6.18}$$

מהגדרת הצפיפות המותנית מתקיים תנאי ההתחלה

$$p(y,0 \mid x) = \delta(y-x) \tag{6.19}$$

כיוון שבזמן $t=0$ מתקיים $x_{t_0} = x_0$.

ניתן לרשום משוואה זו ברישום מקוצר כ-$\frac{\partial}{\partial t}p(y,t \mid x) = Lp(y,t \mid x)$. חשוב לשים לב כי אופרטור הגזירה פועל ביחס למשתנה ההתניה (תנאי ההתחלה): במשוואה זו משתנה הצפיפות y הוא קבוע.

המשוואה (6.4) נקראת, בגלל משוואה (6.18), המשוואה האחורית של קולמוגורוב.

<u>נוסחת דינקין</u> לפתרון x_t של המשוואה

$$dx_t = m(x_t)\,dt + \sigma(x_t)\,dw_t \tag{6.20}$$

נפעיל את נוסחת איטו. עבור פונקציה $v(x)$ אפשר לרשום

$$v(x_t) = v(x_0) + \int_0^t \left[v'_x(x_s)m(x_s) + \frac{1}{2}v''_{xx}(x_s)\sigma^2(x_s)\right]\,ds + \int_0^t v'_x(x_s)\sigma(x_s)\,dw_s \tag{6.21}$$

$$= v(x_0) + \int_0^t (Lv)(x_s)\,ds + \int_0^t v'_x(x_s)\sigma(x_s)\,dw_s\ . \tag{6.22}$$

אם τ הוא זמן עצירה, אזי (בתנאים מתאימים על הפונקציה וזמן העצירה) נקבל מהמשוואה הקודמת

$$v(x_{t\wedge\tau}) = v(x_0) + \int_0^t \mathbf{1}_{\{\tau\}}(s)(Lv)(x_s)\,ds + \int_0^t \mathbf{1}_{\{\tau\}}(s)v'_x(x_s)\sigma(x_s)\,dw_s \tag{6.23}$$

$$\mathbb{E}\,v(x_{t\wedge\tau}) = v(x_0) + \mathbb{E}\int_0^{t\wedge\tau}(Lv)(x_s)\,ds\ . \tag{6.24}$$

כעת אם נשאיף את $t\to\infty$ נקבל

$$\mathbb{E}\,v(x_\tau) = v(x_0) + \mathbb{E}\int_0^{\tau}(Lv)(x_s)\,ds\ . \tag{6.25}$$

זוהי נוסחת דינקין Dynkin.

הערה 6.7 מהשוויון (6.21)--(6.22) אפשר להגיע למסקנה הבאה. לכל $f\in C_0^2$ התהליך $M_f(t)$ המוגדר על ידי

$$M_f(t) \doteq f(x_t) - \int_0^t (Lf)(x_s)\,ds \tag{6.26}$$

הוא מרטינגל.

מסתבר כי גם הכיוון ההפוך מכון: נניח שנתון אופרטור דיפרנציאלי מסדר שני L שהוא מהצורה

$$L = m(x)\frac{\partial}{\partial x} + \frac{1}{2}a(x)\frac{\partial^2}{\partial x^2} \tag{6.27}$$

כאשר $a(x) = \sigma^2(x)$ עבור פונקציות כלשהן m,σ. אם התהליך $\{x_t\}$ מקיים את התכונה הבאה: לכל $f\in C_0^2$, התהליך $M_f(t)$ הוא מרטינגל, אזי בהכרח $\{x_t\}$ פותר את המשוואה הסטוכסטית המוגדרת על ידי m,σ

$$dx_t = m(x_t)\,dt + \sigma(x_t)\,dw_t\ . \tag{6.28}$$

גישה זו למשוואות סטוכסטיות נקראת "ניסוח המרטינגל" *martingale formulation*. היא פותחה על ידי *D. Stroock, S.R.S. Varadhan*. חשיבות גישה זו היא כפולה. ראשית היא נותנת כלי לבדוק כי גבול מסויים מקיים מד"ס. שנית, היא מאפשרת הרחבת המושג מד"ס כאשר הטווח של התהליך אינו בהכרח מרחב אוקלידי.

הזמן הממוצע לצאת מקבוצה חסומה. **נתונה קבוצה G פתוחה וחסומה במרחב $\mathbb{R}^n$. נסמן ב-∂G את השפה של G: אלו הנקודות הנמצאות הן בסגור של G והן במשלים של G (שהוא סגור). נסמן**

ב- $\tau_G = \tau = \inf\{t : x_t \in \partial G\}$ **את הזמן הראשון שהתהליך הגיע לשפה של** G**. נניח שלכל** $x \in G$ **הזמן הממוצע של תהליך** x_t **(הפותר מד"ס) להגיע לשפה סופי:** $\mathbb{E}\,\tau < \infty$.

טענה 6.8 ייצוג של $\mathbb{E}_x\,\tau$: אם קיימת פונקציה $v(x)$ הפותרת את המד"ח

$$(Lv)(x) = -1 \qquad x \in G \tag{6.29}$$

$$v(x) = 0 \qquad x \in \partial G \tag{6.30}$$

אזי $v(x) = \mathbb{E}_x\,\tau$.

הוכחה: מנוסחת דינקין

$$\mathbb{E}\,v(x_{t\wedge\tau}) = v(x_0) - \mathbb{E}\int_0^{t\wedge\tau}\ ds\ . \tag{6.31}$$

נשאיף את $t \to \infty$ **ונקבל עבור מצב התחלתי** $x_0 = x$

$$\mathbb{E}\,v(x_\tau) = v(x) - \mathbb{E}\,\tau\ . \tag{6.32}$$

אבל $x_\tau \in \partial G$ **ולכן** $v(x_\tau) = 0$ **ולכן** $v(x) = \mathbb{E}\,\tau$.

דוגמה 6.9 נתייחס למודל העקיבה שבפרק המבוא. עבור תנועה בראונית, נחשב את זמן הבריחה כאשר $G = [-c, c]$.

$$L = \frac{1}{2}\frac{\partial^2}{\partial x^2} \tag{6.33}$$

$$Lv(x) = -1 \qquad x \in (-c, c) \tag{6.34}$$

$$v(-c) = v(c) = 0\ . \tag{6.35}$$

הפתרון הוא פונקציה ריבועית ב- x

$$v(x) = c^2 - x^2\ . \tag{6.36}$$

דוגמה 6.10 בקליטת שידור מאופנן FM אנו מעוניינים לסנכרן את המקלט עם המשדר. לשם כך נרצה לעקוב אחרי הפאזה של הגל הנושא. זאת ניתן לעשות על ידי מעגל עקיבה דינמי: אם נגלה שגיאה של α_t

בפאזה, נתקן של הערכת הפאזה במהירות $m(\alpha_t)$ כלומר המשוואה עבור השגיאה תהייה

$$\frac{d\alpha_t}{dt} = m(\alpha_t) \ . \tag{6.37}$$

אולם אם יש הפרעות ורעשים במדידה, נקבל את המודל הבא עבור שגיאת הפאזה:

$$d\alpha_t = m(\alpha_t)\, dt + \beta\, dw_t \tag{6.38}$$

כאשר β מאפיין את הגודל הייחסי של הרעשים. גם כאן מעניינת השאלה: כמה זמן בממוצע לוקח עד ל"שבירת נעילה"? נניח שחוג הנעילה מפסיק לפעול אם מגיעים לשגיאה $\alpha_t = c$ (ברור כי $c < 180^o$). לפי נוסחת דינקין, $\mathbb{E}_x\, \tau = v(x)$ כאשר $v(x)$ פותר את

$$m(x)\frac{dv}{dx} + \frac{1}{2}\beta^2\frac{d^2v}{dx^2} = -1 \qquad x \in (-c, c) \tag{6.39}$$

$$v(-c) = v(c) = 0 \ . \tag{6.40}$$

לעיתים חשוב לא רק מתי נצא מקבוצה G אלא גם היכן על שפת הקבוצה נצא. כדי לחשב זאת נשאל: כיצד לחשב את $\mathbb{E}_x(\psi(x_\tau))$ כאשר $\psi(y)$ מודדת "כמה כדאי" לצאת בנקודה $y = x_\tau \in \partial G$. נתבונן בפתרון v של המשוואה

$$Lv(x) = 0 \qquad x \in G \tag{6.41}$$

$$v(x) = \psi(x) \qquad x \in \partial G \ . \tag{6.42}$$

נפעיל את נוסחת דינקין: נקבל, כמו במקרה הקודם $\mathbb{E}_x\, v(x_\tau) = v(x)$. כיוון ש-$x_\tau \in \partial G$ נקבל כי $v(x_\tau) = \psi(x_\tau)$ ולכן $\mathbb{E}_x\, \psi(x_\tau) = v(x)$. בפרט אם נקח באופן פורמלי $\psi(x) = \delta(x - x_0)$ נקבל את צפיפות היציאה בנקודה x_0.

<u>נוסחת פיינמן-כץ Feynman Kac formula</u>.

תהי G קבוצה פתוחה וחסומה, ו-x_t תהליך עם גנרטור L. נגדיר $\tau = \inf\{t : x_t \in \partial G\}$-הזמן הראשון שהתהליך הגיע לשפה של G. נתונות פונקציות ϕ על ∂G ו-g, v על הסגור של G, כאשר $v \geq 0$. נגדיר פונקציה f על ידי הנוסחה

$$f(x) = \mathbb{E}_x \int_0^\tau g(x_t) e^{-\int_0^t v(x_s)\, ds}\ dt + \mathbb{E}_x\, \phi(x_\tau) e^{-\int_0^\tau v(x_s)\, ds} \ . \tag{6.43}$$

תרגיל 6.11 תחת תנאים טכניים מתאימים, הפונציה f לעיל היא הפתרון היחיד של המד"ח

$$(Lf)(x) - v(x)f(x) = -g(x) \qquad x \in G \tag{6.44}$$

$$\lim_{x \to a} f(x) = \phi(a) \qquad a \in \partial G \ . \tag{6.45}$$

רמז: הגדר תהליך דיפוזיה חדש $\tilde{x}_t = \begin{pmatrix} x_t \\ \xi_t \end{pmatrix}$ כאשר $\xi_t = \int_0^t v(x_s)\,ds$ והשתמש בנוסחת דינקין עבור הפונקציה $f(x)e^{-\xi}$ בדומה לפיתוח שעשינו עבור $v = 0$ ו-$g = 1$.

נוסחה (6.43) **נקראת נוסחת פיינמן-כץ.**

תרגיל 6.12 הראה כי פונקצית המומנט f_m מסדר m המוגדרת על ידי

$$f_m(x) \doteq \mathbb{E}_x \left(\int_0^\tau v(x_s)\,ds \right)^m \tag{6.46}$$

מקיימת את המד"ח

$$Lf_m(x) + mv(x)f_{m-1}(x) = 0 \ . \tag{6.47}$$

הסק כי ניתן לחשב את f_m על ידי פתרון סדרה של מד"ח.

<u>הנוסחה הקידמית של קולמוגורוב, היא נוסחת פוקר-פלנק Fokker-Planck equation.</u>

הנוסחה האחורית נתנה משוואה אשר מקיימת הצפיפות המותנית: זוהי מד"ח במשתנה הזמן t ובמשתנה ההתניה (נקודת ההתחלה). קיימת גם נוסחה קדמית - שהיא מד"ח חלקית במשתנה הזמן t ובמשתנה הקדמי - המצב. לצורך הפיתוח נניח קיום צפיפות $p(y,t \mid x)$ בעלת נגזרות רציפות כנדרש.

טענה 6.13 תחת תנאים מתאימים, הצפיפות של פתרון המד"ס מקיימת

$$\frac{\partial p}{\partial t}(y,t \mid x) = -\frac{\partial}{\partial y}[m(y)p(y,t \mid x)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(y)p(y,t \mid x)] \tag{6.48}$$

הערה 6.14 האופרטור $L^* \doteq -\frac{\partial}{\partial y}m(y) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\sigma^2(y)$ נקרא האופרטור הדואלי (הפורמלי) של הגנרטור $L = m(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}$. שים לב שקיום האופרטור L^* דורש קיום נגזרות ל- m, σ בעוד שתנאי זה אינו דרוש להגדרת L. באופן כללי המשוואה האחורית קיימת בתנאים חלשים יותר מזו הקדמית.

הוכחה: נבחר פונקציה $f \in C_0^2$ ונפעיל את נוסחת איטו: כיוון ש- x_t מקיים את המד"ס,

$$\frac{1}{\varepsilon}\,\mathbb{E}[f(x_{t+\varepsilon}) - f(x_t)] = \frac{1}{\varepsilon}\,\mathbb{E}\int_t^{t+\varepsilon} (Lf)(x_s)\,ds \tag{6.49}$$

וכאשר $\varepsilon \downarrow 0$ נקבל

$$\frac{\partial}{\partial t}\,\mathbb{E}[f(x_t)] = \mathbb{E}(Lf)(x_t)\ . \tag{6.50}$$

כעת, אם קיימת צפיפות מותנית $p(y,t \mid x)$ אזי אפשר לרשום נוסחה זו כך:

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} f(y)p(y,t \mid x)\,dy = \int_{-\infty}^{\infty}\left[p(y,t \mid x)\left(m(y)\frac{df}{dy}(y) + \frac{1}{2}\sigma^2(y)\frac{d^2 f}{dy^2}(y)\right)\right]\,dy\ . \tag{6.51}$$

נבצע כעת אינטגרציה בחלקים על האיבר הראשון בצד ימין:

$$\begin{aligned}\int_{-\infty}^{\infty} m(y)p(y,t \mid x)\frac{df}{dy}(y)\,dy& \\ &= m(y)p(y,t \mid x)f(y)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{\partial}{\partial y}[m(y)p(y,t \mid x)]f(y)\,dy\ .\end{aligned} \tag{6.52}$$

אולם $f(y) = 0$ עבור $|y|$ גדול מספיק, ולכן הביטוי הראשון מימין מתאפס. בצורה דומה כיוון ש-$df/dy = 0$ עבור $|y|$ גדול מספיק, נקבל על ידי שתי אינטגרציות בחלקים

(6.53)

$$\int_{-\infty}^{\infty}\sigma^2(y)p(y,t \mid x)\frac{d^2 f}{dy^2}(y)\,dy = \sigma^2(y)p(y,t \mid x)\frac{df}{dy}(y)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{d}{dy}[\sigma^2(y)p(y,t \mid x)]\frac{df}{dy}\,dy$$

$$= -\frac{d}{dy}[\sigma^2(y)p(y,t \mid x)]f(y)|_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\frac{d^2}{dy^2}[\sigma^2(y)p(y,t \mid x)]f(y)\,dy\ . \tag{6.54}$$

נציב את שתי התוצאות האחרונות ב-(6.51):

$$\begin{aligned}\frac{\partial}{\partial t}\int_{-\infty}^{\infty} p(y,t \mid x)f(y)\,dy& \\ &= -\int_{-\infty}^{\infty}\frac{\partial}{\partial y}[m(y)p(y,t \mid x)]f(y)\,dy + \frac{1}{2}\int_{-\infty}^{\infty}\frac{d^2}{dy^2}[\sigma^2(y)p(y,t \mid x)]f(y)\,dy\ .\end{aligned} \tag{6.55}$$

כיוון שהשוויון נכון לכל $f \in C_0^2$ נסיק כי

$$\frac{\partial}{\partial t}p(y,t \mid x) = -\frac{\partial}{\partial y}[m(y)p(y,t \mid x)] + \frac{1}{2}\frac{d^2}{dy^2}[\sigma^2(y)p(y,t \mid x)] \tag{6.56}$$

וסיימנו את ההוכחה.

מהמשוואה הקדמית (6.48) אפשר לקבל משוואה עבור הפילוג (הבלתי מותנה) של x_t. נסמן את הצפיפות של המשתנה x_0 ב-$p_0(x)$. אזי הצפיפות של x_t נתונה על ידי

$$p(y,t) = \int_{-\infty}^{\infty} p(y,t \mid x) p_0(x)\, dx \ . \tag{6.57}$$

נבצע אינטגרציה כזו על (6.48), ולאחר החלפת סדר גזירות ואינטגרציות נקבל

$$\frac{\partial p}{\partial t}(y,t) = -\frac{\partial}{\partial y}[m(y)p(y,t)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(y)p(y,t] \ . \tag{6.58}$$

זו מד"ח עבור הצפיפות של x_t כאשר תנאי ההתחלה של המשוואה הם $p(y,0) = p_0(y)$.

כעת אפשר לשאול: האם ניתן לבחור פילוג (צפיפות) התחלתי p_0 כך ש-$p(y,t) = p_0(y)$ לכל t, כלומר כך שהצפיפות החד מימדית אינה תלויה בזמן? כיוון שהתהליך הוא מרקובי והומוגני בזמן, תכונה כזו מבטיחה כי התהליך הוא סטציונרי. כדי לבדוק נניח שאכן קיימת צפיפות כזו. ואז $\frac{\partial}{\partial t}p(y,t) = 0$ וכן

$$\int_{-\infty}^{\infty} p(y,t \mid x) p_0(x)\, dx = p_0(y) \ . \tag{6.59}$$

בנוסף אם $p_0(y)$ שווה ל-$p(y,t)$ אזי p_0 מקיים את (6.48), ואם נרשום זאת במפורש ונבצע אינטגרציה ב- y נקבל

$$0 = \frac{\partial p_0}{\partial t}(y) = -\frac{\partial}{\partial y}[m(y)p_0(y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(y)p_0(y)] \tag{6.60}$$

$$\frac{1}{2}\frac{\partial}{\partial y}[\sigma^2(y)p_0(y)] = m(y)p_0(y) + c \tag{6.61}$$

עבור קבוע c כלשהו. נפתור תחילה משוואה זו למקרה $c = 0$ ובהמשך נשתמש בפתרון כדי לפתור את המקרה הכללי. עבור $c = 0$ נחשב

$$\frac{d}{dy}[\log \sigma^2(y)p_0(y)] = \frac{1}{\sigma^2(y)p_0(y)}\frac{d}{dy}[\sigma^2(y)p_0(y)] = 2\frac{m(y)}{\sigma^2(y)} \tag{6.62}$$

$$p_0(y) = \frac{c'}{\sigma^2(y)} \exp\left[\int_0^y 2\frac{m(z)}{\sigma^2(z)}\, dz\right] \tag{6.63}$$

כאשר הקבוע c' נקבע על ידי הדרישה ש- p_0 יהיה צפיפות, כלומר $\int_{-\infty}^{\infty} p_0(y)\, dy = 1$.

כדי למצוא פתרון עבור המקרה $c \neq 0$ ננחש פתרון מהצורה $p_0(y)c(y)$ כאשר p_0 הוא הפתרון שחישבנו לעיל עבור המקרה $c = 0$. מתוך המשוואה (6.61)

$$\frac{1}{2}\frac{\partial}{\partial y}[\sigma^2(y)p_0(y)c(y)] = m(y)p_0(y)c(y) + c \tag{6.64}$$

$$c(y)\frac{1}{2}\frac{\partial}{\partial y}[\sigma^2(y)p_0(y)] + \frac{1}{2}\sigma^2(y)p_0(y)\frac{\partial}{\partial y}c(y) = m(y)p_0(y)c(y) + c \tag{6.65}$$

$$c(y)\left[\frac{1}{2}\frac{\partial}{\partial y}[\sigma^2(y)p_0(y)] - m(y)p_0(y)\right] + \frac{1}{2}\sigma^2(y)p_0(y)\frac{\partial}{\partial y}c(y) = c \tag{6.66}$$

$$\frac{1}{2}\sigma^2(y)p_0(y)\frac{\partial}{\partial y}c(y) = c \tag{6.67}$$

כי p_0 מקיימת את המשוואה (6.61) עם $c = 0$. קיבלנו

$$\frac{\partial}{\partial y}c(y) = \frac{2c}{\sigma^2(y)p_0(y)} \ . \tag{6.68}$$

כלומר נחשב תחילה עבור $c = 0$, ואז נפתור את המשוואה האינטגרלית עבור $c(y)$ והצפיפות היא $p_0(y)c(y)$.

דוגמה 6.15 עבור המשוואה הלינארית

$$dx_t = mx_t\,dt + \beta\,dw_t \tag{6.69}$$

(m, β קבועים) נקבל

$$p_0(y) = \frac{c}{\beta^2}\exp\left[\frac{2}{\beta^2}\cdot m \cdot \int_0^y z\,dz\right] = \frac{c}{\beta^2}e^{my^2/\beta^2} \ . \tag{6.70}$$

כלומר, אם למשוואה לינארית יש צפיפות סטציונרית, אזי הצפיפות היא בהכרח גאוסית! שים לב כי אם $m \geq 0$ אזי p_0 הנתון על ידי משוואה זו אינה צפיפות, ובפרט אין צפיפות סטציונרית במקרה זה. זה בהחלט צפוי, שכן במקרה זה המשוואה, אפילו ללא רעש, איננה יציבה. עבור $m < 0$ יש פתרון והוא אכן צפיפות סטציונרית.

נניח שאנו קולטים אות $\{x_t,\ 0 \le t \le 1\}$ ואנו מעוניינים להחליט בין שתי השערות: השערה H_0 היא שקלטנו רעש בלבד, כלומר $\{x_t = w_t,\ 0 \le t \le 1\}$, והשערה H_1 היא שקלטנו אות בתוספת רעש, כאשר האות הוא דטרמיניסטי ונקלט דרך אינטגרל: $\left\{x_t = \int_0^t \phi_s\, ds + w_t,\ 0 \le t \le 1\right\}$.

אם נתייחס למדידות בזמנים בדידים בלבד, כלומר למדידות $\{a_{t_1}, a_{t_2}, \ldots , a_{t_n}\}$ של ערכי התהליך בזמנים אלו, ואם בנוסף הערכים האפשריים כתוצאת מדידה הם בדידים, אזי אפשר לקבל החלטה על סמך יחס הסבירות:

$$\Lambda = \Lambda(a_{t_1}, a_{t_2}, \ldots , a_{t_n}) = \frac{\mathbb{P}\left[x_{t_1} = a_{t_1}, x_{t_2} = a_{t_2}, \ldots , x_{t_n} = a_{t_n}\ \middle|\ H_1\right]}{\mathbb{P}\left[x_{t_1} = a_{t_1}, x_{t_2} = a_{t_2}, \ldots , x_{t_n} = a_{t_n}\ \middle|\ H_0\right]} \tag{7.1}$$

כלומר על ידי השוואה, עבור הערך הנתון של המדידות, האם סביר יותר שמקורן בהנחה אחת או אחרת.

לא ברור כיצד להרחיב זאת למדידות שאינן בדידות-כלומר למדידה של תהליך, וכן למדידות שערכיהן אינם בדידים. אחת הגישות היעילות לכך היא דרך משפט הייצוג של רדון-ניקודים-לבג.

נתון מרחב מדיד $(\Omega, \mathcal{F})$ ושתי מידות הסתברות $\mathbb{P}_0, \mathbb{P}_1$ ננסה ליצג את $\mathbb{P}_1$ בעזרת $\mathbb{P}_0$. נסמן את התוחלות המתאימות ב-$\mathbb{E}_0, \mathbb{E}_1$.

משפט 7.1 *Radon-Nikodym-Lebesgue*. נתונים $\mathbb{P}_0, \mathbb{P}_1, (\Omega, \mathcal{F})$. אזי קיים משתנה אקראי a על $(\Omega, \mathcal{F})$ ומאורע $N \in \mathcal{F}$ כך ש-$\mathbb{P}_0(N) = 0$ ולכל $A \in \mathcal{F}$

$$\mathbb{P}_1(A) = \int_A a(\omega)\, d\,\mathbb{P}_0(\omega) + \mathbb{P}_1(A \cap N) \tag{7.2}$$

או, בצורה שקולה,

$$\mathbb{P}_1(A) = \mathbb{E}_0[a(\omega)\mathbf{1}_{\{A\}}] + \mathbb{P}_1(A \cap N)\ . \tag{7.3}$$

בנוסף, אם x הוא מ"א איטגרבילי ביחס ל- $\mathbb{E}_1$ אזי

$$\mathbb{E}_1\, x = \int_\Omega x(\omega)\, d\,\mathbb{P}_1(\omega) = \int_{\Omega\setminus N} a(\omega)x(\omega)\, d\,\mathbb{P}_0(\omega) + \int_N x(\omega)\, d\,\mathbb{P}_1(\omega)\ . \tag{7.4}$$

פירוש הסימון $A \setminus B$ הוא - כל הנקודות ב- A שאינן ב- B, כלומר $A \setminus B = A \cap B^c$.

הגדרה 7.2 $a(\omega)$ נקראת נגזרת רדון-ניקודים של $\mathbb{P}_1$ לפי $\mathbb{P}_0$, מקובל לסמן

$$a(\omega) = \frac{d\,\mathbb{P}_1(\omega)}{d\,\mathbb{P}_0(\omega)} \tag{7.5}$$

אם $N = \emptyset$ נאמר ש- $\mathbb{P}_1$ רציפה בהחלט *absolutely continuous* ביחס ל- $\mathbb{P}_0$---נסמן זאת $\mathbb{P}_0 \ll \mathbb{P}_1$. הגדרה שקולה לרציפות בהחלט: לכל $B \in \mathcal{F}$ אם $\mathbb{P}_0(B) = 0$ אזי גם $\mathbb{P}_1(B) = 0$.

כמובן שנגזרת רדון ניקודים אינה נגזרת במובן של נגזרת של פונקציה.
אם $\mathbb{P}_1$ רציפה בהחלט absolutely continuous **ביחס ל- $\mathbb{P}_0$ אזי**

$$\mathbb{P}_1(A) = \int_A a(\omega)\,d\,\mathbb{P}_0(\omega) \tag{7.6}$$

$$= \mathbb{E}_0[a(\omega)\mathbf{1}_{\{A\}}] \tag{7.7}$$

$$\mathbb{E}_1\,x = \mathbb{E}_0[a(\omega)x(\omega)]\ . \tag{7.8}$$

בפרט נקבל

$$1 = \mathbb{E}_1\,1 = \mathbb{E}_0\,a(\omega)\ . \tag{7.9}$$

דוגמה 7.3 נניח ש- $\Omega = (-\infty, \infty)$ ולמידות יש צפיפויות p_0, p_1, ונניח ש- $\mathbb{P}_0\{(-\infty, 0]\} = 0$ וכן $p_0(\omega) > 0$ לכל $\omega > 0$. אזי $N = (-\infty, 0]$ ו-

$$a(\omega) = \begin{cases} 0 & \omega \in N \\ \frac{p_1(\omega)}{p_2(\omega)} & \omega \notin N\ . \end{cases} \tag{7.10}$$

תרגיל 7.4 נתונות פונקציות הפילוג הבאות על $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathbb{B})$

$$F_0(\omega) = \begin{cases} 0 & \omega < -0.4 \\ 1/2 & -0.4 \le \omega < 0.2 \\ 1 & 0.2 \le \omega\ , \end{cases} \tag{7.11}$$

$$F_1(\omega) = \begin{cases} 0 & \omega < -0.4 \\ 0.8 & -0.4 \le \omega < 0.2 \\ 1 & 0.2 \le \omega\ . \end{cases} \tag{7.12}$$

חשב את N ואת נגזרת רדון ניקודים $a(\omega)$ של $\mathbb{P}_1$ לפי $\mathbb{P}_0$. חשב את $\bar{a}(\omega) = d\,\mathbb{P}_0 / d\,\mathbb{P}_1$ שהיא נגזרת רדון ניקודים של $\mathbb{P}_0$ לפי $\mathbb{P}_1$. הראה כי ניתן לבחור את $a, \bar{a}$ כך ש- $a(\omega) \cdot \bar{a}(\omega) \equiv 1$.

<u>הוכחת משפט רדון-ניקודים-לבג.</u> **להוכחת המשפט נעזר במשפט הייצוג של ריץ Riesz-משפט 8.19. נגדיר מידת הסתברות חדשה על $(\Omega, \mathcal{F})$ על ידי**

$$\mu(A) \doteq \frac{1}{2}[\mathbb{P}_0(A) + \mathbb{P}_1(A)] \ . \tag{7.13}$$

נסמן ב- H את אוסף המ"א המקיימים

$$\int x^2(\omega)\, d\mu = \int x^2(\omega)\, d\,\mathbb{P}_0 + \int x^2(\omega)\, d\,\mathbb{P}_1 < \infty \tag{7.14}$$

או, בצורה שקולה,

$$\|x\|_\mu^2 \doteq \mathbb{E}_\mu\, x^2 = \frac{1}{2}\,\mathbb{E}_0\, x^2 + \frac{1}{2}\,\mathbb{E}_1\, x^2 < \infty \ . \tag{7.15}$$

זהו מרחב הילברט: המכפלה הפנימית במרחב זה היא

$$(x_1, x_2) = \mathbb{E}_\mu[x_1 \cdot x_2] \ . \tag{7.16}$$

נגדיר פונקציונל

$$f(x) \doteq \int x\, d\,\mathbb{P}_1 = \mathbb{E}_1\, x \ . \tag{7.17}$$

זהו בבירור פונקציונל לינארי, והוא חסום שכן מאי שוויון ינסן ואי שוויון שוורץ,

$$|\,\mathbb{E}_1\, x| \le \mathbb{E}_1\, |x| \le \mathbb{E}_0\, |x| + \mathbb{E}_1\, |x| = 2\,\mathbb{E}_\mu\, |x| = 2(|x|, 1) \le 2\|x\|_\mu \ . \tag{7.18}$$

לכן ממשפט ריץ נובע כי קיים אלמנט ב- H (כלומר מ"א) λ כך ש-

$$f(x) = \mathbb{E}_1\, x = (x, \lambda) = \int x\lambda\, d\mu = \mathbb{E}_\mu[x\lambda] \ . \tag{7.19}$$

כמובן ש- λ אינו תלוי ב- x. **נשים לב כי** $0 \le \lambda \le 2$ **בהסתברות** 1 **לפי** μ, **כי אם** $A = \{omega : \lambda(\omega) < 0\}$ **אזי** $\mathbf{1}_{\{A\}}$ **מ"א ו-**

$$0 \le \mathbb{E}_1\, \mathbf{1}_{\{A\}} = \int \mathbf{1}_{\{A\}}\lambda\, d\mu \le 0 \ . \tag{7.20}$$

לכן $\mathbb{P}_1(A)=0$ וכיוון ש- $\lambda\mathbf{1}_{\{A\}}<0$ לכל ω קיבלנו כי $\mu(A)=0$. מצד שני אם $B=\{\omega:\lambda>2+\varepsilon\}$ עבור $\varepsilon>0$ כלשהו, אזי

$$\mathbb{E}_1\,\mathbf{1}_{\{\lambda>2+\varepsilon\}}=\mathbb{E}_\mu[\lambda\mathbf{1}_{\{\lambda>2+\varepsilon\}}] \tag{7.21}$$

$$=\frac{1}{2}\,\mathbb{E}_0[\lambda\mathbf{1}_{\{\lambda>2+\varepsilon\}}]+\mathbb{E}_1[\lambda\mathbf{1}_{\{\lambda>2+\varepsilon\}}] \tag{7.22}$$

$$\geq\left(1+\frac{\varepsilon}{2}\right)\mathbb{E}_1\,\mathbf{1}_{\{\lambda>2+\varepsilon\}}\ . \tag{7.23}$$

אך זה יתכן רק אם $\mathbb{E}_1\,\mathbf{1}_{\{\lambda>2+\varepsilon\}}=0$. אבל אז

$$0=\mathbb{E}_1\,\mathbf{1}_{\{\lambda>2+\varepsilon\}}=\mathbb{E}_\mu\,\lambda\mathbf{1}_{\{\lambda>2+\varepsilon\}}\geq\mathbb{E}_\mu\,\mathbf{1}_{\{\lambda>2+\varepsilon\}}\geq 0 \tag{7.24}$$

ולכן גם $\mu(B)=0$.

כעת לכל $x\in H$

$$\int x\,d\,\mathbb{P}_1=\int x\lambda\,d\mu=\int\frac{1}{2}x\lambda\,d\,\mathbb{P}_0+\int\frac{1}{2}x\lambda\,d\,\mathbb{P}_1\ . \tag{7.25}$$

נפעיל נוסחה זו על המ"א $x\lambda/2$ ונקבל

$$\int x\,d\,\mathbb{P}_1=\frac{1}{2}\int x\lambda^2\,d\mu+\frac{1}{2}\int x\lambda\,d\,\mathbb{P}_0 \tag{7.26}$$

$$=\frac{1}{4}\int x\lambda^2\,d\,\mathbb{P}_1+\int[x\lambda/2+x\lambda^2/4]\,d\,\mathbb{P}_0 \tag{7.27}$$

$$=\int x(\lambda/2)^n\,d\,\mathbb{P}_1+\int x[\lambda/2+(\lambda/2)^2+\ldots+(\lambda/2)^n]\,d\,\mathbb{P}_0\ . \tag{7.28}$$

נסמן $N\doteq\{\omega:\lambda(\omega)=2\}$. המשוואה האחרונה נכונה לכל מ"א ב- H: נשתמש בה עבור המ"א $x=\mathbf{1}_{\{N\}}$. נקבל

$$\int\mathbf{1}_{\{N\}}\,d\,\mathbb{P}_1=\int\mathbf{1}_{\{N\}}\,d\,\mathbb{P}_1+\int n\mathbf{1}_{\{N\}}\,d\,\mathbb{P}_0 \tag{7.29}$$

וזה יתכן רק אם $\int\mathbf{1}_{\{N\}}\,d\,\mathbb{P}_0=\mathbb{P}_0(N)=0$. נגדיר כעת

$$a(\omega)=\begin{cases}\sum_{n=1}^{\infty}\left(\frac{\lambda(\omega)}{2}\right)^n & \omega\notin N\\ b(\omega) & \omega\in N\end{cases} \tag{7.30}$$

כאשר b הוא מ"א שרירותי. כיוון ש- $\lambda(\omega) < 2$ עבור $\omega \notin N$ הטור מתכנס, וקל לראות שלפי הגדרה זו a הוא מ"א. נציב כעת מ"א x כלשהוא ב- (7.26) ונקבל

$$\int x \, d\mathbb{P}_1 = \lim_{n\to\infty} \int x(\lambda/2)^n \, d\mathbb{P}_1 + \lim_{n\to\infty} \int x[\lambda/2 + (\lambda/2)^2 + \ldots + (\lambda/2)^n] \, d\mathbb{P}_0 \tag{7.31}$$

$$= \lim_{n\to\infty} \int_N x(\lambda/2)^n \, d\mathbb{P}_1 + \lim_{n\to\infty} \int_{\Omega\setminus N} x(\lambda/2)^n \, d\mathbb{P}_1 + \int xa \, d\mathbb{P}_0 \tag{7.32}$$

$$= \int_N x \, d\mathbb{P}_1 + \int xa \, d\mathbb{P}_0 \tag{7.33}$$

כאשר ההצדקה למעברים היא כדלהלן. החלוקה בשורה הראשונה לשני גבולות ובשורה השנייה לשלושה גבולות תוצדק כשנראה ששלושת הגבולות קיימים. הביטוי הראשון אינו תלוי ב- n. הביטוי השני מתכנס לגבול הנתון בגלל התכנסות נשלטת (ע"י $2|x|$). כיוון שצד שמאל של המשוואה לא תלוי ב- n הביטוי השלישי מתכנס בהכרח. על ידי הפעלת נוסחה זו עבור החלק החיובי של x ובנפרד עבור החלק השלילי נקבל שגבולות אלו קיימים וסופיים, ומהפעלת התכנסות נשלטת על כל אחד נוכל להחליף סדר אינטגרל וגבול ולקבל את הביטוי האחרון.
נשים לב כי, כיוון ש- $\mathbb{P}_0(N) = 0$,

$$\int xa \, d\mathbb{P}_0 = \int_{\Omega\setminus N} xa \, d\mathbb{P}_0 \tag{7.34}$$

נחזור כעת לבעיית הגילוי. מהו חוק החלטה? נזכר שעלינו לבחור בין שתי מידות הסתברות $\mathbb{P}_0, \mathbb{P}_1$. בדוגמה שנתנו, $\mathbb{P}_0$ מתאר את פילוג האות $\{x_t = w_t\}$, ומידת ההסתברות $\mathbb{P}_1$ מתארת את פילוג האות $\{x_t = \int_0^t \phi_s \, ds + w_t$ עבור פונקציה ידועה ϕ.

חוק החלטה יהיה חלוקה של Ω לשני מאורעות A ו- A^c כך שמתוך המדידות נוכל לבחור את A - ובכך להחליט על H_0 - או לבחור את A^c ובכך להחליט על H_1. בבעיה זו ההנחה היא שידוע פילוג התחלתי על שתי ההשערות, כלומר ידועה הסתברות א-פריורי (לפני שביצענו מדידות) לנכונות כל אחת מההשערות. הסתברות השגיאה הכוללת היא לכן

$$\mathbb{P}(e) = \mathbb{P}[e \mid H_0] \cdot \mathbb{P}(H_0) + \mathbb{P}[e \mid H_1] \cdot \mathbb{P}(H_1) = \int_{A^c} d\mathbb{P}_0 \cdot \mathbb{P}(H_0) + \int_A d\mathbb{P}_1 \cdot \mathbb{P}(H_1) \ . \tag{7.35}$$

הדרך שבידנו למזער את הסתברות השגיאה היא על ידי מציאת A שתמזער את $\mathbb{P}(e)$. לפי משפט

$$\mathbb{P}(e) = \int_{A^c} d\,\mathbb{P}_0 \cdot \mathbb{P}(H_0) + \int_A a(\omega) d\,\mathbb{P}_0 \cdot \mathbb{P}(H_1) \cdot \mathbb{P}(H_1) + \mathbb{P}_1(N \cap A) \cdot \mathbb{P}(H_1)\ . \tag{7.36}$$

כיון ש- $\mathbb{P}_0(N) = 0$, לכל A אם נוריד ממנו את N כלומר נבחר את A כך ש- $N \subset A^c$ אזי שני האנטגרלים לא ישתנו, והביטוי האחרון יקטן. לכן תמיד כדאי לבחור את A לפי כלל זה. נבדוק כעת מקרה פרטי בו $\mathbb{P}(H_0) = \mathbb{P}(H_1) = 1/2$ כלומר ההתסברויות המוקדמות של שתי ההשערות שוות. נקבל במקרה זה

$$\mathbb{P}(e) = \frac{1}{2}\left[\int_{A^c} d\,\mathbb{P}_0 + \int_A a(\omega)\, d\,\mathbb{P}_0\right] \tag{7.37}$$

ולכן החלוקה של Ω ל- A ו- A^c שתמזער את השגיאה צריכה להתבסס על המבחן $a(\omega) < 1$ או $a(\omega) > 1$. כלומר $A = \{\omega : a(\omega < 1\}$ ואז

$$\mathbb{P}(e) = \frac{1}{2}\int_\Omega (a(\omega) \wedge 1)\, d\,\mathbb{P}_0(\omega)\ . \tag{7.38}$$

תרגיל 7.5 פתור את בעיית ההחלטה כאשר $\mathbb{P}(H_0) \neq \mathbb{P}(H_1)$.

תרגיל 7.6 יהי $\Omega = \mathbb{R}$ ו-

$$\mathbb{P}_0(\omega : \omega \leq \alpha) = \frac{1}{\sqrt{2\pi}}\} \int_0^\alpha e^{-\theta^2/2}\, d\theta\ . \tag{7.39}$$

נגדיר מידה חדשה $\mathbb{P}_1$ על ידי הדרישה $\mathbb{P}_1 \ll \mathbb{P}_0$ וכן, עבור קבועים c ו- $\sigma > 0$,

$$\frac{d\,\mathbb{P}_1}{d\,\mathbb{P}_0}(\omega) = \frac{\frac{1}{\sigma}\exp\left(-\frac{(\omega-c)^2}{2\sigma^2}\right)}{\exp\left(-\omega^2/2\right)}\ . \tag{7.40}$$

המשתנה האקראי $x(\omega) = \omega$ הוא גאוסי ומנורמל לפי $\mathbb{P}_0$. הראה כי $\tilde{x}(\omega) = \frac{x(\omega)-c}{\sigma^2}$ הוא גאוסי ומנורמל לפי $\mathbb{P}_1$. חשב את $\frac{d\,\mathbb{P}_1}{d\,\mathbb{P}_0}$.

תרגיל 7.7 הוכח או תן דוגמה נגדית לטענות הבאות.

א. אם $\mathbb{P}_0 \ll \mathbb{P}_1$ וכן $\mathbb{P}_1 \ll \mathbb{P}_2$ אזי $\mathbb{P}_0 \ll \mathbb{P}_2$ ובנוסף מתקיים

$$\frac{d\,\mathbb{P}_2}{d\,\mathbb{P}_0} = \frac{d\,\mathbb{P}_2}{d\,\mathbb{P}_1} \cdot \frac{d\,\mathbb{P}_1}{d\,\mathbb{P}_0} \ . \tag{7.41}$$

ב. אם $\mathbb{P}_0 \ll \mathbb{P}_1$ וכן $\mathbb{P}_1 \ll \mathbb{P}_0$ אזי $\mathbb{P}_1 = \mathbb{P}_0$.

אם $\mathbb{P}_0 \ll \mathbb{P}_1$ וכן $\mathbb{P}_1 \ll \mathbb{P}_0$ אזי נסמן $\mathbb{P}_1 \sim \mathbb{P}_0$ ונאמר כי $\mathbb{P}_0$ ו- $\mathbb{P}_1$ הן אקויולנטיות.

תרגיל 7.8 אם $\mathbb{P}_1 \sim \mathbb{P}_0$ אזי $\mathbb{P}_0(A) = 1$ אם ורק אם $\mathbb{P}_1(A) = 1$ ומתקיים בהסתברות 1 לפי כל אחת מהמידות

$$\frac{d\,\mathbb{P}_1}{d\,\mathbb{P}_0}(\omega) \cdot \frac{d\,\mathbb{P}_0}{d\,\mathbb{P}_1}(\omega) = 1 \ . \tag{7.42}$$

בבעית הגילוי היה נתון $(\Omega, \mathcal{F}, \mathbb{P}_0, \mathbb{P}_1)$ כאשר $\mathbb{P}_0$ ו- $\mathbb{P}_1$ מתארים רעש, ואות + רעש בהתאמה. נניח כעת כי נתונות מדידות רק עד זמן t. נסמן $G = \mathcal{F}_t$ כך ש- G מתאר מאורעות שקרו עד זמן t. נניח ש- $\mathbb{P}_1 \ll \mathbb{P}_0$: אזי בעית הגילוי בהנתן המידע $\mathcal{F}$ כרוכה בחישוב נגזרת רדון ניקודים a. ננסח כעת את הבעיה כאשר ידוע רק G.

יהי $\tilde{\mathbb{P}}_0$ ($\tilde{\mathbb{P}}_1$) הצימצום של $\mathbb{P}_0$ (של $\mathbb{P}_1$ בהתאמה) ל- G, כלומר אלו מידות הסתברות המוגדרות רק עבור מאורעות ב- G, וכך ש- $\tilde{\mathbb{P}}_0(A) = \mathbb{P}_0(A)$ וכן $\tilde{\mathbb{P}}_1(A) = \mathbb{P}_1(A)$ לכל $A \in G$. אזי אפשר לנסח את בעיית הגילוי דרך $(\Omega, \mathcal{F}, \tilde{\mathbb{P}}_0, \tilde{\mathbb{P}}_1)$ ולהגיע למסקנה כי, כיוון ש- $\tilde{\mathbb{P}}_1 \ll \tilde{\mathbb{P}}_0$, הבעיה כרוכה בחישוב נגזרת רדון ניקודים המתאימה, שנסמן ב- $\tilde{a}(\omega)$.

תרגיל 7.9 אם $\mathbb{P}_1 \ll \mathbb{P}_0$ אזי $\tilde{\mathbb{P}}_1 \ll \tilde{\mathbb{P}}_0$.

באופן כללי יותר, בהנתן סיגמה-שדה $G \subset \mathcal{F}$, נגדיר $\mathbb{P}_i,\ i = 0, 1$ ו- $\tilde{a}$ כמו לעיל. השאלה היא, מהו הקשר בין $a = d\,\mathbb{P}_1 / d\,\mathbb{P}_0$ לבין $\tilde{a} = d\tilde{\mathbb{P}}_1 / d\tilde{\mathbb{P}}_0$? בפרט, האם ניתן לחשב את $\tilde{a}$ על סמך ידיעת a?

טענה 7.10 $\tilde{a}(\omega) = \mathbb{E}_0[a(\omega) \mid G]$.

הוכחה: יהי x מ"א חסום ומדיד על G. אזי אם $\mathbb{P}_1 \ll \mathbb{P}_0$

$$\mathbb{E}_1(x) = \tilde{\mathbb{E}}_1(x) = \int_\Omega \tilde{a}(\omega) x(\omega)\, d\tilde{\mathbb{P}}_0(\omega) = \tilde{\mathbb{E}}_0(\tilde{a}x) = \mathbb{E}_0(\tilde{a}x) \tag{7.43}$$

כיוון שלפי ההגדרה, עבור מ"א אקראי מדיד G, המידות $\mathbb{E}_i$ ו- $\tilde{\mathbb{E}}_i$ יתנו אותה תוחלת, ובמקרה שלנו x ו- $\tilde{a}$ מדידים G. מצד שני, מהגדרת a מתקיים

$$\mathbb{E}_1(x) = \mathbb{E}_0(ax) = \mathbb{E}_0\left[\mathbb{E}_0(ax \mid G)\right] = \mathbb{E}_0\left[\mathbb{E}_0(a \mid G)x\right] \ . \tag{7.44}$$

לכן $\mathbb{E}_0(\tilde{a}x) = \mathbb{E}_0\left[\mathbb{E}_0(a \mid G)x\right]$ לכל x חסום ומדיד והטענה הוכחה.

תרגיל 7.11

א. הראה כי השוויון האחרון נובע, בהסתברות 1 לפי $\mathbb{P}_0$, משתי המשוואות הקודמות.

ב. אם $\mathbb{P}_0 \ll \mathbb{P}_1$ אזי a יחיד בהסתברות 1 לפי $\mathbb{P}_0$.

תרגיל 7.12 נניח כי $\omega = (x_1, x_2)$ כאשר $x_1, x_2 \in \mathbb{R}$. נניח כי ל- $\mathbb{P}_0$ צפיפות $p_0(x_1, x_2)$ ול- $\mathbb{P}_1$ צפיפות $p_1(x_1, x_2)$.

א. בטא את התנאי $\mathbb{P}_0 \ll \mathbb{P}_1$ על ידי תנאים על הצפיפויות.

ב. מצא ביטוי עבור a.

ג. יהי H אוסף הקבוצות

$$H = \{(x_1 \le \alpha,\ x_2 \in \mathbb{R}),\ \alpha \in \mathbb{R}\} \ . \tag{7.45}$$

חשב את $\tilde{a}$ והראה כי $\tilde{a} = \mathbb{E}_0[a \mid H]$.

כעת יהי $\{x_t,\ 0 \le t \le T\}$ תהליך אקראי. נסמן $\mathcal{F}_t = \sigma\{x_s : 0 \le s \le t\}$ ונסמן ב- $\tilde{\mathbb{P}}_{it}$ את הצמצום של $\mathbb{P}_i$ ל- $\mathcal{F}_t$. אזי מהטענה הקודמת, $\tilde{a}_t = d\tilde{\mathbb{P}}_{1t}/d\tilde{\mathbb{P}}_{0t}$ הוא $\mathcal{F}_t$ מרטינגל לפי $\mathbb{P}_0$. בנוסף $\tilde{a}_t \ge 0$ ומתכונת המרטינגל $\mathbb{E}_0\,\tilde{a}_t = \mathbb{E}_0\,a = \mathbb{E}_1\,1 = 1$.

הגדרה 7.13 המידות $\mathbb{P}_0$ ו- $\mathbb{P}_1$ נקראות ניצבות (סימון: $\mathbb{P}_0 \perp \mathbb{P}_1$) אם קיים מאורע N כך ש- $\mathbb{P}_0(N) = 0$ ו- $\mathbb{P}_1(N) = 1$ (השווה למשפט רדון-ניקודים-לבג *7.1*).

נשים לב כי אין קשר בין ניצבות במובן זה, ניצבות של מ"א, ואי תלות סטטיסטית.
אם המידות ניצבות אזי בהכרח

$$\mathbb{P}_0(N^c) = \mathbb{P}_1(N) = 1, \quad \mathbb{P}_0(N) = \mathbb{P}_1(N^c) = 0, \quad N \cup N^c = \Omega \ . \tag{7.46}$$

משפט 7.14 יהי $\{x_t\}$ ת"א גאוסי לפי $\mathbb{P}_0$ וגם לפי $\mathbb{P}_1$, כאשר הסיגמה-שדה נוצר על ידי התהליך. אזי או ש- $\mathbb{P}_1 \sim \mathbb{P}_0$ או ש $\mathbb{P}_0 \perp \mathbb{P}_1$.

תרגיל 7.15 נניח שלפי $\mathbb{P}_0$ התהליך $\{x_t\}$ הוא $x_t = w_t$ ואלו לפי $\mathbb{P}_1$ מתקיים $x_t = \alpha w_t$ כאשר $|\alpha| \neq 1$. אזי $\mathbb{P}_0 \perp \mathbb{P}_1$. רמז: הגדר את N להיות המאורע שהתנודה הריבועית של התהליך באינטרוול $[0,1]$ היא 1.

כיצד נחשב תוחלת מותנית לפי $\mathbb{P}_1$ מתוך ידיעת $\mathbb{P}_0$? נניח $\mathbb{P}_1 \ll \mathbb{P}_0$ ויהי a נגזרת רדון ניקודים כך ש- $\mathbb{E}_1 x = \mathbb{E}_0 ax$ לכל מ"א x. יהי $\mathcal{F}_1$ תת-סיגמה-שדה של $\mathcal{F}$, ו- x מ"א המקיים $\mathbb{E}_1 |x| < \infty$. נסמן $\hat{x} = \mathbb{E}_1(x \mid \mathcal{F}_1)$. אזי $\hat{x}$ מאופיין על ידי הדרישה כי הוא מדיד $\mathcal{F}_1$ ולכל מ"א y שהוא חסום ומדיד $\mathcal{F}_1$ מתקיים $\mathbb{E}_1[y(x-\hat{x})] = 0$. לכן $\mathbb{E}_0[ay(x-\hat{x})] = 0$ כאשר $a = d\,\mathbb{P}_1 / d\,\mathbb{P}_0$. על ידי החלקה נקבל מכאן

$$\mathbb{E}_0\left[\mathbb{E}_0(ayx - ay\hat{x}) \mid \mathcal{F}_1)\right] = 0 \tag{7.47}$$

$$\mathbb{E}_0\left[y\,\mathbb{E}_0(ax \mid \mathcal{F}_1)\right] = \mathbb{E}_0\left[y\hat{x}\,\mathbb{E}_0(a \mid \mathcal{F}_1)\right] \ . \tag{7.48}$$

כיוון שהשוויון מתקיים לכל y מדיד $\mathcal{F}_1$ וחסום נסיק כי

$$\mathbb{E}_0(ax \mid \mathcal{F}_1) = \hat{x}\,\mathbb{E}_0(a \mid \mathcal{F}_1) = \mathbb{E}_1(x \mid \mathcal{F}_1)\,\mathbb{E}_0(a \mid \mathcal{F}_1) \ . \tag{7.49}$$

לכן נוכל להגדיר

$$\mathbb{E}_1(x \mid \mathcal{F}_1) = \begin{cases} \frac{\mathbb{E}_0(ax|\mathcal{F}_1)}{\mathbb{E}_0(a|\mathcal{F}_1)} & \omega : \mathbb{E}_0(a \mid \mathcal{F}_1)(\omega) \neq 0 \\ 0 & \omega : \mathbb{E}_0(a \mid \mathcal{F}_1)(\omega) = 0 \end{cases} \tag{7.50}$$

כאשר הבחירה של הערך 0 בביטוי האחרון היא שרירותית. כדי שההגדרה תהייה עקבית צריך להוכיח כי $\mathbb{P}_1\{\omega : \mathbb{E}_0(a \mid \mathcal{F}_1)(\omega) = 0\} = 0$. נעשה זאת בעזרת משפט ההחלקה:

$$\mathbb{P}_1\{\omega : \mathbb{E}_0(a \mid \mathcal{F}_1)(\omega) = 0\} = \mathbb{E}_1\,\mathbf{1}_{\{\omega:\mathbb{E}_0(a|\mathcal{F}_1)(\omega)=0\}} \tag{7.51}$$

$$= \mathbb{E}_0[a\mathbf{1}_{\{\omega:\mathbb{E}_0(a|\mathcal{F}_1)(\omega)=0\}}] \tag{7.52}$$

$$= \mathbb{E}_0\Big\{\mathbb{E}_0[a\mathbf{1}_{\{\omega:\mathbb{E}_0(a|\mathcal{F}_1)(\omega)=0\}} \mid \mathcal{F}_1]\Big\} \tag{7.53}$$

$$= \mathbb{E}_0\Big\{\mathbf{1}_{\{\omega:\mathbb{E}_0(a|\mathcal{F}_1)(\omega)=0\}}\,\mathbb{E}_0[a \mid \mathcal{F}_1]\Big\} = 0 \ . \tag{7.54}$$

תרגיל 7.16 אם $\mathbb{P}_1 \ll \mathbb{P}_0$ אזי $(M_t, \mathcal{F}_t)$ הוא מרטינגל ביחס ל- $\mathbb{P}_1$ אם ורק אם $(a_t M_t, \mathcal{F}_t)$ הוא מרטינגל ביחס ל- $\mathbb{P}_0$ כאשר $a_t = \mathbb{E}_0[d\,\mathbb{P}_1 / d\,\mathbb{P}_0 \mid \mathcal{F}_t]$.

יהי $(\Omega, \mathcal{F}, \mathbb{P}_0)$ מרחב הסתברות ו- $\{\mathcal{F}_t\},\ 0 \le t \le T$ פילטרציה. תהי $(y_t, \mathcal{F}_t)$ תנועת בראון סטנדרטית. יהי ϕ_t תהליך אקראי מתואם, וכך ש- $\int_0^t \phi^2(s)\, ds < \infty$ בהסתברות 1 לפי $\mathbb{P}_0$. נגדיר

$$a_t = \exp\left(\int_0^t \phi_s\, dy_s - \frac{1}{2}\int_0^t \phi^2(s)\, ds\right) . \tag{7.55}$$

משפט 7.17 ($Girsanov$ גירסנוב) נגדיר מידה $\mathbb{P}_1$ על ידי

$$\mathbb{P}_1(A) = \int_A a_T(\omega)\, d\,\mathbb{P}_0(\omega) . \tag{7.56}$$

אם a_t הוא $\mathbb{P}_0$-מרטינגל אזי התהליך

$$\tilde{y}_t = y_t - \int_0^t \phi_s\, ds \tag{7.57}$$

הוא תנועה בראונית סטנדרטית לפי $\mathbb{P}_1$.

הוכחה: ניתן הוכחה חלקית במובן שנדלג עם פטים טכניים (ההוכחה מלאה למשל אם ϕ חסום). ברור כי a_t אינו שלילי ומתכונת המרטינגל $\mathbb{E}_0\, a_t = 1$ לכל t. כזי $\mathbb{P}_1$ היא מידת הסתברות. קל לראות כי $\mathbb{P}_1 \ll \mathbb{P}_0$ וכן כי $a_T = d\,\mathbb{P}_1 / d\,\mathbb{P}_0$. כיוון שכל התהליכים להלן רציפים, נשתמש במפשט לוי: מספיק להוכיח כי

$$\left(y_t - \int_0^t \phi_s\, ds\right) a_t = b_t a_t \tag{7.58}$$

$$\left(\left[y_t - \int_0^t \phi_s\, ds\right]^2 - t\right) a_t \tag{7.59}$$

הם תהליכי מרטינגל לפי $\mathbb{P}_0$. אולם לפי $\mathbb{P}_0$ התהליך y_t הוא תנועת בראון. ממשפט איטו, a_t מקיים

$$a_t = 1 + \int_0^t a_s\, dy_s . \tag{7.60}$$

נפעיל את נוסחת איטו הדו-ממדית על הפונקציה

$$u(a,b) = a \cdot b \tag{7.61}$$

$$da_t = a_t \phi_t \, dy_t \tag{7.62}$$

$$db_t = dy_t - \phi_t \, dt \tag{7.63}$$

$$u(a_t, b_t) \left(y_t - \int_0^t \phi_s \, ds \right) a_t \tag{7.64}$$

נקבל

$$a_t b_t = \int_0^t a_s \, db_s + \int_0^t b_s \, da_s + \frac{1}{2} \int_0^t 2(1 \cdot a_s \phi_s) \, ds \tag{7.65}$$

$$= \int_0^t a_s \, dy_s + \int_0^t b_s a_a \phi_s \, dy_s \ . \tag{7.66}$$

מתכונות האיטגרל הסטוכסטי, $a_t b_t$ הוא מרטינגל (לא נביא כאן הוכחה שהאינטגרנד הוא ב- L^2) והוכחנו עבור (7.58).

כדי לטפל ב- (7.59) נשים לב כי מנוסחת איטו

$$b_t^2 = 2 \int_0^t b_s \, db_s + t \tag{7.67}$$

ולכן נותר להראות כי

$$a_t \cdot \int_0^t b_s \, db_s \doteq a_t \cdot c_t \tag{7.68}$$

הוא מרטינגל, כאשר

$$da_t = a_t \phi_t \, dy_t \ , \qquad dc_t = b_t \, dy_t - b_t \phi_t \, dt \ . \tag{7.69}$$

נפעיל שוב את נוסחת איטו ונקבל

$$a_t c_t = \int_0^t c_s a_s \phi_s \, dy_s + \int_0^t a_s b_s \, dy_s - \int_0^t a_s b_s \phi_s \, ds + \frac{1}{2} \int_0^t 2(a_s \phi_s b_s) \, ds \tag{7.70}$$

$$= \int_0^t a_s (c_s \phi_s + b_s) \, dy_s \tag{7.71}$$

וקיבלנו שוב אינטגרל סטוכסטי שהוא מרטינגל.

כדי לפתח הבנה אינטואיטיבית למשפט גירסנוב, נניח ש- ϕ היא פונקציה דטרמיניסטית (לא אקראית) ורציפה. נניח שלפי $\mathbb{P}_0$ התהליך הוא תנועת בראון כלומר $y_t = w_t$, ואילו לפי $\mathbb{P}_1$ התהליך הוא סיגנל בתוספת תנועת בראון

$$y_t = w_t + \int_0^t \phi_s \, ds \ .$$

נקבע $\epsilon > 0$ ונסמן $\Delta y_{n\epsilon} \doteq y_{n\epsilon+\epsilon} - y_{n\epsilon}$ ו-

$$y^\epsilon \doteq \begin{pmatrix} \Delta y_\epsilon \\ \Delta y_{2\epsilon} \\ \vdots \\ \Delta y_{t-\epsilon} \end{pmatrix} \tag{7.72}$$

אזי לפי $\mathbb{P}_0$ וכן לפי $\mathbb{P}_1$ הווקטור y^ϵ הוא וקטור גאוסי, עם רכיבים בת"ס. בנוסף,

$$\mathbb{E}_0 \, \Delta y_{n\epsilon} = 0 \ , \qquad \mathbb{E}_1 \, \Delta y_{n\epsilon} \approx \epsilon \phi_{n\epsilon} \ . \tag{7.73}$$

כיוון של- y^ϵ יש צפיפות לפי שתי המידות, יחס הסבירות הוא יחס הצפיפויות, שהן גאוסיות ולכן

$$\frac{d\,\mathbb{P}_1^\epsilon}{d\,\mathbb{P}_0^\epsilon} = \frac{\exp\left(-\frac{1}{2\epsilon}\sum(\Delta y_{n\epsilon} - \epsilon\phi_{n\epsilon})^2\right)}{\exp\left(-\frac{1}{2\epsilon}\sum(\Delta y_{n\epsilon})^2\right)} \tag{7.74}$$

$$= \exp\left(-\frac{1}{2\epsilon}\sum\left[(\Delta y_{n\epsilon})^2 + (\epsilon\phi_{n\epsilon})^2 - 2\Delta y_{n\epsilon}\cdot\epsilon\phi_{n\epsilon} - (\Delta y_{n\epsilon})^2\right]\right) \tag{7.75}$$

$$= \exp\left(\sum \phi_{n\epsilon}\cdot\Delta y_{n\epsilon} - \frac{1}{2}\sum(\phi_{n\epsilon})^2\cdot\epsilon\right) \tag{7.76}$$

$$\to \exp\left(\int_0^t \phi_s \, dy_s - \frac{1}{2}\int_0^t \phi_s^2 \, ds\right) \tag{7.77}$$

כאשר $\epsilon \to 0$.

דוגמה 7.18 בבעית הגלוי קולטים את $x_0^T \doteq \{x_s, \ 0 \le s \le T\}$ כאשר

לפי $\mathbb{P}_0$ נקלט רעש $x_t = w_t$

לפי $\mathbb{P}_1$ נקלט אות בתוספת רעש כלומר $x_t = w_t + \int_0^t \phi_s \, ds$.

נשתמש במשפט גרסנובץ נסמן $y_t = w_t + \int_0^t \phi_s \, ds$. נניח תחילה כי ϕ הוא דטרמיניסטי. לפי משפט גרסנוב,

המידה $\tilde{\mathbb{P}}_1$ המוגדרת דרך

$$\frac{d\tilde{\mathbb{P}}_1}{d\,\mathbb{P}_0} = \exp\left\{\int_0^t \phi_s\,dy_s - \frac{1}{2}\int_0^t \phi_s^2\,ds\right\} \tag{7.78}$$

היא מידה של תהליך אשר אפשר לתאר כ-

$$y_t = \tilde{w}_t + \int_0^t \phi_s\,ds \tag{7.79}$$

כאשר $\tilde{w}_t$ היא תנועת בראון. לכן $\tilde{\mathbb{P}}_1 = \mathbb{P}_1$. כעת ניתן להרחיב את התוצאה למקרה בו ϕ_t תלוי במסלולי המדידה, כלומר הוא תהליך אקראי

$$\phi_t = \phi(t, y_0^t)\ . \tag{7.80}$$

תוצאות כאלו חשובות לבעיות תקשורת עם משוב. המסקנה היא כפולה: ראשית, $\mathbb{P}_1 \ll \mathbb{P}_0$-וזאת מהשוויון $\tilde{\mathbb{P}}_1 = \mathbb{P}_1$ ומהגדרת $\tilde{\mathbb{P}}_1$. שנית, את $d\,\mathbb{P}_1 / d\,\mathbb{P}_0$ ניתן לחשב בצורה מפורשת על סמך המדידה.

דוגמה 7.19 בתקשורת ספרתית עלינו להבחין בין שני סיגנלים, או בין שתי מידות:

$$\mathbb{P}_1 : x_t = w_t + \int_0^t s_a(s)\,ds \tag{7.81}$$

$$\mathbb{P}_2 : x_t = w_t + \int_0^t s_b(s)\,ds\ . \tag{7.82}$$

נגדיר שוב כי $x_t = w_t$ לפי $\mathbb{P}_0$, ואז

$$\frac{d\,\mathbb{P}_2}{d\,\mathbb{P}_1} = \frac{d\,\mathbb{P}_2}{d\,\mathbb{P}_0} \bigg/ \frac{d\,\mathbb{P}_1}{d\,\mathbb{P}_0} \tag{7.83}$$

$$= \exp\left\{\int_0^T (s_b(s) - s_a(s))\,dx_s - \frac{1}{2}\int_0^T (s_b^2(s) - s_a^2(s))\,ds\right\}\ . \tag{7.84}$$

הערה 7.20 אחת הדרישות במשפט גרסנוב היא ש- a_t יהיה מרטינגל. תנאים טכניים לכך ראה למשל בספרו של *Elliot* ע' 176--170.

<u>הרחבה ספציפית של משפט גרסנוב.</u>

יהיה Ω אוסף זוגות פונקציות רציפות על $[0,T]$ (ניתן להחליש את דרישת הרציפות). נסמן פונקציות על ידי $x_0^T = \{x_s,\ 0 \le s \le T\}$ ו- $w = \{x_0^T, y_0^T\}$. יהי $\mathcal{F}_t = \sigma(x_0^t, y_0^t)$ ונניח כי לפי $\mathbb{P}_0$

$$\mathbb{P}_0 : \quad \begin{cases} x_t = \phi_t \\ y_t = w_t \end{cases} \tag{7.85}$$

כאשר לפי $\mathbb{P}_0$ התהליכים מקיימים $w \perp \phi$. יהי ψ תהליך מדיד על $\mathcal{F}_t$ ונניח כי

$$a_t = \exp\left\{\int_0^t \phi_s\, dy_s - \frac{1}{2}\int_0^t \phi_s^2\, ds\right\} \tag{7.86}$$

הוא מרטינגל לפי $\mathbb{P}_0$. אזי $\mathbb{P}_1$ (המוגדר כמיקודם על ידי $d\,\mathbb{P}_1/d\,\mathbb{P}_0 = a_t$) מקיים

$$\mathbb{P}_1 : \begin{cases} x_t = \phi_t \\ y_t = \int_0^t \phi_s\, ds + \tilde{w}_s \end{cases} \tag{7.87}$$

כאשר חוק ההסתברות של $(\phi_0^T, \tilde{w}_0^T)$ לפי $\mathbb{P}_1$ (כאשר $\tilde{w}_t = y_t - \int_0^t \phi_s\, ds$) זהה לחוק ההסתברות של (ϕ_0^T, w_0^T) לפי $\mathbb{P}_0$.

כדי להוכיח זאת נשים לב כי לפי משפט גרסנוב $\tilde{w}_t$ היא תנועה בראונית. נותר לכן להוכיח כי לפי $\mathbb{P}_1$ התהליכים $\tilde{w}, \Phi$ הם בת"ס וכן שחוק הפילוג של ϕ לא השתנה. נבחר שתי פונקציות דטרמיניסטיות h, f. נסמן

$$c_t = \exp\left\{i\int_0^t f_s\phi_s\, ds + i\int -)^t h_s\, d\tilde{w}_s\right\} \tag{7.88}$$

$$B_t = \exp\left\{i\int_0^t f_s\phi_s\, ds + i\int -)^t h_s\, dy_s\right\}\ . \tag{7.89}$$

מספיק להוכיח כי לכל בחירה של h, f ממשיות וחסומות $\mathbb{E}_1\, c_t = \mathbb{E}_0\, B_t$, כי אלו הפונקציונלים האפייניים של (y, ϕ) לפי $\mathbb{P}_0$ ושל $(y - \int \phi_s\, ds, \phi$ לפי $\mathbb{P}_1$. נחשב

(7.90)

$$\mathbb{E}_1\, c_t = \mathbb{E}_0\, a_t c_t = \mathbb{E}_0[\mathbb{E}_0(a_t c_t \mid \phi_0^t)]$$

$$\mathbb{E}_0(a_t c_t \mid \phi_0^t) = \mathbb{E}_0\left(\exp\left[\int_0^t \phi_s\, dy_s - \frac{1}{2}\int_0^t \phi_s^2\, ds + i\int_0^t f_s\phi_s\, ds + i\int_0^t h_s\, d\tilde{w}_s\right]\middle|\, \phi_0^t\right)\ . \tag{7.91}$$

כיון ש- $d\tilde{w}_s = dy_s - \phi_s\,ds$ נקבל

$$\begin{aligned}
&\mathbb{E}_0(a_t c_t \mid \phi_0^t)\\
&\quad = \mathbb{E}_0\left(\exp\left[\int_0^t (\phi_s + ih_s)\,dy_s - \frac{1}{2}\int_0^t (\phi_s^2 - h_s^2 + 2ih_s\phi_s)\,ds - \frac{1}{2}\int_0^t h_s^2\,ds + i\int_0^t f_s\phi_s\,ds\right]\middle|\,\phi_0^t\right)\\
&\quad = \mathbb{E}_0\left(\exp\left[\int_0^t (\phi_s + ih_s)\,dy_s - \frac{1}{2}\int_0^t (\phi_s + ih_s)^2\,ds\right]\middle|\,\phi_0^t\right)\cdot\exp\left[-\frac{1}{2}\int_0^t h_s^2\,ds + i\int_0^t f_s\phi_s\,ds\right]
\end{aligned}$$

כיוון ש- ϕ מדיד על עצמו ו- f, h דטרמיניסטים. נסמן

$$z_t = \exp\left[\int_0^t (\phi_s + ih_s)\,dy_s - \frac{1}{2}\int_0^t (\phi_s + ih_s)^2\,ds\right]\ . \tag{7.92}$$

אזי ממשפט איטו

$$z_t = 1 + \int_0^t z_s(\phi_s + ih_s)\,dy_s\ . \tag{7.93}$$

אך מכאן נובע כי $\mathbb{E}_0(z_t \mid \phi_0^t) = 1$ כי בגבול (כאשר $\epsilon \downarrow 0$), מהגדרת אינטגרל איטו,

$$\mathbb{E}_0(z_t \mid \phi_0^t) = 1 + \mathbb{E}_0\left[\sum z_{i\epsilon}(\phi_{i\epsilon} + ih_{i\epsilon})(y_{i\epsilon+\epsilon} - y_{i\epsilon}) \mid \phi_0^t\right]\ . \tag{7.94}$$

אולם

$$\mathbb{E}_0[\mathbb{E}_0(z_{i\epsilon} \mid \phi_0^t)(\phi_{i\epsilon} + ih_{i\epsilon})(y_{i\epsilon+\epsilon} - y_{i\epsilon}) \mid \phi_0^t y_0^{i\epsilon})] = 0 \tag{7.95}$$

כיוון שתחת $\mathbb{P}_0$ התהליך y_t הוא תנועה בראונית בת"ס ב- ϕ_0^t, ושאר התהליכים מדידים על ההתניה.
לכן $\mathbb{E}_0(z_t \mid \phi_0^t) = 1$ ולסיכום

$$\mathbb{E}_0[a_t c_t \mid \phi_0^t] = \exp\left[-\frac{1}{2}\int_0^t h_s^2\,ds + i\int_0^t \phi_s f_s\,ds\right]\ . \tag{7.96}$$

מצד שני

$$\mathbb{E}_0\,B_t = \mathbb{E}_0\exp\left[i\int_0^t f_s\phi_s\,ds + i\int_0^t h_s\,dy_s\right] \tag{7.97}$$

ועל ידי התניה ב- ϕ_0^t, כיוון ש- h_s דטרמיניסטי

$$\mathbb{E}_0\,B_t = \mathbb{E}_0\exp\left[i\int_0^t f_s\phi_s\,ds - i\int_0^t h_s^2\,ds\right] \tag{7.98}$$

כי y_s הוא תנועת בראון לפי $\mathbb{E}_0$. לסיכום

$$\mathbb{E}_1\, c_t = \mathbb{E}_0 \exp\left[-\frac{1}{2}\int_0^t h_s^2\, ds + i\int_0^t \phi_s f_s\, ds\right] = \mathbb{E}_0\, B_t \ . \tag{7.99}$$

8 נספח: מרחבים טופולוגיים, מטריים ומרחבי הילברט

בנספח זה כמה נושאי רקע מתחום האנליזה ואנליזה פונקציונלית.

8.1 טופולוגיה ומרחבים מטריים

מרחב, או קבוצה, הם שמות עבור אוסף של עצמים. אנו נקרא לאברי הקבוצה נקודות ונסמן את המרחב ב-S. דוגמה---הישר הממשי $\mathbb{R}$ הוא אוסף כל המספרים הממשיים.

הגדרה 8.1 עבור מרחב נתון, טפולוגיה $\mathbb{T}$ היא אוסף קבוצות המוכלות במרחב, אשר מקיימות את התכונות הבאות.

1. $S \in \mathbb{T},\ \emptyset \in \mathbb{T}$

2. $\mathbb{T}$ סגור תחת חיתוכים סופיים ואיחודים כלשהם. כלומר אם $\{O_\alpha,\ \alpha \in A\}$ הוא אוסף אברים ב-$\mathbb{T}$ אזי גם $O_{\alpha_1} \cap O_{\alpha_2} \in \mathbb{T}$ וכן $\cup_\alpha O_\alpha \in \mathbb{T}$.

מרחב טופולוגי הוא מרחב שעליו מוגדרת טופולוגיה.

דוגמה - אוסף הקבוצות הפתוחות על הישר הממשי היא טופולוגיה.

הגדרה 8.2 מרחב יקרא מטרי אם מוגדרת פונקציה ממשית d המוגדרת על זוגות ב- S ומקיימת:

1. $d(x,y) \geq 0$

2. $d(x,y) = 0$ אם ורק אם $x = y$

3. $d(x,y) = d(y,x)$

4. אי שוויון המשולש $d(x,y) \leq d(x,z) + d(z,y)$.

דוגמה - המרחק האוקלידי הוא מטריקה.

במרחב מטרי, כדור פתוח $B(x,\varepsilon)$ הוא אוסף הנקודות במרחב אשר מרחקן מהנקודה x קטן ממש מ- ε. קבוצה פתוחה היא קבוצה אשר סביב כל נקודה שלה יש כדור פתוח המוכל בקבוצה.

אוסף הקבוצות הפתוחות במרחב מטרי הוא טופולוגיה (בדוק!). המרחבים בהם נעסוק יהיו תמיד מרחבים מטריים, ונשתמש תמיד בטופולוגיה המוגדרת על ידי הקבוצות הפתוחות.

התכנסות: אם $x_n \in S$, $n = 1, 2, \ldots$, נאמר כי x_n מתכנס ל- x (כאשר $n \to \infty$) ונסמן זאת $x_n \to x$ אם $d(x_n, x) \to 0$.

הסדרה $\{x_n\}$ תקרא "סדרת קושי" (Cauchy) אם לכל ϵ קיים $N = N(\epsilon)$ כך ש- $d(x_m, x_n) < \epsilon$ לכל $m, n > N$. אם x_n מתכנסת אזי היא קושי, כי $d(x_n, x_m) \le d(x_n - x) + d(x - x_m)$.
מרחב מטרי נקרא שלם אם כל סדרת קושי מתכנסת.

8.2 משפט נקודת שבת

משפט נקודת השבת של בנך Banach Fixed Point Theorem ישמש אותנו להוכחת קיום פתרונות למשוואות סטוכסטיות.

הגדרה 8.3 יהי (S, d) מרחב מטרי (S הוא מרחב עליו מוגדרת מטריקה d). אופרטור T הוא פונקציה מ-S ל-S. אופרטור T יקרא "מכווץ" *"Contraction operator"* עם קבוע $q < 1$ אם הוא מקיים, לכל $x_1, x_2 \in S$ את אי השוויון

$$d(Tx_1, Tx_2) \le q \cdot d(x_1, x_2) \ . \tag{8.1}$$

לדוגמה הפונקציה $f(x) = qx + \alpha$ עם $|q| < 1$ היא אופרטור מכווץ על הישר הממשי.

נשים לב כי אופרטור מכווץ הוא בהכרח פונקציה רציפה, כלומר אם $y_n \to y$ אזי $Ty_n \to Ty$. זה נובע מכך שלפי ההגדרה $d(Ty_n, Ty) \le q \cdot d(y_n, y)$.

משפט 8.4 משפט נקודת השבת של בנך. אם (S, d) הוא מרחב מטרי שלם ו-T אופרטור מכווץ עם קבוע q אזי קיים פתרון יחיד למשוואה

$$Tx = x \ . \tag{8.2}$$

פתרון זה נקרא נקודת השבת של T. בנוסף, לכל x_0 מתקיים

$$T^n x_0 = T(T^{n-1} x_0) \to x \tag{8.3}$$

כלומר האיטרציות מתכנסות לנקודת השבת, וקצב ההתכנסות הוא גאומטרי במובן ש- $d(T^n x_0, x)$ יורד כמו q^n.

הוכחה: נתחיל בהוכחת יחידות. נניח ש-x_1, x_2 הן שתיהן נקודות שבת. אזי

$$d(x_1, x_2) = d(Tx_1, Tx_2) \le q \cdot d(x_1, x_2) \tag{8.4}$$

כאשר השוויון הראשון נובע מכך שהנקודות הן נקודות שבת, והשני מכך שהאופרטור מכווץ. כיוון ש-$q < 1$ אי השיוויון יתכן רק אם $d(x_1, x_2) = 0$ כלומר הנקודות זהות, והראנו יחידות.

קיום נקודת שבת וקצב התכנסות: נבחר $m > n$. מאי שוויון המשולש

$$d(T^n x_0, T^m x_0) \le d(T^n x_0, T^{n+1} x_0) + d(T^{n+1} x_0, T^m x_0) \tag{8.5}$$

$$\le \sum_{i=n}^{m-1} d(T^i x_0, T^{i+1} x_0) \tag{8.6}$$

$$\le \sum_{i=n}^{m-1} q^i d(x_0, Tx_0) \tag{8.7}$$

$$\le d(x_0, Tx_0) \sum_{i=n}^{\infty} q^i \tag{8.8}$$

$$= d(x_0, Tx_0) \frac{q^n}{1-q} \ . \tag{8.9}$$

הביטוי האחרון שואף ל-0 כאשר $n \to \infty$ ולכן הסדרה $T^n x_0$ היא סדרת קושי. כיוון שהמרחב שלם, הסדרה מתכנסת לגבול שנסמן ב-x. כלומר $d(T^n x_0, x) \to 0$. כעת, x הוא נקודת שבת כי

$$x = \lim_n T^n x_0 = \lim_n T(T^{n-1} x_0) = T \lim_n T^{n-1} x_0 = Tx$$

כאשר השוויון השלישי נובע מרציפות. קיבלנו שהגבול הוא אכן נקודת שבת. כעת, כיוון ש-x הוא נקודת שבת, $d(T^n x_0, x) = d(T^n x_0, T^n x) \le q^n \cdot d(x_0, x)$ וקצב ההתכנסות הוא כמובטח.

תרגיל 8.5 הוכח את ההרחבה הבאה של משפט נקודת השבת. אם קיים שלם חיובי n_0 כך ש-T^{n_0} הוא אופרטור כיווץ, אזי מסקנות המשפט בתוקף: ל- T קיימת נקודת שבת יחידה והאיטרציות מתכנסות בקצב גאומטרי.

תרגיל 8.6 דוגמאות של אופרטורי כיווץ:

א. הראה כי אם T הוא אופרטור מכווץ לינארי אזי נקודת השבת היא 0.

ב. מי מהאופרטורים הבאים (מהישר הממשי) הם מכווצים, ועבור אלו ערכים של הפרמטרים?

$$Tx = \alpha x + \beta \tag{8.10}$$

$$Tx = \sqrt{x} \tag{8.11}$$

$$Tx = x^2 \tag{8.12}$$

$$TX = \alpha \sin(x) \tag{8.13}$$

ג. תהי f פונקציה רציפה ליפשיץ עם קבוע L כלומר $|f(x) - f(y) \leq L|x - y|$. הוכח כי אם $\tau L < 1$ אזי קיים פתרון רציף יחיד על אינטרוול זמן $[0, \tau]$ למשוואה הדיפרנציאלית

$$\frac{dx}{dt} = f(x(t)) \tag{8.14}$$

עם תנאי התחלה נתון $x(0) = x_0$ וכן כי איטרציות פיקרד (להלן) מתכנסות לפתרון. רמז: הגדר אופרטור

$$Tx(t) = x_0 + \int_0^t f(x(s))\, ds \tag{8.15}$$

על מרחב הפונקציות הרציפות, עם מטריקה $\|y\| = \sup_{0\leq t\leq \tau} |y(t)|$. איטרציות פיקרד מוגדרות כך:

$$x^0(t) = x_0, \qquad x^{n+1}(t) = (Tx^n)(t)\ . \tag{8.16}$$

ראה הוכחת קיום ויחידות של משוואות סטוכסטיות.

8.3 מרחבי הילברט

יהי $H = \{h\}$ אוסף אלמנטים h שיקראו וקטורים עבורם מוגדר חיבור וכפל בסקלר. נניח:

1. **לכל $h_1 \in H,\ h_2 \in H$ הסכום $h_1 + h_2 \in H$ ולכל α ממשי $\alpha h_1 \in H$,**
2. **מתקיימות תכונות הלינאריות:**

$$\alpha\,(h_1 + h_2) = \alpha h_1 + \alpha h_2 \tag{8.17}$$

$$(\alpha + \beta)\, h = \alpha h + \beta h \tag{8.18}$$

$$\alpha\,(\beta h) = (\alpha\beta)\, h \tag{8.19}$$

$$1 \cdot h = h \tag{8.20}$$

מרחב בעל כל התכונות הללו נקרא מרחב לינארי (או מרחב וקטורי לינארי). במרחב כזה יש תמיד איבר שנסמן ב- 0 (הוא מוגדר על ידי $h-h$ עבור h שרירותי. בדוק כי $0 \cdot h = 0$ כאשר האפס הראשון הוא סקלר והשני איבר ב- H).

הגדרה 8.7 נורמה $\|\cdot\|$ היא פונקציה ממשית המקיימת את התכונות הבאות:

1. $\|h\| = 0$ אם ורק אם $h = 0$
2. $\|x+y\| \le \|x\| + \|y\|$
3. $\|\alpha x\| = |\alpha \; \|x\|$.

בדוק כי כל נורמה היא מטריקה.

<u>הגדרה:</u> מכפלה סקלרית ב- H היא פונקציה ממשית על הזוגות h_1, h_2 ב- H, שנסמן ב- $(h_1,\, h_2)$, והמקיימת

$$(h_1\,,\,h_2) = (h_2\,,\,h_1) \tag{8.21}$$

$$(h_1\,,\,h_2 + \alpha h_3) = (h_1\,,\,h_2) + \alpha\,(h_1\,,\,h_3) \tag{8.22}$$

$$(h\,,\,h) = 0 \text{ אם ורק אם } h = 0 \tag{8.23}$$

$$(h\,,\,h) \ge 0 \tag{8.24}$$

נסמן $||h|| = (h,h)^{1/2}$. שים לב כי זו אכן נורמה, וכי משתי התכונות האחרונות נובע באופן מיידי כי אם $(h,g)=0$ לכל g אזי בהכרח $h=0$.

משפט 8.8 אי שיוויון קושי שוורץ - *Cauchy Schwartz inequality*: $\quad (h_1,h_2)^2 \le ||h_1||^2||h_2||^2$

הוכחה: מהגדרה, $\|h_1 + \lambda h_2\|^2 \ge 0$ לכל λ ממשי.
מצא λ המביא ביטוי זה למינימום, הצב והתוצאה תתקבל.

משפט 8.9 אי שיוויון המשולש: $\quad \|h_1 + h_2\| \le \|h_1\| + \|h_2\|$

הוכחה:

$$
\begin{aligned}
\|h_1+h_2\|^2 &= ((h_1+h_2),(h_1+h_2)) && \text{מההגדרה} \\
&= ((h_1+h_2),h_1)+((h_1+h_2),h_2) && \text{מלינאריות} \\
&\le \|h_1+h_2\|\cdot\|h_1\|+\|h_1+h_2\|\cdot\|h_2\| && \text{מאי שיוויון שוורץ} \\
&= \|h_1+h_2\|\cdot(\|h_1\|+\|h_2\|)
\end{aligned}
$$

בגלל תכונות (8.21)--(8.24) ואי שיוויון המשולש, הגודל $\|\cdot\|$ המושרה על ידי המכפלה הסקלרית הוא נורמה - מדד של "אורך" על H.

הערה 8.10 לאוסף אלמנטים יכולה להיות יותר ממכפלה סקלרית אחת, ולכן גם יותר מנורמה אחת - ראה תרגיל למטה.

התכנסות: אם $h_n \in H$, $n=1,2,\ldots$, נאמר כי h_n מתכנס ל- h (ונסמן זאת $h_n \to h$) כאשר $n\to\infty$ אם $||h_n-h||\to 0$. זוהי "התכנסות חזקה". בצורה דומה $d(x_n,y)\to 0$ משמעו התכנסות במטריקה d לגבול y.

הסדרה $\{h_n\}$ תקרא "סדרת קושי" (Cauchy) אם לכל ϵ קיים $N=N(\epsilon)$ כך ש- $\|h_m-h_n\|<\epsilon$ לכל $m,n>N$. אם h_n מתכנסת אזי היא קושי, כי $\|h_n-h_m\|\le\|h_n-h\|+\|h-h_m\|$.
מרחב עם נורמה נקרא שלם אם כל סדרת קושי מתכנסת.

הגדרה 8.11 מרחב הילברט הוא מרחב ווקטורי לינארי עם מכפלה סקלרית שהוא שלם, כאשר הנורמה היא זו המוגדרת על ידי המכפלה הסקלרית. תת מרחב לינארי של מרחב הילברט שהוא שלם תחת אותה מכפלה סקלרית נקרא תת מרחב הילברט של המרחב המקורי.

תרגיל 8.12 יהי H המרחב האוקלידי התלת מימדי, כלומר אוסף הווקטורים $h=(x^1,x^2,x^3)$ כאשר x^i ממשיים. נגדיר $(h_1,h_2)=\sum_{i=1}^3 x_1^i x_2^i$. הוכח כי H היא מרחב הילברט.

כעת נגדיר נורמה חדשה $\|h\|_1=\sum_{i=1}^3|x^i|$. הוכח כי תחת נורמה זו, H הוא מרחב בנך, כלומר, מרחב וקטורי לינארי שהוא בעל נורמה ושלם תחת נורמה זו. הוכח כי במקרה זה המרחב H עם $\|\cdot\|_1$ איננו מרחב הילברט. רמז: הוכח כי לכל מרחב הילברט, ניתן לייצג:

$$(h_1,h_2)=(1/4)(\|h_1+h_2\|^2-\|h_1-h_2\|^2)$$

תרגיל 8.13 נגדיר $(h_1,h_2)_a=\sum_{i,j=1}^{3}a_{ij}x_1^ix_2^j$ כאשר $\{a_{ij}\}$ אברי מטריצה 3×3.
מתי זוהי מכפלה סקלרית?

הערה 8.14 במקרה שבתרגילים, אוסף הווקטורים מהצורה $(0,x^2,x^3)$ מהווים תת מרחב הילברט.
נקדים את המאוחר ונציין כי הסיבה העיקרית להתענינות שלנו במרחבי הילברט היא
<u>הדוגמה:</u> (ללא הוכחה) יהי אוסף המ"א $\{X\}$ כך ש $\mathbb{E}\,X^2<\infty$.
נגדיר $(X_1,X_2)=\mathbb{E}\,X_1X_2$. אזי ניתן להראות כי H מרחב הילברט עם מכפלה סקלרית $(\cdot,\cdot)$. במרחב
זה, אם $\mathbb{E}\,(X-Y)^2=0$ אזי $X=Y$!

יהי $\{X_i, i=1,2,\cdots,N\}$ אוסף אלמנטים ב-H. נגדיר:
$H_1=\{\mathbb{E}\,h^2<\infty$ האוסף h כך שקיימת פונקציה כלשהי f עבורה $h=f(X_1,\cdots,X_n)$ ובנוסף$\}$
אזי H_1 תת מרחב הילברט של H. זהו תת המרחב ה"נפרש" על ידי פונקציות של
$\{X_i, i=1,2,\cdots,N\}$. תתי מרחב כאלו ידונו בהמשך.

<u>דוגמה:</u> אם H הוא אוסף הפונקציות הממשיות על האינטרוול $[a,b]$ המקיימות

$$\int_a^b f^2(t)\mu(t)dt<\infty \tag{8.25}$$

נגדיר

$$(f,g)=\int_a^b f(t)g(t)\mu(t)dt, \tag{8.26}$$

כאשר $\mu(t)\geq 0$. בהנחות טכניות על הפונקציות והגדרה נכונה (לבג) של האינטגרל, מתקבל מרחב
הילברט. גם כאן, אם

$$\int_a^b \left(f_1(t)-f_2(t)\right)^2\mu(t)dt=0 \tag{8.27}$$

אזי f_1 זהה ל-f_2 כאלמנט ב-H. כך למשל, אם $\mu(t)=0$ עבור t בקטע $[c,d]$ המוכל ב-$[a,b]$, אזי אם
f_1 ו-f_2 שונים רק בקטע $[c,d]$ הם זהים כאלמנטים ב-H. בנוסף, אם הם שונים בנקודות מבודדות
$t_i, i=1,\cdots,M$ הם עדין זהים כאלמנטים ב-H.

המשפט החשוב ביותר מבחינתנו הוא

משפט 8.15 (משפט ההשלכה) יהי H מרחב הילברט, ו-H_1 תת מרחב הילברט של H. יהי $h \in H$ נתון, ונגדיר $d = \inf_{h_1 \in H_1}\{\|h - h_1\|\}$ אזי

א. קיים $\hat{h} \in H_1$ כך ש- $\|h - \hat{h}\| = d$.

ב. כל $\hat{h}$ כזה מקיים $\left(h - \hat{h}, h_1\right) = 0$ לכל $h_1 \in H_1$ ("ניצבות השגיאה").

ג. $\hat{h}$ המקיים את ב. הוא יחיד.

הוכחה:

א. נבחר סדרה $\{h_n\}$ **כך ש-** $h_n \in H_1$ **ו-** $\|h_n - h\| \rightarrow d$.

$$\begin{aligned}\|h_n - h_m\|^2 &= (h_n - h_m, h_n - h_m) \\ &= 2\left((h_n - h), (h_n - h)\right) + 2\left((h_m - h), (h_m - h)\right) \\ &\quad - \left((h_n + h_m - 2h), (h_n + h_m - 2h)\right) \\ &= 2\|h_n - h\|^2 + 2\|h_m - h\|^2 - 4\|h - \frac{1}{2}(h_n + h_m)\|^2\end{aligned}$$

כיוון ש- $\frac{1}{2}(h_n - h_m) \in H_1$, **מהגדרת** d, **מתקיים** $\|h - \frac{1}{2}(h_n + h_m)\| \geq d$, **ולכן**

$$\|h_n - h_m\|^2 \leq 2\|h_n - h\|^2 + 2\|h_m - h\|^2 - 4d^2$$

אבל צד ימין שואף ל-0 כאשר n **ו-**m **שואפים ל-**∞**. לכן** $\{h_n\}$ **סדרת קושי ב-** H_1**, ולכן מהשלמות קיים לה גבול ב-**H_1 **שנסמן ב-**$\hat{h}$.
כיוון ש-

$$\|h - \hat{h}\| \leq \|h - h_n\| + \|h_n - \hat{h}\|$$

נובע כי $\|\hat{h} - h\| = d$.

ב. נניח בדרך השלילה כי קיים $h_1 \in H_1$ **ו-** $\delta > 0$ **כך ש-** $\left(\left(h - \hat{h}\right), h_1\right) = \delta$ **עבור** $\hat{h}$ **המקיים**

$d = \|h - \hat{h}\|$. נגדיר $h_0 = h_1/\delta$. יהי $\epsilon = 1/\|h_0\|^2$. אזי

$$\begin{aligned}
\|h - \hat{h} - \epsilon h_0\|^2 &= \left(\left(h - \hat{h} - \epsilon h_0\right), \left(h - \hat{h} - \epsilon h_0\right)\right) \\
&= \|h - \hat{h}\|^2 + \epsilon^2 \|h_0\|^2 - 2\epsilon \\
&= d^2 + \frac{1}{\|h_0\|^2} - 2\frac{1}{\|h_0\|^2} \\
&< d^2
\end{aligned}$$

וזאת בסתירה להגדרה של d, כי $\hat{h} + \epsilon h_0 \in H_1$.
במקרה בו $\delta < 0$ החישוב זהה, כי אז נבחר $h_1' = -h_1$.

ג. נניח בשלילה כי $\hat{h}$ ו- $\tilde{h}$ ב-H_1 מקיימים שניהם את תכונת הניצבות. אזי $\left(h - \tilde{h}, \hat{h} - \tilde{h}\right) = 0$ כי $\hat{h} - \tilde{h} \in H_1$. לכן,

$$0 = \left(\left(h - \hat{h}\right) + \left(\hat{h} - \tilde{h}\right), \left(\hat{h} - \tilde{h}\right)\right) = \left(h - \hat{h}, \hat{h} - \tilde{h}\right) + \|\hat{h} - \tilde{h}\|^2 = 0 + \|\hat{h} - \tilde{h}\|^2$$

ולכן $\hat{h} = \tilde{h}$.

הערה 8.16 משפט ההשלכה תקף גם אם H_1 היא קבוצה קעורה וסגורה - ההוכחה זהה.

הגדרה 8.17 פונקציונל הוא פונקציה ממשית על H. פונקציונל יקרא לינארי אם $f(h_1 + \alpha h_2) = f(h_1) + \alpha f(h_2)$ ויקרא חסום אם קיים $K < \infty$ כך שלכל $h \in H$, $|f(h)| < K\|h\|$.

הערה 8.18 המרחב $\mathbb{R}$ הוא מרחב הילברט, ולכן פונקציה ממשית של $\mathbb{R}$ היא פונקציונל. הפונקציונל $g(x) = 2x$ הוא לינארי וחסום על $\mathbb{R}$, לפי ההגדרה לעיל, אך אין לבלבל זאת עם הגדרת <u>פונקציה</u> חסומה – $g(x)$ איננה פונקציה חסומה!

<u>דוגמה:</u> נבחר $h_0 \in H$ ונגדיר $f(h) = (h_0, h)$. אזי f לינארי (בדוק!) וחסום כי

$$|f(h)| = |(h_1, h_0)| \leq \|h\|\|h_0\| \qquad (8.28)$$

משפט 8.19 (משפט ריץ - $Riesz$) אם f פונקציונל לינארי חסום ב-H אזי קיים $h_0 \in H$ יחיד כך שלכל $h \in H$ מתקיים $f(h) = (h, h_0)$ כלומר, ניתן לייצג את f בעזרת h_0, ע"י מכפלה סקלרית.

הגדרה 8.20 מרחב הילברט H יקרא ספרבילי אם קיימת משפחה אורתונורמלית שלמה, כלומר קיימת סידרה $\{h_m, m = 1, 2, \dots\}$ שהיא

אורתונורמלית – $(h_i, h_j) = \delta_{ij}$

שלמה – $h \in H$ ו- $(h_i, h) = 0$ לכל i גורר $h = 0$

או, בצורה שקולה:

לכל $h \in H$ קיים פיתוח פוריה $h = \sum_{i=1}^{\infty}(h_i, h)h_i$, כלומר $\lim_{N\to\infty} \|h - \sum_{i=1}^{N}(h_i, h)h_i\| = 0$.

<u>דוגמה:</u> **יהי $L^2(0,1)$ אוסף הפונקציות הממשיות על $(0,1)$ המקיימות $\int_0^1 f^2(t)dt < \infty$. אזי קיים** משפט פוריה, ולכן המרחב $L^2(0,1)$ **ספרבילי.**