Remove Watermark No

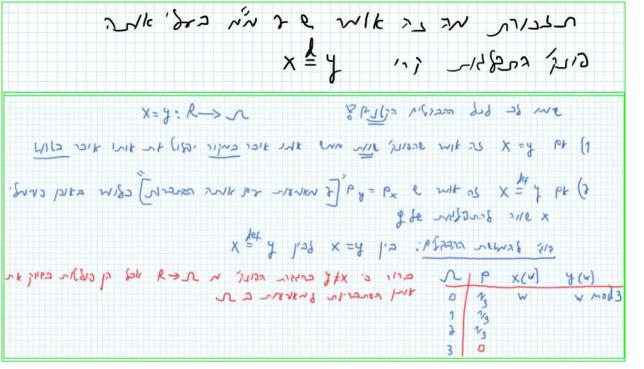
יסודות תורת ההסתברות

אוהד נ. פלדהיים אורי גוראל גורביץ'

מכון איינשטיין למתמטיקה

האוניברסיטה העברית בירושלים © כל הזכויות שמורות לכותבים

2019 בינואר 3



פרק 10

סדרות של התפלגויות

כמעט שאינני מכיר דבר-מה אשר מותיר רושם כה עז על הדמיון כמו תבנית הסדר הקוסמי המתבטאת באמצעות חוק שכיחות הסטיות. החוק היה זוכה להאנשה על ידי היוונים אילו הכירוהו. הוא מושל בשלווה ובצניעות מושלמת בלב התוהו הפראי ביותר. ככל שרב האספסוף, ניכרת יותר האנארכיה, כך נשעית שליטתו מושלמת. זהו החוק העליון שבשיגעון.

– סר פרנסיס גלטון על משפט הגבול המרכזי, "תורשה טבעית", 1889

במובן מסויים של התכנסות משתנים מקריים נתקלנו בפרק 3, בדוגמא 3.42. שם רצינו להראות שהתפלגות מספר המטבעות, מחוטלים לפני תוצאה ראשונה של עץ (או סיום הטלת כל המטבעות), כאשר מטילים N מטבעות, מתכנסת להתפלגותו של משתנה גיאומטרי. קשרים דומים שימשו אותנו לניסוח הקשר בין התפלגות בינומית לפואסונית (בעיה 3.27) ולניסוח החוק החלש של המספרים הגדולים (משפט 5.8). בפרק זה נגדיר מושג זה של התכנסות המכונה התכנסות התפלגויות או התכנסות בהתפלגות ונחקור תכונות אוניברסליות שלו.

10.1 התכנסות התפלגויות

את מושג ההתכנסות עבור התפלגויות נגדיר במונחים של פונקציות התפלגות מצטברות.

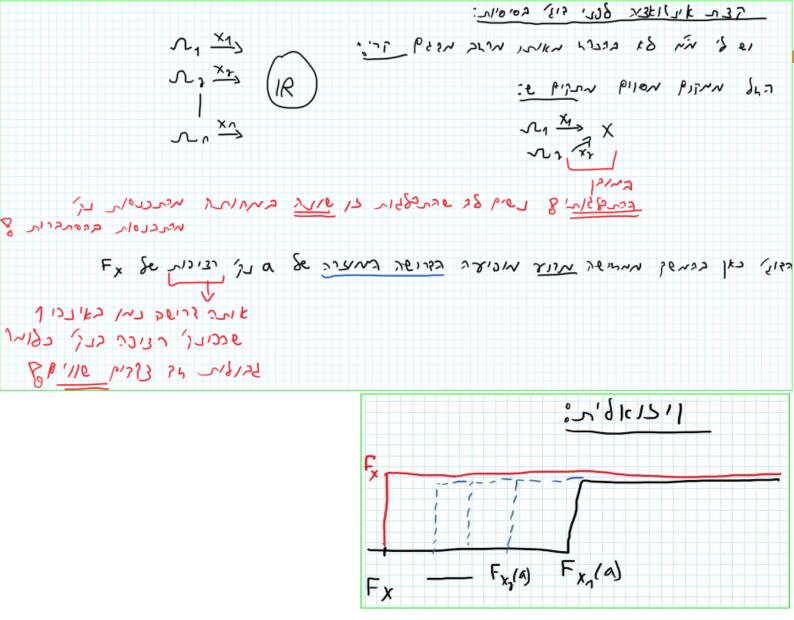
הגדרה 10.1 (התכנסות בהתפלגות למשתנה מקרי בדיד). תהי $(X_n)_{n\in\mathbb{N}}$ סדרת משתנים מקריים, לא בהכרח (in distribution) על אותו מרחב הסתברות, ויהי X משתנה מקרי. נאמר כי $(X_n)_{n\in\mathbb{N}}$ מתכנסת בהתפלגות משתנה X אם לכל X_n שהיא נקודת רציפות של X_n מתקיים X_n אם לכל X_n שהיא נקודת רציפות של X_n מתקיים

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

ב- ממצאת ב-I נמצאת ב-I לכל קטע ב-I התכנסות בהתפלגות היא תכונה חלשה יותר מ-I המורכבת המוטיבציה להגדרה המורכבת והחלשה יותר, נסתכל על הדוגמא הבאה.

דוגמא 10.2 תהי $\stackrel{\mathrm{a.s.}}{=} c_n$ סדרת משתנים מקריים קבועים, כאשר $c_n o 0$ ויהי $X_n \stackrel{\mathrm{a.s.}}{=} c_n$ פונקציית ההתפלגות

$$f_{x_n}(a) = 11(a \ge \frac{1}{n}) \xrightarrow{n-3n-3} f_{x}(a) = 11(a \ge 0)$$



פרק 10 160

המצטברות של הסדרה הן

$$F_{X_n}(a) = \begin{cases} 0 & a < c_n \\ 1 & a \ge c_n \end{cases}$$

ואומנם $\lim_{n \to \infty} F_{X_n}(a) = F_X(a)$ בכל נקודה מלבד 0, ניתן לומר בכל נקודה של $\lim_{n \to \infty} F_{X_n}(a) = F_X(a)$ שההגדרה המורכבת שבחרנו נועדה לאפשר גם להתכנסות כזו להיקרא התכנסות בהתפלגות, ואכן לפי הגדרה $X_n \stackrel{\mathrm{d}}{\to} X \mathfrak{N}$

דוגמא 10.3 (דוגמאות פשוטות להתכנסות בהתפלגות).

- $X_n \stackrel{\mathrm{d}}{\to} X$ אם X_n כולם משתנים מקריים שווי-התפלגות, אז (א
- $c_n o c$ אם ורק אם $X_n\stackrel{ ext{d}}{ o} X$ אז $\mathbb{P}(X=c)=1$ ר- ו $\mathbb{P}(X_n=c_n)=1$ אם (ב)
- נשים לב . $X_n \sim U([\frac{1}{n},\frac{2}{n}])$ ו $X_n \sim U([0,\frac{1}{n}])$ כך גם לגבי . $X_n \stackrel{\mathsf{d}}{\to} 0$ אז $X_n \sim U([-\frac{1}{n},\frac{1}{n}])$ גו אם (ג) שבדוגמאות אלה הערכים של $F_{X_n}(0)$ שונים ולא חייבים להתכנס ל- $F_{X_n}(0)$. כיוון ש-0 היא נקודת אי-רציפות של F_X הדבר עולה בקנה אחד עם הגדרת התכנסות בהתפלגות.

כעת נראה כיצד ההתפלגויות הרציפות בהחלט שלמדנו להכיר בפרק 9, מתקבלות כגבול של התפלגויות



- דוגמא 10.4. a<0 או a>1 או a>1
 - : נחשב את הגבול אר פל $k\in\mathbb{N}_0$ לכל $X\sim\operatorname{Po}(1)$, $X_n\sim\operatorname{Bin}(n,\frac{1}{n})$

$$\mathbb{P}(X_n = k) = \binom{n}{k} \frac{1}{n^k} \left(1 - \frac{1}{n} \right)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{1}{k!} \left(1 - \frac{1}{n} \right)^{n-k} \overset{n \to \infty}{\to} \frac{e^{-1}}{k!} = \mathbb{P}(X = k).$$

עבור k טבעי, אז $a\in(k,k+1)$ אם נקח . $\mathbb{R}\setminus\mathbb{N}$ היא אר F_X טבעי, אז

$$F_{X_n}(a) = \sum_{j=0}^k \mathbb{P}(X_n = j) \overset{n \to \infty}{ o} \sum_{j=0}^k \mathbb{P}(X = j) = F_X(a),$$

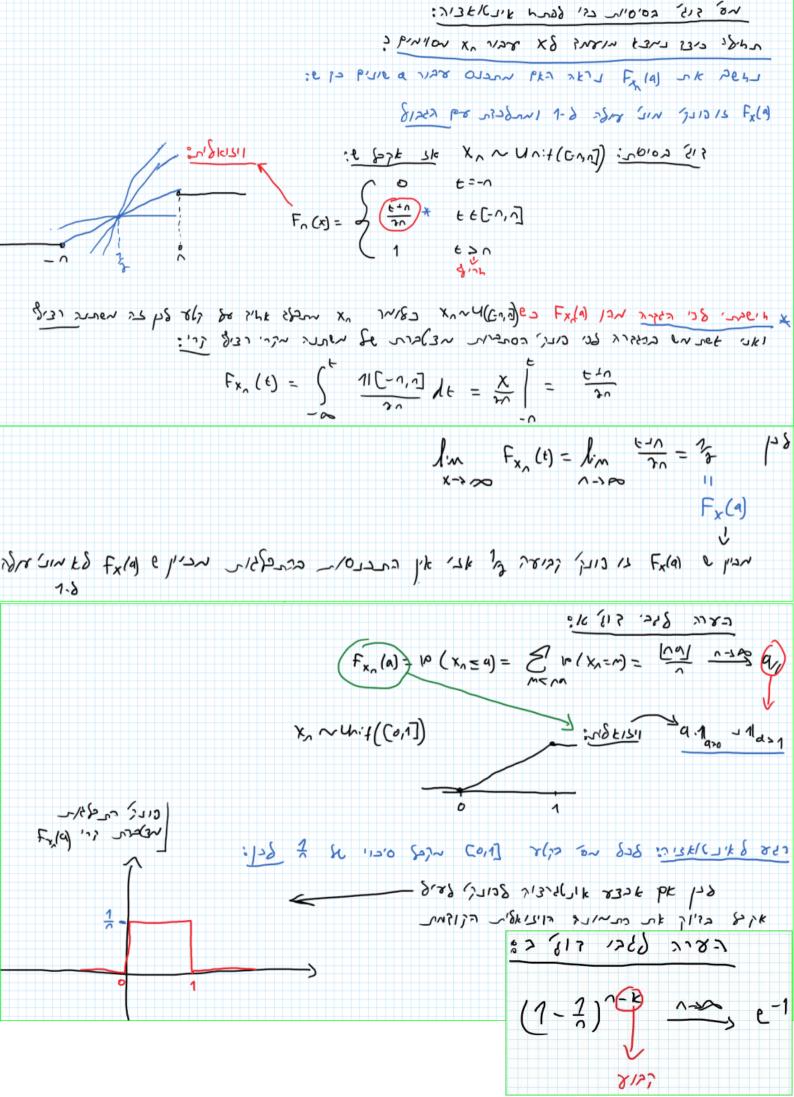
 F_X אם בנקודות האי-רציפות של $F_{X_n}(a) o F_X(a)$ או בדוגמה זו ולכן $X_n \overset{ ext{d}}{ o} X$

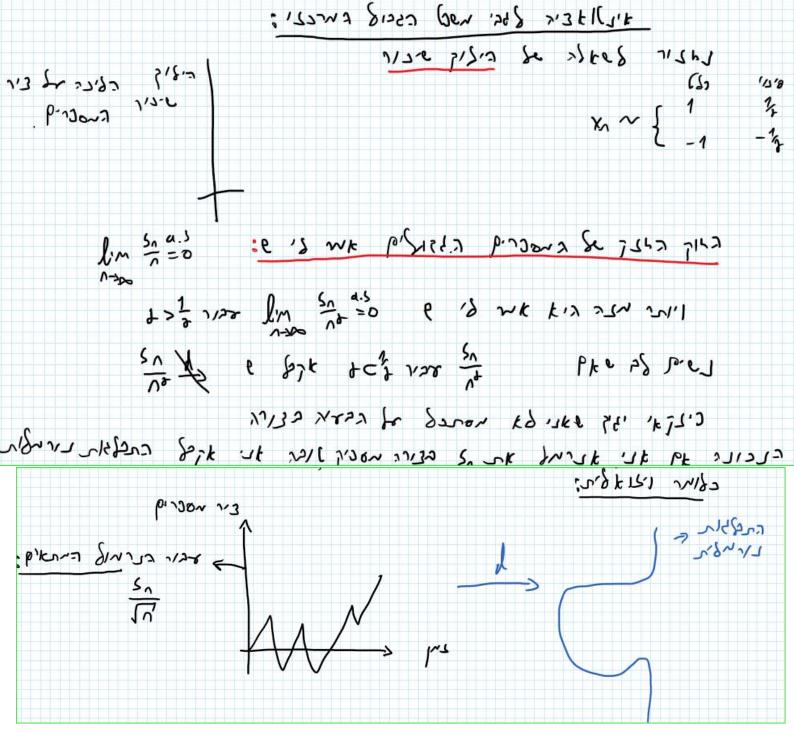
a<0 או a>1 עבור F_X עבור רציפות אינ נקודת מל $X\sim U([0,1])$ עבור $X_n\simeq U([\frac{1}{n},\frac{2}{n},\dots,1])$ (א) או $\mathbb{P}(X_n\leq a)=\frac{\lfloor na\rfloor}{n}\overset{n\to\infty}{\to}a=\mathbb{P}(X\leq a)$ מתקיים 0< a<1 ולכן יש

$$.a>0$$
 נקח $.Y_n:=rac{1}{n}X_n\stackrel{ ext{d}}{ o}X$ אזי איז $X\sim \operatorname{Exp}(1)$ נקח $X_n\sim \operatorname{Geo}(rac{1}{n})$ אם (ד

$$\mathbb{P}(Y_n \le a) = \sum_{k=1}^{\lfloor na \rfloor} \left(1 - \frac{1}{n}\right)^k \frac{1}{n} = \frac{1 - \left(1 - \frac{1}{n}\right)^{na+1}}{1 - \left(1 - \frac{1}{n}\right)} \frac{1}{n} \xrightarrow{n \to \infty} 1 - e^{-a} = F_X(a).$$

נפקדת יחידה מרשימת דוגמאות זו היא ההתפלגות הנורמלית. ההתכנסות להתפלגות זו מבוארת בפרק הבא.





סדרות של התפלגויות

10.2 משפט הגבול המרכזי

אחת התוצאות המרשימות ביותר של תורת ההסתברות, אשר תהילתה מובאת בציטוט שבתחילת הפרק, היא משפט הגבול המרכזי. משפט זה קובע כי כאשר מספר גדול של גורמים בלתי-תלויים בעלי שונות נסכמים, אז לאחר נרמול בסטיית התקן של הסכום - סטייתו מהתוחלת שואפת תמיד להתפלגות נורמלית. זו דוגמא לתופעה של אוניברסליות, כלומר מצב שבו תנאי ההתחלת המתמטיים אינם משפיעים על התוצאה. תוצאות כאלו זוכות לעיתים קרובות לשימושים רבים בכל תחומי המדע.

גרסא בסיסית של המשפט הייתה ידועה כבר לאברהם דה-מואבר ב-1733, והוא הורחב לתנאים מקלים יותר ויותר, כאשר גרסאותיו המודרניות מאפשרות גם סכימה של משתנים בעלי תלויות ובלבד שהללו תהיינה חלשות דיו. כאן נביא את גרסתו הקלאסית והמוכרת ביותר של המשפט.

, משפט הגבול המרכזיי). תהי אויי הדרת משתנים מקריים בלתי-תלויים ושווי התפלגות משפט 10.5 (משפט הגבול המרכזיי). תהי

בעלי תוחלת 0 ושונות 1. אזי

 $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \stackrel{\mathrm{d}}{\to} Z$

כאשר Z הוא משתנה מקרי נורמלי סטנדרטי.

מתקיים $a\in\mathbb{R}$ מתקיים

$$\mathbb{P}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_{i} \leq a\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}e^{-x^{2}/2}dx.$$

(Carl-Gustav וקרל-גוסטב אסין (Andrew Berry) גרסא כמותית של המשפט הוכחה על ידי אנדרו ברי .Esseen

משפט ברי-אסין). תהי תהפלגות, סדרת משתנים מקריים בלתי-תלויים ושווי התפלגות, בעלי (משפט ברי-אסין). תהי תהפלגות, סדרת משתנים מקריים מקריים $a\in\mathbb{R}$ אזי לכל $ho=\mathbb{E}\left[\left|X_1^3\right|\right]<\infty$ מתקיים

$$\left| \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \le a \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx \right| \le \frac{\rho}{\sqrt{n}}.$$

 $rac{X_n-pn}{\sqrt{p(1-p)n}}\stackrel{ ext{d}}{
ightarrow}$ כדרת משתנים מקריים בינומים ויהי $X\sim N(0,1)$. נראה כי $X_n\sim ext{Bin}(n,p)$. נראה כי $X_n\sim ext{Bin}(n,p)$

נשים לב כי . $Y_{\widehat{\mathbf{I}}} \sim \mathrm{Ber}(p)$ הייו כך שמתקיים בלתי-תלויים בלתי-תלויים משתנים בלתי-תלויים מתקיים אווי התפלגות ושווי בלתי-תלויים ושווי התפלגות כך שמתקיים בלתי-תלויים בלתי-תליים בלתי-תלויים ב

$$\frac{X_n - pn}{\sqrt{p(1-p)n}} = \sum_{i=1}^n \frac{Y_i - \mathbb{E}(Y_i)}{\sqrt{\operatorname{Var}(Y_i)}}$$

הוין באן מאג אין לעיתל במו שהפיציאי יבר הלמנת BEDRYS ENDED ESCEL ENCES 3 1 = ~ VIL 2 8)= var(y:)-0 Var(y:) = 1,1 Var (2;) = Var 6 = TVOT ; 1201 Var (2) = var x = Vurl y:)
Var (7:) 80,514 019 ching To 829/1 Alut KLIEILZ

162 פרק

נסמן $Z_i = \frac{Y_i - \mathbb{E}(Y_i)}{\sqrt{\mathrm{Var}(Y_i)}}$ נסמן משפט הגבול משפט בלתי-תלויים בעלי תוחלת $Z_i = \frac{Y_i - \mathbb{E}(Y_i)}{\sqrt{\mathrm{Var}(Y_i)}}$

$$\frac{X_n - pn}{\sqrt{p(1-p)n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \stackrel{d}{\to} X,$$

כנדרש.

10.3 התכנסות בהתפלגות והתכנסות בהסתברות

X ייהי משתנים מקריים מדרת התכנסות ההתכנסות החכנסות החכנסות ווררת התכנסות בהסתברות ווררת התכנסות בהתפלגות). אז אז אז $X_n \stackrel{\mathrm{d}}{\to} X$ משתנה מקרי, כולם על אותו מרחב הסתברות. אם $X_n \stackrel{\mathrm{d}}{\to} X$

 $\epsilon > 0$ ויהי , F_X ויהי מוכחה. תהי a נקודת רציפות של

$$\mathbb{P}(X_n \le a) = \mathbb{P}(X_n \le a, X \le a + \epsilon) + \mathbb{P}(X_n \le a, X > a + \epsilon)$$

$$\leq \mathbb{P}(X \le a + \epsilon) + \mathbb{P}(|X - X_n| > \epsilon)$$

(1)

דומה דומה . $F_{X_n}(a) \leq F_X(a+\epsilon) + \mathbb{P}(|X-X_n| > \epsilon)$ ולכן

 $\mathbb{P}(X \le a - \epsilon) = \mathbb{P}(X \le a - \epsilon, X_n \le a) + \mathbb{P}(X \le a - \epsilon, X_n > a) \le \mathbb{P}(X_n \le a) + \mathbb{P}(|X - X_n| > \epsilon)$

ולכן
$$F_X(a-\epsilon) \leq F_{X_n}(a) + \mathbb{P}(|X-X_n| > \epsilon)$$
 ולכן

$$F_X(a-\epsilon) - \mathbb{P}(|X-X_n| > \epsilon) \le F_{X_n}(a) \le F_X(a+\epsilon) + \mathbb{P}(|X-X_n| > \epsilon)$$

נשאיף את לכל לכל מהגדרת ההתכנסות מהגדרת מהגדרת לעלכל לכל מהגדרת לאינסוף נקבל מהגדרת לאינסוף מחלבים לכל מהגדרת לאינסוף ונקבל מהגדרת ההתכנסות בהסתברות כי לכל ל

$$F_X(a-\epsilon) \leq \liminf_{n\to\infty} F_{X_n}(a) \leq \limsup_{n\to\infty} F_{X_n}(a) \leq F_X(a+\epsilon).$$

 $.F_X(a) \leq \lim_{n o \infty} F_{X_n}(a) \leq F_X(a)$ ומכך שa נקודת רציפות נקבל

ההיפך אינו נכון. למשל, אם $(X_n)_{n\in\mathbb{N}}$ הם משתנים מקריים בלתי תלויים ושווי התפלגות שאינם קבועים, אז $X_n \stackrel{p}{ o} X$ כאשר X משתנה מקרי נוסף בעל אותה התפלגות, אך $X \stackrel{p}{ o} X$. עם זאת, מתקיים המשפט הבא.

משפט 10.9 (התכנסות בהתפלגות לקבוע גוררת התכנסות בהסתברות). תהי א סדרת משתנים מקריים, מקריים מקריים $X_n \stackrel{\mathrm{p}}{\to} c$ אז אז $C \in \mathbb{R}$ אז משתנה מקרי ויהי X

הוכחה. תהי c>0 פונקצית ההתפלגות המצטברת של המשתנה המקרי הקבוע לכל F_c מתקיים הוכחה. תהי $F_{X_n}(c+\epsilon)\to 1$ ולכן אם אם $F_c(c+\epsilon)\to 1$ ולכן אם $F_{X_n}(c+\epsilon)\to 1$ ולכן המקריים איז פונקצית ההתפלגות המצטברת איז פונקצית החתפלגות המצטברת המצטברת המצטברת המקריים המקריים פונקצית החתפלגות המצטברת המצטברת המצטברת המצטברת המצטברת המצטברת המצטברת המצטברת המקריים פונקצית החתפלגות המצטברת המ

$$\mathbb{P}(|X_n - c| \le \epsilon) \ge F_{X_n}(c + \epsilon) - F_{X_n}(c - \epsilon) \to 1,$$

בנדרש.

$$P(X_n \leq a, X = a \leq t) \leq P(X \leq a \leq t) = \frac{1}{2} = \frac{1}{$$

סדרות של התפלגויות

*10.4 משמעותה של התכנסות התפלגויות

הדרישה בהגדרה של התכנסות בהתפלגות של סדרת משתנים מתייחסת רק להתכנסות של בנקודות רציפות. ראינו בדוגמא 10.2 שזוהי דרישה משמעותית. ראשית נציג הגדרה חילופית להתכנסות בהתפלגות שאינה מערבת סוגיות רציפות של F_{X_n} .

 $\epsilon>0$ טענה 10.10. תהי תהי אם ורק אם לכל $(X_n)_{n\in\mathbb{N}}$ מתכנסת בהתפלגות ל-X אם ורק אם לכל ולכל T מתקיים:

$$F_X(r-\epsilon) \le \liminf_{n \to \infty} F_{X_n}(r) \le \limsup_{n \to \infty} F_{X_n}(r) \le F_X(r+\epsilon)$$

הוכחה. קל לראות כי בנקודות רציפות של F_X הטענה שקולה לכך ש- $F_X(r)=F_X(r)$, ולכן התנאי הוכחה. קל לראות כי בנקודות רציפות של F_X הטענה 10.10 חזק יותר מאשר התנאי בהגדרה 10.1. בכיוון ההפוך, עלינו להראות כי אם F_X , אז בכל נקודת אי-רציפות T_X מתקיים תנאי הטענה. נשים לב כי $T_X(r)$ הנה פונקציה מונוטונית עולה ולכן יש לה מספר בן-מניה של נקודות אי-רציפות (ר' טענה 0.4). יהי $T_X(r)=0$ שהנן נקודות רציפות של $T_X(r)=0$ לכן מתקיים לפי תכונת קיימות נקודות $T_X(r)=0$ ההתכנסות בהתפלגות

$$F_X(b-\epsilon) \leq F_X(b_-) = \lim_{n o \infty} F_{X_n}(b_-) \leq \liminf_{n o \infty} F_{X_n}(b) \leq \limsup_{n o \infty} F_{X_n}(b) \leq F_X(b+\epsilon),$$
כנדרש.

$$\mathbb{P}(X \in (a + \epsilon, b - \epsilon)) \le \mathbb{P}(X_n \in (a, b)) \le \mathbb{P}(X \in (a - \epsilon, b + \epsilon)).$$

נציג ללא הוכחה את הטענה הבאה המסכמת את מהותה של התכנסות בהתפלגות.

טענה 10.11. תהי $X_n \stackrel{ ext{d}}{\to} X$ אם ורק אם לכל מקריים ויהי אוי משתנים מקריים ורק אם ורק אם לכל $(X_n)_{n \in \mathbb{N}}$ אם ורק אם לכל $f: \mathbb{R} \to \mathbb{R}$

שבעיה 10.2. להוכיח את טענה 10.10 עבור משתנים מקריים רציפים בהחלט. כ

164

בעיות הרחבה והעשרה

 $Z_n\stackrel{
m d}{ o}Z$ יש להראות כי $Z_n=Y_n-\lfloor Y_n
floor$ וכן $Y_n=\sum_{i=1}^nX_i$ ונסמן ונסמן $X_n\sim {
m Unif}[0,1]$ יש להראות כי $X_n\sim {
m Unif}[0,1]$ עבור $Z\sim {
m Unif}[0,1]$

בעיה 10.4. לחשב את הגבול 🛇

$$\lim_{n\to\infty}e^{-n}\sum_{i=0}^n\frac{n^i}{i!}.$$

רמז: לכל n הביטוי הוא הסתברותו של מאורע הקשור למשתנה מקרי מסוים. ניתן להשתמש במשפט הגבול המרכזי בכדי לחשב את הגבול.

