21. לדגום עם או בלי החזרות? (Should you sample with or without replacement?)[D]

בכד A נמצאים 2 כדורים אדומים וכדור ירוק אחד, ובכד B נמצאים 101 כדורים אדומים ו-100 כדורים ירוקים. בחר כד אחד בצורה אקראי ושלוף שני כדורים באופן אקראי מהכד שנבחר. אתה מנצח אם אתה מזהה נכון אם כדורים נשלפו מכד A או כד B.

חשב את ההסתברויות לניצחון בכל אחד מהחוקים שלהן וקבע איזו שיטה נותן את ההסתברות הגבוהה ביותר לניצחון.

שאלה 1: אתה מחזיר את הכדור הראשון לפני השליפה השנייה.

שאלה 2: אתה לא מחזיר את הכדור הראשון לפני השליפה השנייה.

שאלה 3: לאחר שליפת הכדור הראשון אתה יכול להחליט אם להחזיר אותו או לא.

רמז: כאשר מחשבים הסתברויות:

$$\frac{101}{201} \approx \frac{100}{201} \approx \frac{100}{200} \approx \frac{1}{2}\,.$$

פתרון

יש ארבע תוצאות שנסמן RR, RG, GR, GG. לכל אחד מהחוקים נחשב את ההסתברויות המתנות של ארבעת התוצאות כתלות בזהות הכד A או B שנבחר תחילה. אחר כך נכפיל את ההסתברויות ב-$1/2$ כדי לקחת בחשבון את הבחירה האקראית של הכד.

תשובה 1: שליפה עם החזרה:

$$\begin{array}{rcccc}
P(RR|A) & = & \frac{2}{3}\cdot\frac{2}{3} & = & \frac{4}{9}\\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}\\
\hline
P(RG|A) & = & \frac{2}{3}\cdot\frac{1}{3} & = & \frac{2}{9}\\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}\\
\hline
P(GR|A) & = & \frac{1}{3}\cdot\frac{2}{3} & = & \frac{2}{9}\\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}\\
\hline
P(GG|A) & = & \frac{1}{3}\cdot\frac{1}{3} & = & \frac{1}{9}\\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

אם התוצאה היא RR ההסתברות שכד A נבחר $(4/9)$ גבוהה מהסתברות שכד B נבחר $(1/4)$; אחרת, ההסתברות ש-B נבחר גבוהה יותר. לכן:

$$P(\text{ניצחון}) = \frac{1}{2}\left(\frac{4}{9}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}\right) = \frac{43}{72} \approx 0.5972\,.$$

תשובה 2: שליפה ללא החזרה:

$$\begin{array}{rcccl}
P(RR|A) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A) & = & \frac{1}{3}\cdot 1 & = & \frac{1}{3} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A) & = & \frac{1}{3}\cdot 0 & = & 0 \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

אם התוצאה היא GG ההסתברות שכד B נבחר גבוהה יותר (כמובן!) מההסתברות שכד A נבחר; אחרת, ההסתברות שכד A נבחר גבוהה יותר. לכן:

$$P(\text{ניצחון}) = \frac{1}{2}\left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{4}\right) = \frac{5}{8} = 0.6250\,,$$

שהיא גבוהה יותר מההסתברות לניצחון כאשר שליפה היא עם החזרה.

תשובה 3: ההחלטה נעשית על סמך התוצאות של השליפה הראשונה.

אם השליפה הראשונה היא מכד A ההסתברויות חייבות להיות מותנות גם בהחלטה להחזיר או לא. אם השליפה הראשונה היא מכד B היא לא משפיעה על ההסתברויות בגלל הקירוב ברמז.

$$\begin{array}{lcccl}
P(RR|A,w) & = & \frac{2}{3}\cdot\frac{2}{3} & = & \frac{4}{9} \\
P(RR|A,w/o) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A,w) & = & \frac{2}{3}\cdot\frac{1}{3} & = & \frac{2}{9} \\
P(RG|A,w/o) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A,w) & = & \frac{1}{3}\cdot\frac{2}{3} & = & \frac{2}{9} \\
P(GR|A,w/o) & = & \frac{1}{3}\cdot 1 & = & \frac{1}{3} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A,w) & = & \frac{1}{3}\cdot\frac{1}{3} & = & \frac{1}{9} \\
P(GG|A,w/o) & = & \frac{1}{3}\cdot 0 & = & 0 \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

אם כדור אדום נשלף ראשונה אזי $\frac{4}{9} > \frac{1}{4}$ ו-$\frac{2}{9} < \frac{1}{4}$ עם החזרה, לעומת $\frac{1}{3} > \frac{1}{4}$ ו-$\frac{1}{3} > \frac{1}{4}$ ללא החזרה, לכן הכדור השני יכול לעזור בזיהוי הכד רק אם השליפה נעשית **עם** החזרה: אם אדום כד A ואם ירוק כד B.
נבחר שליפה עם החזרה:

$$P(\text{ניצחון אם אדום ראשון}) = \frac{1}{2}\left(\frac{4}{9} + \frac{1}{4}\right) = \frac{25}{72}\,.$$

אם כדור ירוק נשלף ראשון אזי $\frac{2}{9} < \frac{1}{4}$ ו-$\frac{1}{9} < \frac{1}{4}$ עם החזרה, לעומת $\frac{1}{3} > \frac{1}{4}$ ו-$0 < \frac{1}{4}$ ללא החזרה, לכן הכדור השני יכול לעזור בזיהוי הכד רק אם השליפה נעשית **בלי** החזרה:

$$P(\text{ניצחון אם ירוק ראשון}) = \frac{1}{2}\left(\frac{1}{3} + \frac{1}{4}\right) = \frac{7}{24}\,.$$

ההסתברות לניצחון היא:

$$P(\text{ניצחון}) = \frac{25}{72} + \frac{7}{24} = \frac{23}{36} \approx 0.6389\,.$$

ההסתברות הגבוהה ביותר לניצחון מתקבלת כאשר ההחלטה לשלוף עם או בלי החזרה מתקבלת בהתאם לתוצאה של השליפה הראשונה.

סימולציה

```
With replacement:
Expectation of winning = 0.5972
Average wins           = 0.5976
Without replacement:
Expectation of winning = 0.6250
Average wins           = 0.6207
Decide after first draw:
Expectation of winning = 0.6389
Average wins           = 0.6379
```

22. הקלפי (The ballot box)

שני מועמדים A ו-B מתמודדים בבחירות. A קיבל a קולות ו-B קיבל b קולות, $a > b$. הקולות נספרים אחד-אחד וסכומי הביניים $(a_i, b_i), 1 \leq i \leq a+b$ מתעדכנים לאחר ספירת כל קול. מה ההסתברות שעבור לפחות i אחד, $a_i = b_i$?

שאלה 1: פתור עבור $a = 3, b = 2$ על ידי הכנת רשימה (a_i, b_i) עבור $1 \leq i \leq 5$.

שאלה 2: פתור לכל $a > b$.

רמז 1: מה ניתן להגיד על זהות המועמד המוביל עד לתיקו **הראשון**?

Hint 2: מה החשיבות של הקול הראשון שנספר.

פתרון

תשובה 1: מספר הדרכים לסדר את סכומי הביניים הוא $10 = \binom{5}{3} = \binom{5}{2}$ כי המיקום הקולות עבור מועמד אחד קובע את מיקום הקולות של המועמד השני. בטבלה שלהלן רשומים הסידורים האפשריים של הקולות וסכומי הביניים כאשר התיקו הראשון מודגש:

A	A	A	B	B	$(1,0)$	$(2,0)$	$(3,0)$	$(3,1)$	$(3,2)$
A	A	B	A	B	$(1,0)$	$(2,0)$	$(2,1)$	$(3,1)$	$(3,2)$
A	B	A	A	B	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(3,1)$	$(3,2)$
B	A	A	A	B	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(3,1)$	$(3,2)$
A	A	B	B	A	$(1,0)$	$(2,0)$	$(2,1)$	$(\mathbf{2},\mathbf{2})$	$(3,2)$
A	B	A	B	A	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(2,2)$	$(3,2)$
B	A	A	B	A	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(2,2)$	$(3,2)$
A	B	B	A	A	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(1,2)$	$(2,2)$	$(3,2)$
B	A	B	A	A	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(1,2)$	$(2,2)$	$(3,2)$
B	B	A	A	A	$(0,1)$	$(0,2)$	$(1,2)$	$(\mathbf{2},\mathbf{2})$	$(3,2)$

קיימים מצבי תיקו בכל הסידורים פרט לשני הראשונים ולכן:

$$P(\text{קיים תיקו עם } (3,2) \text{ קולות}) = \frac{8}{10} = \frac{4}{5}\,.$$

תשובה 2: נתחיל עם דיון איך לגשת לפתרון של השאלה השנייה. הנה מספר סידורים עבור $(a,b) = (3,2)$ עד לקבלת **התיקו הראשון**:

B מוביל עד תיקו	A מוביל עד תיקו
$B\quad A$	$A\quad B$
$B\quad B\quad A\quad A$	$A\quad A\quad B\quad B$

לכל סידור בו A מוביל עד לתיקו הראשון, קיים סידור שהוא תמונת ראי בו B מוביל עד לקבלת התיקו הראשון. השיקוף מתקבל על ידי החלפת כל A ב-B ולהפך. לפני התיקו הראשון אחד מהמועמדים חייב להוביל. אם הקול הראשון שנספר הוא עבור B חייב להיות תיקו בהמשך הספירה כי $a > b$.

ההסתברות שהקול הראשון היא עבור B היא:

$$P(\text{קול ראשון עבור } B) = \frac{b}{a+b}\,.$$

על ידי שיקוף המיקומים של הקולות, מספר הסידורים שמובילים לתיקו שמתחילים בקול עבור A שווה למספר הסידורים שמובילים לתיקו שמתחילים בקול עבור B. לכן:

$$P(\text{קיים תיקו }) = 2\cdot\frac{b}{a+b}\,.$$

בדיקה:

$$P(\text{קיים תיקו עם } (3,2) \text{ קולות}) = 2\cdot\frac{2}{2+3} = \frac{4}{5}\,.$$

סימולציה

```
For a =  3, b =  2:
Probability of a tie = 0.8000
Proportion of ties   = 0.8118
For a = 10, b =  8:
Probability of a tie = 0.8889
Proportion of ties   = 0.8977
For a = 20, b = 18:
Probability of a tie = 0.9474
Proportion of ties   = 0.9354
```

23. תיקו בהשוואת מטבעות (Ties in matching pennies)^D

הטל זוג מטבעות הוגנות N פעמים עבור N זוג, ורשום את מספר הפעמים שהזוגיות היא זוגי (עץ-עץ או פלי-פלי) ומספר ההפעמים שהזוגיות היא אי-זוגי (עץ-פלי או פלי-עץ). מה ההסתברות לקבל תיקו (לא כולל התיקו $0-0$ בהתחלה)?

שאלה 1: פתור עבור $N=4$ על ידי רישום כל התוצאות האפשריות.

שאלה 2: פתור עבור $N=6$ על ידי פיתוח נוסחה להסתברות.

שאלה 3: הסבר מדוע ההסתברות למספר אי-זוגי $N+1$ שווה להסתברות של המספר הזוגי N.

רמז: השתמש בפתרון לבעיה 22.

פתרון

תשובה 1: נסמן את ההטלות עם זוגיות זוגי ב-E וההטלות עם זוגיות אי-זוגי ב-O. בעשרה מתוך ששה עשר סידורי ההטלות יש תיקו (מודגש):

EOOO	**EOOE**	**EOEO**	**EOEE**	**EEOO**	EEOE	EEEO	EEEE
OOOO	OOOE	OOEO	**OOEE**	**OEOO**	**OEOE**	**OEEO**	**OEEE**

תשובה 2: לפי בעיה 22:

$$P(i \text{ תיקו בהטלה}) = \begin{cases} 2i/N & i \le N/2 \text{ אם} \\ 2(N-i)/N & i \ge N/2 \text{ אם}, \end{cases} \tag{12}$$

כי בבעיית הקלפי הראנו שהערך הנמוך יותר קובע את ההסתברות.

ההסתברות ל-i זוגיים ניתן על ידי ההתפלגות הבינומית:

$$P(i \text{ זוגיים}) = \binom{N}{i}\left(\frac{1}{2}\right)^{i}\left(\frac{1}{2}\right)^{N-i} = \binom{N}{i}\left(\frac{1}{2}\right)^{N} = 2^{-N}\binom{N}{i}. \tag{13}$$

ההסתברות לתיקו היא הסכום מעל i של ההסתברות לקבל i זוגיים כפול ההסתברות לתיקו בהטלה ה-i (משוואה 12). עבור $N = 6$ ותוך שימוש ב-$\binom{N}{i} = \binom{N}{N-i}$:

$$\begin{aligned}
P(\text{תיקו}) &= \frac{2 \cdot 2^{-6}}{6} \sum_{i=0}^{6} i \binom{i}{6} \\
&= \frac{1}{192} \left(0 \cdot \binom{6}{0} + 1 \cdot \binom{6}{1} + 2 \cdot \binom{6}{2} + 3 \cdot \binom{6}{3} + 4 \cdot \binom{6}{2} + 5 \cdot \binom{6}{1} + 6 \cdot \binom{6}{0} \right) \\
&= \frac{1}{192} \left(2 \cdot \binom{6}{1} + 4 \cdot \binom{6}{2} + 3 \cdot \binom{6}{3} \right) \\
&= \frac{132}{192} \approx 0.6875\,.
\end{aligned}$$

תשובה 3: התיקו הראשון בהטלה ה-$N+1$ מתרחש רק עם הספירה כמעט זהות לאחר ההטלה ה-N:

$$\begin{aligned}
&((N/2) - 1, (N/2) + 1) \\
&((N/2), (N/2)) \\
&((N/2) + 1, (N/2) - 1)
\end{aligned}$$

אבל ללא תלות התוצאת ההטלה הסופית הערכים לא יהיו שווים.

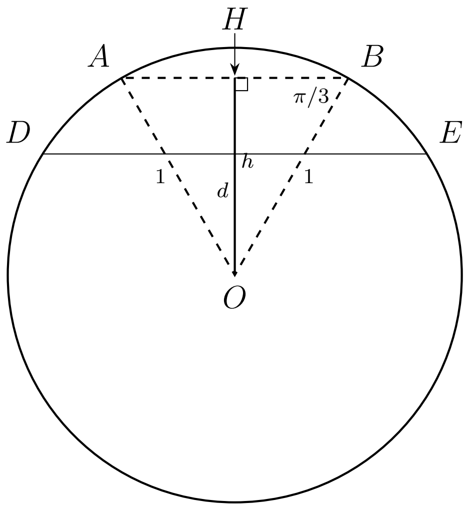
סימולציה

```
For  4 tosses:
Probability of ties = 0.6250
Proportion of ties  = 0.6192
For  6 tosses:
Probability of ties = 0.6875
Proportion of ties  = 0.6900
For  7 tosses:
Probability of ties = 0.6875
Proportion of ties  = 0.6811
For 10 tosses:
Probability of ties = 0.7539
Proportion of ties  = 0.7559
For 20 tosses:
Probability of ties = 0.8238
Proportion of ties  = 0.8255
```

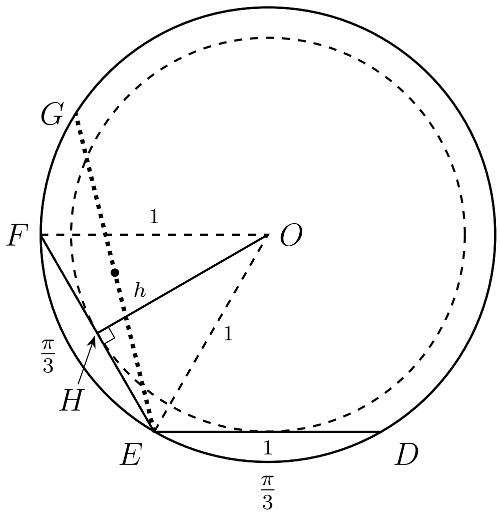
25. אורכים של מיתרים אקראיים (Lengths of random chords)

בחר מיתר אקראי במעגל היחידה. מה ההסתברות שאורכו של המיתר גדול מ-1?

כדי לפתור את הבעיה עליך להחליט איך לבחור מיתר אקראי. פתור את הבעיה עבור מהאפשרויות שלהלן:



איור 3(א) מרחק של מיתר מהמרכז בהתפלגות אחידה ב-$(0,1)$



איור 3(ב) נקודת האמצע של מיתר בהתפלגות אחידה במעגל ונקודות הקצה של המיתר בהתפלגות אחידה בהיקף

שאלה 1: התפלגות אחידה של מרחק המיתר מהמרכז המעגל.

שאלה 2: התפלגות אחידה של הנקודה האמצעית של המיתר בתוך המעגל.

שאלה 3: התפלגות נקודות הקצה של המיתר על היקף המעגל.

פתרון

תשובה 1: מיתר ארוך יותר מהרדיוס אם הוא קרוב יותר למרכז ממיתר באורך 1. יהי $\overline{AB}$ מיתר באורך 1 ובנה גובה $h = \overline{OH}$ מהמרכז O אל המיתר (איור 3(א)). בגלל ש-$\triangle AOB$ שווה-צלעות $\triangle OHB$ הוא משולש ישר-זווית:

$$h = \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}\,.$$

יהי d המרחק של המיתר $\overline{DE}$ מהמרכז. לפי ההנחה ההתפלגות של d אחידה ב-$(0,1)$ ולכן:

$$P(\overline{DE} > 1) = P(d < h) = \frac{h}{1} = \frac{\sqrt{3}}{2} \approx 0.8660\,.$$

תשובה 2: בנה מעגל עם מרכז O ורדיוס h כאשר h הוא אורכו של גובה לגובה למיתר באורך 1. משיק לכל נקודה על המעגל יהיה מיתר $\overline{FE}$ שאורכו 1. אורכו של כל מיתר $\overline{EG}$ שנקודת האמצע שלו נמצאת בתוך המעגל יהיה גדול מ-1 (איור 3(ב)), ולכן ההסתברות שאורכו של מיתר גדול מ-1 היא היחס בין השטחים של שני המעגלים:

$$P(\overline{EG} > 1) = \frac{\pi \cdot h^2}{\pi \cdot 1^2} = h^2 = \frac{3}{4}\,.$$

הסתברות זו היא הריבוע של ההסתברות שחישבנו בשאלה הקודמת.

תשובה 3: בנקודה שרירותית על ההיקף של מעגל היחידה (E באיור 3(ב)). כל נקודה אחרת על ההיקף (כגון G באיור) מגדירה מיתר שאורכו גדול מאחד אלא אם הנקודה שנבחרה נופל על הקשתות $\widehat{EF}$ או $\widehat{ED}$. לכן ההסתברות היא היחס בין הקשת $\widehat{FD}$ להיקף המעגל:

$$P(\overline{EG} > 1) = \frac{(2\pi - (2\pi/3)) \cdot 1}{2\pi \cdot 1} = \frac{2}{3}\,.$$

סימולציה הסימולציה היא עבור בחירת שתי נקודות על ההיקף.

```
Probability of long chords = 0.6667
Proportion of long chords  = 0.6627
```

26. הממהרים לדו-קרב (The hurried duelers)

A ו-B מגיעים לנקודת מפגש בזמן אקראי עם התלפגות אחידה בתוך פרק זמן של שעה. אם A מגיע קודם ו-B לא מגיע במשך 5 דקות, A עוזב. באופן דומה אם B מגיע קודם ו-A לא מגיע במשך 5 דקות, B עוזב. מה ההסתברות שהם יפגשו?

בפרק הזמן של שעה הזמן הוא **רציף** בתחום $[0, 1]$, כלומר, אי-אפשר לספור מספר בדיד של דקות או שניות כדי לחשב הסתברויות. כן ניתן לחשב הסתברויות של פרקי זמן.

רמז: צייר גרף כאשר ציר ה-x הוא זמן הגעתו של A וציר ה-y הוא זמן הגעתו של B.

פתרון

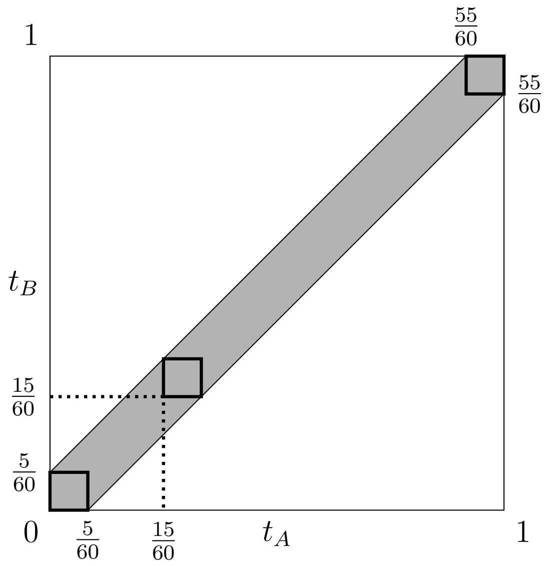
ללא הגבלת הכלליות הנח ש-A מגיע קודם. אם A מגיע ב-0 $t_A = 0$ ואם B מגיע לפני $t_B = 5/60$ הם נפגשים, אחרת הם לא נפגשים. מצב זה מוצג באיור 4 על ידי ריבוע קטן בראשית הצירים. אם A מגיע יותר מאוחר אזי גם B חייב להגיע באותו איחור. למשל, אם A מגיע ב-15 $t_A = 15$, B חייב להגיע בין $t_B = 15$ ו-$t_B = 20$. לכן הפגישה תתקיים בריבוע המתקבל על ידי הזזת הריבוע ב-15 מ-$(0, 0)$ ל-$(15/60, 15/60)$.

ההסתברות לפגישה היא היחס בין השטח האפור בגרף לשטחו של הריבוע הגדול. קל יותר לחשב את המשלים שהוא היחס בין שטח המשולשים הלבנים לשטחו של הריבוע הגדול:

$$\begin{aligned} P(\text{נפגשים } A, B) &= 1 - P(\text{לא נפגשים } A, B) \\ &= 1 - 2 \cdot \left(\frac{1}{2} \cdot \frac{55}{60} \cdot \frac{55}{60}\right) = \frac{23}{144} \approx 0.1597\,. \end{aligned}$$

סימולציה

```
Probability of meeting    = 0.1597
Proportion of meetings    = 0.1549
```



איור 4: זמנים המבטיחים מפגש בין A ל-B

27. לתפוס את הזייפן הזהיר (Catching the cautious counterfeiter)

נתון n קופסאות ובכל אחת n מטבעות כאשר מטבע אחד בכל קופסה מזויף. שלוף מטבע אחת מכל קופסה ובדוק אם היא מזויפת או אמיתית. מה ההסתברות שכל המטבעות שנשלפות מזויפות?

שאלה 1: פתור עבור $n = 10$.

שאלה 2: פתור עבור $n = 100$.

שאלה 3: פתח נוסחה עבור ההסתברות כאשר n שרירותי.

שאלה 4: פתח נוסחה עבור ההסתברות כאשר n שואב לאיסוף.

פתרון

השליפות בלתי תלויות ולכן ההסתברות שכל המטבעות אמיתיות היא מכפלת ההסתברות עבור שליפה אחת.

תשובה 1:

$$P(\text{כל המטבעות אמיתיים}) = \left(\frac{9}{10}\right)^{10} \approx 0.3487\,.$$

תשובה 2:

$$P(\text{כל המטבעות אמיתיים}) = \left(\frac{99}{100}\right)^{100} \approx 0.3660\,.$$

תשובה 3:

$$P(\text{כל המטבעות אמיתיים}) = \left(\frac{n-1}{n}\right)^{n}\,.$$

תשובה 4:

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n=\frac{1}{e}\,. \tag{14}$$

ניתן להוכיח את הגבול באמצעות חשבון דיפרנציאלי. תחילה ניתן לחשב את הגבול של הלוגריתם של הצד השמולי של משוואה 14:

$$\lim_{n\to\infty}\ln\left(1-\frac{1}{n}\right)^n=\lim_{n\to\infty}n\ln\left(1-\frac{1}{n}\right)=\lim_{n\to\infty}\frac{\ln\left(1-\frac{1}{n}\right)}{1/n}\,.$$

אם נחשב את הגבול נקבל $0/0=(\ln\ 1)/0$ אבל לפי חוק l'Hôpital נוכל להחליף את הביטוי בחילוק הנגזרות:

$$\lim_{n\to\infty}\ln\left(1-\frac{1}{n}\right)^n=\lim_{n\to\infty}\frac{\left(1-\frac{1}{n}\right)^{-1}(-(-n^{-2}))}{-n^{-2}}=-1$$
$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n=e^{-1}\approx 0.3679\,.$$

סימולציה

```
For  10 boxes:
Probability of all real = 0.3487
Proportion all real     = 0.3480
For 100 boxes:
Probability of all real = 0.3660
Proportion all real     = 0.3730
For 200 boxes:
Probability of all real = 0.3670
Proportion all real     = 0.3690
```

28. לתפוס את הזייפן החמדן (Catching the greedy counterfeiter)

נתון n קופסאות ובכל אחת n מטבעות מהן m מזוייפות. שלוף מטבע אחת מכל קופסה ובדוק אם היא מזויפת או אמיתית. מה ההסתברות $P(n,m,r)$ ש-r מתוך המטבעות מזוייפות?

שאלה 1: פתח נוסחה עבור $P(n,m,r)$.

שאלה 2: חשב $P(20,10,2), P(20,10,8), P(20,5,2), P(20,5,4)$.

פתרון

תשובה 1: יש $\binom{n}{r}$ תת-קבוצות של קבוצת הקופסאות מהן המטבעות המזוייפות נשלפו. לפי ההתפלגות הבינומית:

$$P(n,m,r)=\binom{n}{r}\left(\frac{m}{n}\right)^r\left(\frac{n-m}{n}\right)^{n-r}\,.$$

תשובה 2:

$$\begin{aligned}
P(20,10,2) &= \binom{20}{2}\left(\frac{10}{20}\right)^2\left(\frac{10}{20}\right)^{18} \approx 0.0002\\
P(20,10,8) &= \binom{20}{8}\left(\frac{10}{20}\right)^8\left(\frac{10}{20}\right)^{12} \approx 0.1201\\
P(20,5,2) &= \binom{20}{2}\left(\frac{5}{20}\right)^2\left(\frac{15}{20}\right)^{18} \approx 0.0669\\
P(20,5,4) &= \binom{20}{4}\left(\frac{5}{20}\right)^4\left(\frac{15}{20}\right)^{12} \approx 0.1952\,.
\end{aligned}$$

סימולציה

```
For 10 bad coins,  2 draws:
Probability of counterfeit  = 0.0002
Proportion counterfeit      = 0.0002
For 10 bad coins,  8 draws:
Probability of counterfeit  = 0.1201
Proportion counterfeit      = 0.1181
For  5 bad coins,  2 draws:
Probability of counterfeit  = 0.0669
Proportion counterfeit      = 0.0688
For  5 bad coins,  4 draws:
Probability of counterfeit  = 0.1897
Proportion counterfeit      = 0.1905
```

Mosteller משתמש בגבול מבעיה 27 כדי להראות שעבור m, r נתון, כאשר n שואף לאינסוף:

$$\lim_{n\to\infty} P(n,m,r) = \frac{e^{-m}m^r}{r!}\,. \tag{15}$$

הנה השוואה של ההסתברויות עבור $m = 10, r = 8$ עבור ערכים עולים של n:

```
Limit = 0.11259903, binomial = 0.11482303, n = 100
Limit = 0.11259903, binomial = 0.11282407, n = 1000
Limit = 0.11259903, binomial = 0.11262155, n = 10000
Limit = 0.11259903, binomial = 0.11259926, n = 1000000
```

29. עובש בג'לטין (Moldy gelatin)

נתון לוח מלבני שמחולק ל-n משבצות ריבועיות קטנות. בכל משבצת יש r חיידקים בממוצע.

שאלה 1: פתח נוסחה להסתברות שיש בדיוק r חיידקים ב-n המשבצות.

שאלה 2: חשב את ההסתברות עבור $n = 100, r = 3$.

רמז: בעיה זו דומה לבעיה 28.

פתרון

תשובה 1: p, ההסתברות שבמשבצת אחת נמצא חידק (נתעלם מהאפשרות שחיידק אחד מוכל באופן חלקי בשתי משבצות או יותר), היא m/n. $P(n,m,r)$, ההסתברות שיש בדיוק r חיידקים ב-n משבצות' ניתנת על ידי ההתפלגות הבינומית:

$$P(n,m,r) = \binom{n}{r} \left(\frac{m}{n}\right)^r \left(\frac{n-m}{n}\right)^{n-r} .$$

תשובה 2:

$$P(10,3,3) = \binom{100}{3} \left(\frac{3}{100}\right)^3 \left(\frac{97}{100}\right)^{97} \approx 0.2275 .$$

סימולציה

```
For  20 squares:
Probability of exactly  3 microbes  = 0.2428
Proportion of exactly   3 microbes  = 0.2436
Probability of exactly  5 microbes  = 0.2023
Proportion of exactly   5 microbes  = 0.1954
For 100 squares:
Probability of exactly  3 microbes  = 0.2275
Proportion of exactly   3 microbes  = 0.2247
Probability of exactly  5 microbes  = 0.1800
Proportion of exactly   5 microbes  = 0.1851
```

משוואה 15 מתאים גם כאן לחשב את הגבול כאשר n שואף לאינסוף:

$$\begin{aligned}
\lim_{n\to\infty} P(n,m,r) &= \frac{e^{-m}m^r}{r!} \\
\lim_{n\to\infty} P(n,3,3) &= \frac{e^{-3}\cdot 3^3}{3!} \approx 0.2240 \\
\lim_{n\to\infty} P(n,5,5) &= \frac{e^{-5}\cdot 5^5}{5!} \approx 0.1755 .
\end{aligned}$$

31. ימי הולדת זהים (Birthday pairings)

שאלה 1: עבור n אנשים מה ההסתברות $P(n)$ שלניים מהם או יותר יש יום הולדת זהה?

שאלה 2: מה מערך הקטן ביותר עבור n כך ש-$P(n) > 0.5$?

הנח שההתפלגות של ימי ההולדת היא אחידה בטווח $[1, 365]$.

רמז: חשב את המשלים להסתברות של-n אנשים ימי הולדת **שונים**.

פתרון

תשובה 1: בחר n אנשים אחד אחרי השני ובדוק אם יש להם אותו יום הולדת כמו הקודמים שנבחרו. עבור הראשון יש 365 ימים, עבור השני 364 ימים וכך הלאה:

$$1 - P(n) = \overbrace{\frac{365}{365} \cdot \frac{364}{365} \cdot \cdots \cdot \frac{365-(n-2)}{365} \cdot \frac{365-(n-1)}{365}}^{n} = \frac{365!/(365-n)!}{365^n} .$$

תשובה 2: חשב ערך זה עבור ערכים שונים של n או השתמש בתכנית מחשב כדי לחפש מ-1 ל-365 כדי למצוא את הערך הראשון עבורו $1 - P(n) < 0.5$. הערך הוא 23:

$$1 - P(23) = \frac{365!}{365^{23} \cdot 342!} \approx 0.4927 .$$

רוב האנשים מופתעים שהתשובה היא ערך כל קטן.

מחשבון מודרני מסוגל לחשב $1 - P(n)$ אבל שווה לחשב אותו באמצעות הקירוב של Stirling שהוא $\ln n! \approx n \ln n - n$:

$$\begin{aligned}
\ln(1 - P(n)) &= \ln\left(\frac{365!}{342! \cdot 365^{23}}\right) = \ln 365! - \ln 342! - 23 \ln 365 \\
&\approx (365 \ln 365 - 365) - (342 \ln 342 - 342) - 23 \ \ln 365 \\
&\approx -0.7404 \\
1 - P(n) &\approx e^{-0.7404} = 0.4769 .
\end{aligned}$$

הקורא מוזמן לחשב את ההסתברות עם הקירוב המדוייק יותר:

$$\ln n! \approx n \ln n - n + \frac{1}{6}\left(8n^3 + 4n^2 + n + \frac{1}{30}\right) + \frac{1}{2} \ln \pi .$$

סימולציה

```
For 21 people:
Expectation of no pairs = 0.5563
Average no pairs        = 0.5497
For 22 people:
Expectation of no pairs = 0.5243
Average no pairs        = 0.5237
For 23 people:
Expectation of no pairs = 0.4927
Average no pairs        = 0.4933
```

```
For 24 people:
Expectation of no pairs = 0.4617
Average no pairs        = 0.4576
For 25 people:
Expectation of no pairs = 0.4313
Average no pairs        = 0.4345
```

32. למצוא עמית ליום הולדת (Finding your birthmate)

עמית ליום הולדת, בקיצור עמית, הוא אדם עם יום הולדת זהה לשלך.

מדוע מציאת עמית היא בעיה שונה ממציאת זוג עם ימי הולדת זהים?

שאלה 1: כמה אנשים עליך לשאול עד ש-$Q(n)$, ההסתברות למציאת עמית, גבוהה מ-0.5?

שאלה 2: פתור את הבעיה על ידי שימוש במשוואה 14.

פתרון

להרבה אנשים יכול להיות יום הולדת זהה שנחשב כהצלחה במציאת זוג, אבל לא הצלחה במציאת עמית אלא אם יום ההולדת של האדם האחר זהה לשלך.

תשובה 1: נמצא את n, המספר הקטן ביותר של אנשים, כך ש-$1 - Q(n)$, ההסתברות שלאף אחד מהם הוא לא עמית, פחות מ-0.5. ההסתברות שהאדם הראשון שאתה שואל אינו עמית היא $364/365$, אבל זאת גם ההסתברות שהשני, השלילי, ..., אינו עמית. הפתרון הוא ה-k הקטן ביותר כך ש:

$$1 - Q(n) = \left(\frac{364}{365}\right)^n < \frac{1}{2}\,.$$

בחיפוש עם מחשב נמצא ש-$n = 253$:

$$\left(\frac{364}{365}\right)^{253} \approx 0.4995\,.$$

תשובה 2: משוואה 14 היא:

$$\lim_{n\to\infty}\left(\frac{n-1}{n}\right)^n = \frac{1}{e}\,,$$

וניתן להשתמש בה לחשב את ההסתברות:

$$\begin{aligned}1 - Q(n) &= \left(\frac{365-1}{365}\right)^n = \left[\left(\frac{364}{365}\right)^{365}\right]^{n/365}\\ &\approx e^{-n/365}\\ e^{-253/365} &\approx 0.5000\,.\end{aligned}$$

סימולציה

```
For 251 people:
Probability of no match = 0.5023
Proportion no match     = 0.5120
For 252 people:
Probability of no match = 0.5009
Proportion no match     = 0.5055
For 253 people:
Probability of no match = 0.4995
Proportion no match     = 0.4984
For 254 people:
Probability of no match = 0.4982
Proportion no match     = 0.4987
For 255 people:
Probability of no match = 0.4968
Proportion no match     = 0.5078
```

33. השוואת הבעיית יום הולדת זהה לבעיית עמית ליום הולדת

(Relating the birthday pairings and the birthmate problems)

סמן ב-$P(r)$ את ההסתברות למצוא זוג עם יום הולדת זהה עם r אנשים (בעיה 31), וב-$Q(n)$ את ההסתברות שמתוך n אנשים לפחות אחד הוא עימית שלך (בעיה 32). נתון r עבור איזה n, $P(r) \approx Q(n)$?

פתרון 1

הפתרון מבוסס על [7].

מהפתרון לבעיית 31 מתקבל:

$$\begin{aligned}
P(r) &= \frac{365}{365} \cdot \frac{365-1}{365} \cdot \frac{365-2}{365} \cdot \;\cdots\; \cdot \frac{365-(r-1)}{365} \\
&= 1\left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right) \cdot \;\cdots\; \cdot \left(1-\frac{r-1}{365}\right) \\
&\approx 1-\frac{1}{365}-\frac{2}{365}-\cdots-\frac{r-1}{365} \\
&= 1-\frac{1+2+3+\cdots+(r-1)}{365} \\
&= 1-\frac{r(r-1)/2}{365},
\end{aligned}$$

כאשר הקירוב במשוואה השלישית מתקבל מהשמטת חזקות של $1/365$ גדולות מאחת כי הן קטנות מדי להשפיע באופן מהותי על התוצאה.

באמצעות באותו קירוב, מהפתרון לבעיה 32 מתקבל:

$$\begin{aligned} P_{\text{אין עמית}}(n) &= \overbrace{\left(1-\frac{1}{365}\right)\left(1-\frac{1}{365}\right)\cdots\left(1-\frac{1}{365}\right)}^{n} \\ &\approx 1-\overbrace{\frac{1}{365}-\frac{1}{365}\cdots-\frac{1}{365}}^{n} \\ &\approx 1-\frac{n}{365} \end{aligned}$$

לכן $P(r)\approx Q(n)$ כאשר:

$$n=\frac{r(r-1)}{2}\,.$$

עבור $r=23$, $n=(23\cdot 22)/2=253$.

פתרון 2

Mosteller מביא את הפתרון האיטואיטיבי שלהלן:

> כאשר משווים בין בעיית הזוג והעמית, אנו שמים לב שעבור r אנשים בבעיית הזוג, קיימים $r(r-1)/2$ זוגות או **הזדמנויות** לידי הולדת זהים; לעומת זאת, אם שואלים n אנשים בבעיית העמית קיימות רק n הזדמנויות כדי שאמצא עמית אחד או יותר [7, p. 322].

מכאן הוא מסיק ש-$n\approx r(r-1)/2$.

ניתן להבין את הטיעון כך: בבעיית הזוג בחר תאריך שרירותי ושאל אם לשניים מתוך r **תאריך זה** הוא יום ההולדת שלהם. יש:

$$\binom{r}{2}=\frac{r!}{2!(r-2)!}=\frac{r(r-1)}{2}$$

דרכים שזה אפשרי. עבור בעיית העמית יום ההולדת שלך נתון, ולכל אחד מתוך n אנשים יש אותו סיכוי ליום הולדת זהה. על ידי השוואת שני הביטויים נקבל n עבורו $P_(r)\approx Q(n)$.

סימולציה

תוכל להריץ את הסימולציות לבעיות $31, 32$ ולבדוק תוצאה זו.

34. חופש בימי הולדת (Birthday holidays)^D

בית חרושת נסגר בכל יום שהוא יום הולדת של אחד העובדים. אין חופשות נוספות.

שאלה 1: כמה עובדים כדאי להעסיק כדי למקסם את מספר ימי-העבודה בשנה אחת?

שאלה 2: מה התוחלת של היחס בין מספר ימי-העבודה המקסימליים לבין 365^2, מספר ימי-העבודה עם כל אחד מ-365 העובדים עובדים כל יום?

רמז: הוכח על ידי בדיקת מקרי הקצה שמקסימום חייב להתקיים. אחר כך פתח נוסחה של התוחלת של ימי-העבודה ביום אחד.

פתרון

תשובה 1: אם יש רק עובד אחד יהיו 364 ימי-עבודה. אם יש שני עובדים מספר ימי-העבודה הוא $363 + 363 = 726$ (כאשר נתעלם המאפשרות הזניחה שלשניהם אותו יום הולדת). בקצה השני אם יש מיליון עובדים מספר ימי-העבודה יהיה אפס כמעט בוודאות. אם מספר ימי-העבודה עולה ואחר כך חוזר לאפס, חייב להיות מקסימום בין נקודות הקצה.

כדי לפשט את הסימונים נסמן את המספר הימים בשנה ב-N ומספר העובדים ב-n.

יהי $1 - \dfrac{1}{N} = \dfrac{N-1}{N} = p$. ההסתברות שיום נתון הוא יום-עבודה היא ההסתברות שלכל עובד יום הולדת בתאריך אחר:

$$\overbrace{\frac{N-1}{N} \cdot \;\cdots\; \cdot \frac{N-1}{N}}^{n} = \left(1 - \frac{1}{N}\right)^n = p^n \,.$$

ביום עבודה כל העובדים עובדים וביום חופש אף אחד לא עובד ולכן:

$$E(\text{ימי-עבודה ליום נתון}) = n \cdot p^n + 0 \cdot (1 - p^n) = np^n \,.$$

לכל ימי השנה אותה תוחלת ולכן:

$$E(\text{ימי-עבודה לשנה}) = Nnp^n \,. \qquad (16)$$

כדי למצוא את המקסימום נגזור את משוואה 16 ביחס ל-n ונשמתש ב-$(p^n)' = p^n \ln p$ ניתן להוכיח בעזרת חוק השרשרת:

$$(p^n)' = ((e^{\ln p})^n)' = (e^{n \ln p})' = e^{n \ln p}(n \ln p)' = (e^{\ln p})^n \ln p = p^n \ln p \,.$$

הנגזרת של משוואה 16 היא:

$$(Nnp^n)' = N(p^n + n(p^n)') = N(p^n + np^n \ln p) \,,$$

שהיא 0 כאשר:

$$n = -\frac{1}{\ln p} \,.$$

עבור $N = 365$ מתקבל $n = 364.5$ אבל n הוא מספר שלם ולכן המקסימום מתקבל ב-$n = 364$ או $n = 365$ שנותנים אותו תוחלת של ימי-עבודה:

$$\begin{aligned}
E(\text{ימי-עבודה לשנה}) &= Nnp^n \\
&= 365 \cdot 364 \cdot \left(\frac{364}{365}\right)^{364} \\
&= 365 \cdot 364 \cdot \frac{365}{365} \left(\frac{364}{365}\right)^{364} \\
&= 365 \cdot 365 \cdot \left(\frac{364}{365}\right)^{365} \\
&= 48944 \,.
\end{aligned}$$

תשובה 2: התוחלת של היחס היא:

$$E(\text{ימי-עבודה אפשריים}/\text{ימי-עבודה מקסימליים}) = \frac{365 \cdot 365 \cdot \left(\frac{364}{365}\right)^{365}}{365 \cdot 365} = \left(\frac{364}{365}\right)^{365} \approx 0.3674\,.$$

לפי משוואה 14:

$$\lim_{n\to\infty} E(\text{ימי-עבודה אפשריים}/\text{ימי-עבודה מקסימליים}) = \lim_{N\to\infty}\left(1-\frac{1}{N}\right)^{N} = \frac{1}{e}\,.$$

סימולציה

```
For 100 people
Expectation work-days    =  27742
Average work days        =  27743
Ratio work-days / 365**2 = 0.2082
For 250 people
Expectation work-days    =  45958
Average work days        =  45939
Ratio work-days / 365**2 = 0.3450
For 364 people
Expectation work-days    =  48944
Average work days        =  48936
Ratio work-days / 365**2 = 0.3674
For 365 people
Expectation work-days    =  48944
Average work days        =  48917
Ratio work-days / 365**2 = 0.3674
```

35. על שפת התהום (The cliff-hanger)

חלקיק מוצב על ציר ה-x במקום 1. בכל מקום על ציר ה-x הוא יכול לצעוד צעד ימינה עם הסתברות $2/3$ וצעד שמאלה עם הסתברות $1/3$ (איור 5).

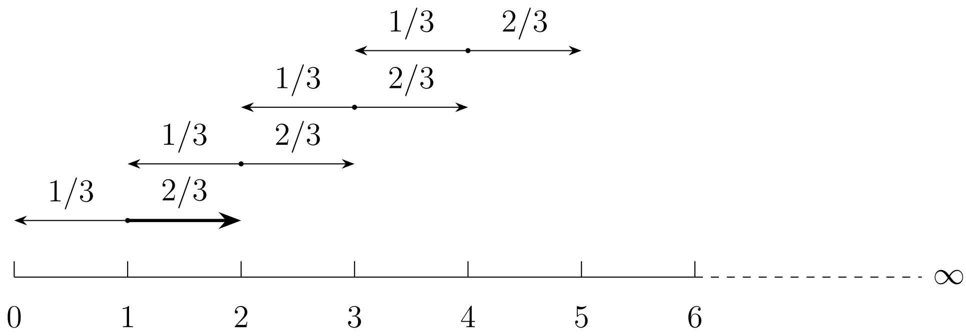
שאלה 1: מה ההסתברות שהחלקיק יגיע למקום 0 בסופו של דבר?

שאלה 2: אם ההסתברות של צעד ימינה היא p וההסתברות של צעד שמאלה היא $1-p$, מה ההסתברות שהחלקיק יגיע למקום 0 בסופו של דבר? נתח את האפשרויות לערכים שונים של p.

רמז: השתמש בהסתרויות מותנות לאחר הצעד הראשון.

פתרון

תשובה 1,2: נסמן צעד שמאלה ב-L וצעד ימינה ב-R. החלקיק יכול להגיע ל-0 ישירות על ידי צעד L עם הסתברות $\frac{1}{3}$, או על ידי צעד RLL עם הסתברות $\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}$, או על ידי צעד $RRLLL$ עם הסתברות $\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2$, ... החישוב נראה כמו סידרה הנדסית פשוטה אבל הוא מתעלם האפשרויות כגון $RLRLL$.



איור 5 : ?האם חלקיק יכול להגיע ל-0 (הציר אינסופי לימין)

נסמן ב-$P(i,j)$ את ההסתברות שהחלקיק מגיע ל-i מ-j. נחשב את ההתסתברות שהחלקיק מגיעה ל-0 מ-1 כתלות בצעד הראשון:

$$\begin{aligned} P(0,1) &= P(0,1|\text{צעד ראשון שמאלה}) + P(0,1|\text{צעד ראשון ימינה}) \\ &= (1-p)\cdot 1 + pP(1,2)P(0,1)\,. \end{aligned}$$

אבל $P(1,2) = P(0,1)$ ומתקבלת משוואה ריבועית ב-$P(0,1)$:

$$\begin{aligned} P(0,1) &= (1-p) + pP(0,1)^2 \\ pP^2(0,1) - P(0,1) + (1-p) &= 0 \\ P(0,1) &= 1,\ (1-p)/p\,. \end{aligned}$$

אם $p \leq 1/2$ אזי $(1-p)/p \geq 1$, כך ש-$P = 1$ הוא הפתרון היחיד ובטוח שהחלקיק יגיע ל-0.

אם $p = 1$ אזי $P = 0$ כי אם החלקיק תמיד צועד ימינה הוא לא יכול להחזור ל-0.

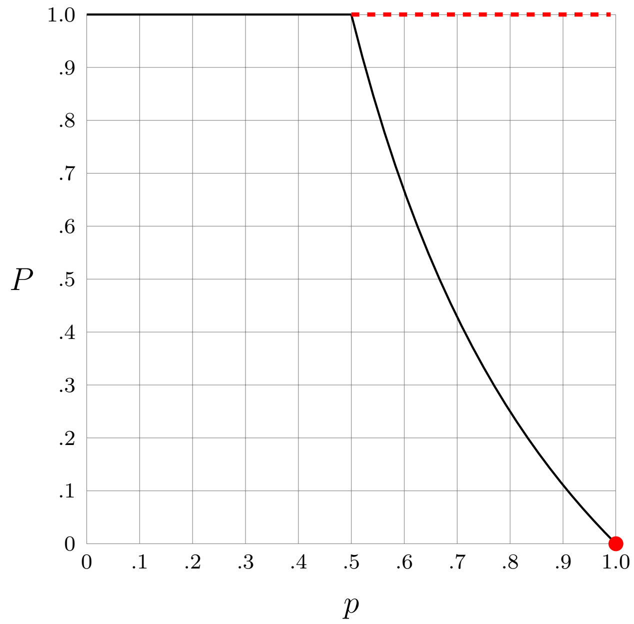
נניח ש-$P(0,1) = 1$ עבור $1/2 < p < 1$, כלומר, $P(0,1)$ **לא תלוי ב-**p. באיור 6 הקו האדום המקווקוו האדום מראה את עבור $P(1,0)$ כאשר p שואף ל-1 והנקודה האדומה מראה ש-$P(1,0) = 0$ כאשר p מגיע ל-1. $P(0,1)$ לא יכולה פתאום "לקפוץ" מ-1 ל-0 ולכן עבור $p > 1/2$ הפתרון היחיד הוא $P = (1-p)/p < 1$.[2]

עבור $p = 2/3, P = 1/2$ ועבור $p = 1/2, P = 1$. זה מפתיע כי לא היינו מצפים שהחלקיק יחזור תמיד ל-0 אם כיוון הצעד נקבע על ידי הטלת מטבע הוגן! אנו זקוקים למטבע ממש לא-הוגן (הסתברות של "עץ" שווה ל-2/3) כדי להשוות את הסיכויים לחזור ל-0 או לא.

סימולציה

```
For probability = 0.2500:
Probability of reaching 0 = 1.0000
Proportion reaching 0     = 1.0000
For probability = 0.5000:
Probability of reaching 0 = 1.0000
```

[2] Mosteller כותב שזה נובע מרציפות אבל הוא לא נותן הוכחה.



איור 6: הגרף של $P=\min(p/(1-p),1)$ עבור $p\in[0,1]$

```
Proportion reaching 0      = 0.9612
For probability = 0.6667:
Probability of reaching 0 = 0.5000
Proportion reaching 0      = 0.5043
For probability = 0.7500:
Probability of reaching 0 = 0.3333
Proportion reaching 0      = 0.3316
For probability = 0.8000:
Probability of reaching 0 = 0.2500
Proportion reaching 0      = 0.2502
```

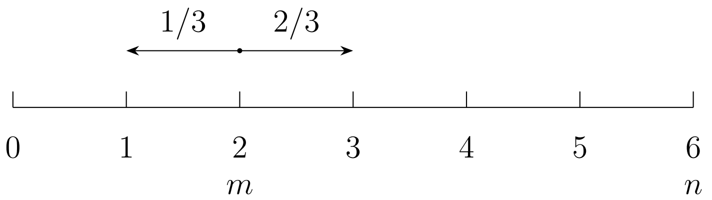
36. פשיטת הרגל של מהמר (Gambler's ruin)^D

חלקיק מוצב על ציר ה-x במקום $m\geq 1$. בכל מקום על ציר ה-x הוא יכול לצעוד צעד ימינה עם הסתברות $p>1/2$ וצעד שמאלה עם הסתברות $1-p$.

שאלה 1: מה ההסתברות שהחלקיק יגיע למקום 0 בסופו של דבר?

שאלה 2: יהי $m>n$. אם החלקיק מגיע למקום 0 או למקום n הוא מפסיק לצעוד (איור 7). מה ההסתברות שהחלקיק יגיע למקום 0? מה ההסתברות שהחלקיק יגיע למקום n?

הערה: בעיה 35 מייצגת מהמר המשחק עם כמות סופית של כסף נגד קזינו עם כמות בלתי מוגבלת של כסף. הבעיה מבקשת את ההסתברות שהמהמר יפסיד את כל כספו. בעיה 36(2) שואלת על מהמר אחד עם m שמשחק נגד מהמר שני עם $n-m$. הבעיה מבקשת את ההסתברויות ש**אחד** מהם מפסיד את כל כספו לשני.



איור 7: האם החלקיק יכול לחזור ל-0 (ציר סופי)?

פתרון

הפתרון מבוסס על [11, Chapter 2, Example 4m].

נסמן ב-$P(i,j)$ את ההסתברות שהחלקיק מגיע ל-i מ-j.

תשובה 1: ראינו בפתרון לבעיה 35 שעבור $p > 1/2$ (כאן נתון), אם חלקיק נמצא במקום 1 ההסתברות שלו להגיע ל-0 היא $r = (1-p)/p$. הסתברות זו לא תלויה במקום האבסולוטי של החלקיק ולכן:

$$P(i,j) = P(i+k, k+1) = P(i-k, j-k)\,,$$

ו-

$$P(0,m) = P(m-1,m)P(m-2,m-1)\cdots P(1,2)P(0,1) = r^m\,. \tag{17}$$

תשובה 2: סמן בקיצור $P_i = P(n,i)$ וחשב את P_i:

$$\begin{aligned}
P_i &= pP_{i+1} + (1-p)P_{i-1}\\
pP_{i+1} &= (p + (1-p))P_i - (1-p)P_{i-1}\\
p(P_{i+1} - P_i) &= (1-p)(P_i - P_{i-1})\\
P_{i+1} - P_i &= r(P_i - P_{i-1})\,.
\end{aligned}$$

$P_0 = 0$ כי אם החלקיק נמצא ב-0 הוא מפסיק לצעוד. לכן:

$$\begin{aligned}
P_2 - P_1 &= r(P_1 - P_0) = rP_1\\
P_3 - P_2 &= r(P_2 - P_1) = r^2 P_1\\
\cdots &= \cdots\\
P_i - P_{i-1} &= r(P_{i-1} - P_{i-2}) = r^{i-1}P_1\,.
\end{aligned}$$

רוב הגורמים בצד השמאלי מצטמצמים כאשר מחברים את המשוואות:

$$\begin{aligned}
P_i - P_1 &= P_1 \sum_{j=2}^{i} r^{j-1}\\
&= P_1 + P_1 \sum_{j=2}^{i} r^{j-1} - P_1\\
P_i &= P_1 \sum_{j=0}^{i-1} r^j = P_1 \left(\frac{1-r^i}{1-r}\right)\,.
\end{aligned} \tag{18}$$

אם חלקיק נמצא ב-n הוא כבר נמצא ב-n כך ש-$P_n = 1$:

$$1 = P_1 \left(\frac{1 - r^n}{1 - r} \right)$$

$$P_1 = \left(\frac{1 - r}{1 - r^n} \right) . \tag{19}$$

ממשוואות 18, 19:

$$P(n, i) = \left(\frac{1 - r^i}{1 - r^n} \right) . \tag{20}$$

בהוכחה סימטרית שמחליפה את p ו-$1 - p$:

$$P(0, i) = \left(\frac{1 - (1/r)^{n-i}}{1 - (1/r)^n} \right) . \tag{21}$$

הקורא מוזמן להראות שהסכום של משוואות 20, 21 הוא 1 כלומר שמובטח שאחד המהמרים ינצח והשני יפסיד. עבור $m = 1, n = 3, p = 2/3$:

$$P(0, 1) = \left(\frac{1 - \left(\frac{1}{2}\right)^1}{1 - \left(\frac{1}{2}\right)^3} \right) = \frac{4}{7}$$

$$P(3, 1) = \left(\frac{1 - 2^2}{1 - 2^3} \right) = \frac{3}{7} .$$

סימולציה

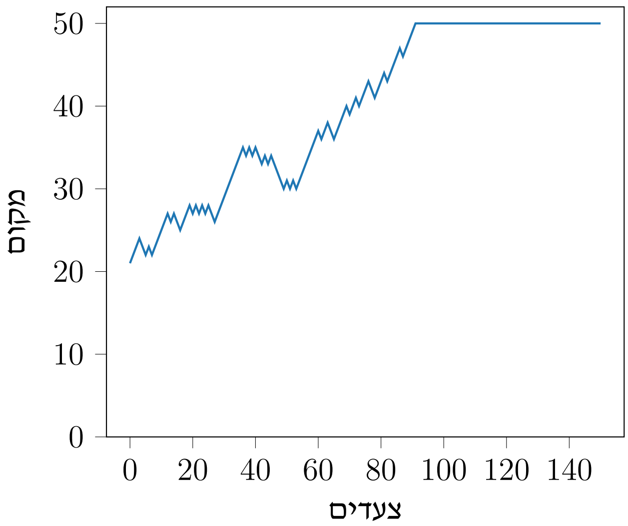
```
For probability = 0.6667:
Probability of reaching (0,10) from 1 = (0.4995,0.5005)
Proportion reaching     (0,10) from 1 = (0.5056,0.4944)
Probability of reaching (0,10) from 4 = (0.0616,0.9384)
Proportion reaching     (0,10) from 4 = (0.0643,0.9357)
Probability of reaching (0,10) from 6 = (0.0147,0.9853)
Proportion reaching     (0,10) from 6 = (0.0123,0.9877)

For probability = 0.7500:
Probability of reaching (0,10) from 1 = (0.3333,0.6667)
Proportion reaching     (0,10) from 1 = (0.3395,0.6605)
Probability of reaching (0,10) from 4 = (0.0123,0.9877)
Proportion reaching     (0,10) from 4 = (0.0115,0.9885)
Probability of reaching (0,10) from 6 = (0.0014,0.9986)
Proportion reaching     (0,10) from 6 = (0.0015,0.9985)
```

ככל שלמהמר בצד שמאל יש יותר והסתברות גבוהה היותר לזכות בכל צעד, כך ההסתברות שלו לניצחון גדלה.

גרף של הצעדים

הגרף נוצר על ידי ספריית `matplotlib` של Python. קוד המקור מופיע בקובץ `36-gamblers-ruin-plot.py`.

פשיטת הרגל של מהמר $p = 0.67$, $n = 50$, $m = 20$



37. משחק נועז או משחק זהיר (Bold play vs. cautious play)

רולט ניתן להמר שהכדור יפול בכיס המסומן במספר זוגי. ההסתברות היא $18/38$ כי יש 18 מספרים זוגיים, 18 מספרים אי-זוגיים ו-2 מספרים ירוקים עליהם הקזינו זוכה.

איזו מהאסטרגיות שלהלן עדיפה?

1. משחק נועז: להמר 20 בסיבוב אחד.
2. משחק זהיר: להמר 1 בכל סיבוב עד שאתה זוכה או מפסיד 20.

רמז: השתמש בתוצאות של בעיה 36. **פתרון**

ההסתברות לזכיה עם משחק נועז היא $18/38 \approx 0.4737$.

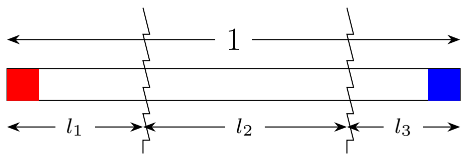
ההסתברות לזכיה עם משחק זהיר היא (משוואה 20):

$$\begin{aligned} r &= \frac{20}{38} \Big/ \frac{18}{38} = \frac{20}{18} \\ P(40,20) &= \frac{1-(20/18)^{20}}{1-(20/18)^{40}} \approx 0.1084\,. \end{aligned}$$

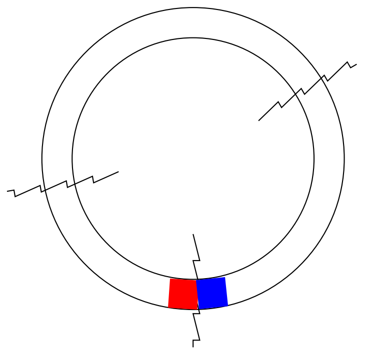
ברור שמשחק נועז עדיף על משחק זהיר.

Mosteller מביא הסבר איטואטיבי: הימור בסיבובים רבים חושף את המהמר לאפשרות שהקזינו מנצח בהסתברות $2/38$.

סימולציה



איור 8(א) חלוקת מקל לשלושה חלקים



איור 8(ב) חלוקת טבעת לשלושה חלקים

```
Probability of bold wins     = 0.4737
Proportion bold wins         = 0.4677
Probability of cautious wins = 0.1084
Proportion cautious wins     = 0.1094
```

39. הכימאי המגושם (The clumsy chemist)

נתון מספר רב של מקלות מזכוכית באורך 1. קצה אחד צבוע באדום ושני בכחול. כאשר זורקים אותם על הרצפה, הם נשברים לשלושה חלקים עם התפלגות אחידה של האורכים (איור 8(א)). מה התוחלת של אורכו של החלק בקצה הכחול?

רמז: במקום מקלות ישרים הנח שקבלת טבעות זכוכית ללא סימנים שגם הם נשברים לשלושה החלקים (איור 8(ב)).

פתרון 1

אין סימטריה במלקות כי הקצות שונים מהחלק האמצני. אולם הטבעת סימטרית ולכן ההתפלגויות של שלושת החלקים יהיו אחידות עם תוחלת $1/3$. על ידי צביעת אחת מנקודות השבירה כפי שמופיע באיור 8(ב), הבעיה הופכת להיות זהה לבעיית המקל כך שההתפלגויות זהות. לכן התוחלת של אורכי החלקים היא גם $1/3$.

פתרון 2

הפתרון אלגנטי שלהלן מבוסס על [4].

נניח שהמקל מייצג את קטע הקו $(0,1)$. המקל נשבר בשני מקומות שניתן לייצג בני משתנים אקראיים בלתי-תלויים עם התפלגות אחידה $X, Y \in (0,1)$. נחשב את ההסתברות $P(|X-Y|>a)$.

טבלה 1 מראה נקודות (x,y) כאשר $x, y \in \{0.0, 0.1, 0.2, \ldots, 0.9\}$ והנקודה העשרונית הושמטה. הערכים בטבלה הם $|X-Y|$. עבור $a=0.6$ הערכים למעלה משמאל מ-$(0,6)$--$(6,9)$, והערכים למטה ומימין מ-$(6,0)$--$(9,6)$, הם התוצאות שמגדירות את $P(|X-Y|>a)$:

$$P(|X-Y|>a) = 2 \cdot \frac{1}{2}(1-a)(1-a) = (1-a)^2 \,.$$

a

	9	9	8	7	6	5	4	3	2	1	0	
	8	8	7	6	5	4	3	2	1	0	1	
	7	7	6	5	4	3	2	1	0	1	2	
a	6	6	5	4	3	2	1	0	1	2	3	
	5	5	4	3	2	1	0	1	2	3	4	
	4	4	3	2	1	0	1	2	3	4	5	
y	3	3	2	1	0	1	2	3	4	5	6	a
	2	2	1	0	1	2	3	4	5	6	7	
	1	1	0	1	2	3	4	5	6	7	8	
	0	0	1	2	3	4	5	6	7	8	9	
		0	1	2	3	4	5	6	7	8	9	

x a

טבלה 1: התפלגות נקודות השבירה ב-$(0,1)\times(0,1)$

עבור $a=0.6$, $P(|X-Y|>0.6)=(0.4)^2=0.16$.

המשלים הוא:

$$P(|X-Y|<a)=1-(1-a)^2\,.$$

הסתברות זו היא ההתפלגות ההסתברות המצטברת cumulative probability distribution (CPD) בקטע. $(0,1)$. ניתן לקבל את פונצקית הסתברות הצפיפות probability density function (PDF) על ידי גזירת ה-CDP:

$$P(|X-Y|=a)=\frac{d}{da}P(|X-Y|<a)=\frac{d}{da}(1-(1-a)^2)=2(1-a)\,.$$

התוחלת היא האינטגרל ה-PDF כפול הערך:

$$E(|X-Y|)=\int_0^1 a\cdot 2(1-a)\,da=2\left(\frac{a^2}{2}-\frac{a^3}{3}\right)\bigg|_0^1=\frac{1}{3}\,.$$

סימולציה

```
Expectation of length of right piece = 0.3333
Average length of right piece        = 0.3359
```

40. האס הראשון (The first ace)

לק קלפים מחפיסה מעורבת היטב עד שמופיע אס. מה התוחלת של מספר הקלפים שיש לחלק?

רמז: חשוב על חפיסת קלפים ללא האסים מסודרת בשורה. **פתרון** הקלפים הם כמו "מקלי" באורך 48

״שנשבר״ על ידי 4 ל-5 חלקים. הפתרון של בעיה 39 מתאים גם כאן והתוחלת של חלק היא $48/5 = 9.6$.

סימולציה

```
Expectation of first ace = 9.6000
Average first ace        = 9.5805
```