פתרון מקוצר לבחינה של ד"ר רון פלד מ 16/06/14

חלק א

סעיף א

יש $n-k+1$ מקומות שבהם יכול להתחיל רצף באורך k. בכל אחד מהם ההסתברות לקיום רצף היא 3^{-k}. תוחלת סכום האינדיקטורים שווה לסכום תוחלתם.

סעיף ב

שונות סכום האינדיקטורים שווה לסכום השונויות שלהם בתוספת פעמיים סכום השונויות המשותפות שביניהם.
השונות של כל אינדיקטור היא $3^{-k}\left(1-3^{-k}\right)$. כאמור יש $n-k+1$ אינדיקטורים.
אין תלות בין מאורעות שבשני מקומות שמרוחקים בלפחות k מקומות יתחיל רצף (כי אין להם תווים משותפים). לכן גם השונות המשותפת בין זוגות כאלה של אינדיקטורים היא אפס.
נחשב את השונות המשותפת שבין שני זוגות שהמרחק בין נקודות ההתחלה שלהם הוא $m<k$: כדי ששניהם יתרחשו צריך רצף של $k+m$ הצלחות במקומות מסוימים. כפי שאמרנו, התוחלת של כל אחד מהם היא 3^{-k}. לכן השונות המשותפת ביניהם היא $3^{-k-m}-3^{-k}\cdot 3^{-k}$. יש $n-k-m+1$ זוגות של רצפים בעלי מרחק m ביניהם.

סעיף ג

המאורע $\left(M>(1+\varepsilon)\log_3 n\right)$ הוא המאורע שעבור $A=\lfloor(1+\varepsilon)\log_3 n+1\rfloor$ יש לפחות רצף אחד באורך A.
ההסתברות שבמקום מסוים מתחיל רצף באורך A הוא 3^{-A}. ההסתברות שלפחות במקום אחד יתחיל רצף כזה אינה גדולה מ $(n-A+1)3^{-A}$ (הסתברות איחוד אף פעם לא גדולה מסכום ההסתברויות). הסתברות זו אינה גדולה מ $n\cdot 3^{-A}$. עבור $A>(1+\varepsilon)\log_3 n$, הגבול של ביטוי זה הוא אפס. המאורע שיש רצף ארוך יותר מוכל במאורע זה.
<u>הערה</u>
אפשר לפתור גם בעזרת אי שיוויון מרקוב.

סעיף ד

המאורע $\left(M>(1-\varepsilon)\log_3 n\right)$ הוא המאורע שעבור $A=\lfloor(1-\varepsilon)\log_3 n+1\rfloor$ יש לפחות רצף אחד באורך A.
הסיכוי לרצף באורך A החל ממקום מסוים שהחל ממנו נשארו עוד לפחות A מקומות הוא 3^{-A}. עבור $A\le(1-\varepsilon)\log_3 n$, סיכוי זה גדול מ $n^{-(1-\varepsilon)}$. עבור $A\le(1-\varepsilon)\log_3 n$ יש לפחות $\frac{n}{\log_3 n}-1$ מקומות שבהם מתחילות מחרוזות זרות באורך A (כל אחת מתחילה אחר סיום האחרת). אין תלות בין הקורה במחרוזות זרות. לכן ההסתברות שבאף אחת מהן לא יהיה

רצף באורך A אינה גדולה מ $\left(1-n^{-(1-\varepsilon)}\right)^{\frac{n}{\log_3 n}-1}$. הגבול של ביטוי זה הוא אפס עבור כל $\varepsilon>0$ (עבור כל $\varepsilon>0$, $\frac{n}{\log_3 n}$ גדל יותר מהר מאשר $n^{(1-\varepsilon)}$).

<u>הערה</u>
ניתן לפתור את השאלה גם על-ידי שימוש בשיטת המומנט השני או אי שיוויון צ'בישב, וכך לקבל תוצאה חזקה יותר עבור כל פונקציה $\log_3 n-f(n)$ כאשר $\lim\limits_{n\to\infty} f(n)=\infty$.

<u>חלק ב</u>

1	2	3	4	5	6	7	8	9	10	11	12
ב	ג	א	ד	ג	א	ג	א	ב	ג	א	ב

<u>הסברים קצרים</u>

<u>שאלה 1</u>

בהינתן המאורע B , X_2 מקבל בסיכוי שווה את כל אחד מהערכים $1,3,4,6,7$ ותוחלתו היא $\frac{1+3+4+6+7}{5}$.

<u>שאלה 2</u>

X_i מקבל את כל אחד מ i ערכים בסיכוי שווה. בדיוק אחד מערכים אלה הוא המכסימום.
בינתן כל קבוצה של i ערכים שונים שמקבלים i המשתנים הראשונים, ההסתברות שהאחרון שביניהם מקבל את הערך המכסימלי היא $\frac{1}{i}$. זה נכון לכל קבוצת ערכים שונים שהם מקבלים. לכן לפי הסתברות שלמה הסיכוי הוא $\frac{1}{i}$.

<u>שאלה 3</u>

בלי שום קשר לסדר של $i-1$ המספרים הקודמים, מתקיים $P(A_i)=\frac{1}{i}$.
עבור כל סידור של $i-1$ המשתנים הראשונים, הסיכוי של המשתנה ה-i להיות מכסימלי הוא $\frac{1}{i}$.

<u>שאלה 4</u>

מתקיים $\sum_{i=1}^{\infty}\frac{1}{i}=\sum_{i=1}^{n}P(A_i)=E(S_n)$ (תוחלת סכום של אינדיקטורים שווה לסכום התוחלות).

עבור ערכי n גדולים מתקיים $\sum_{i=1}^{n}\frac{1}{i}\approx\ln(n)$. לכן החל ממקום מסוים מתקיים למשל $\sum_{i=1}^{n}\frac{1}{i}<7\ln(n)$ (המספר 7 הוא שרירותי).

לפי אי שיוויון מרקוב נקבל עבור אותם ערכי n : $P\left(S_n\ge\sqrt{n}\right)\le\frac{E(S_n)}{\sqrt{n}}<\frac{7\ln(n)}{\sqrt{n}}$.

מתקיים $\lim_{n\to\infty}\frac{7\ln(n)}{\sqrt{n}}=0$.

<u>שאלה 5</u>

בהסתברות $\frac{4}{9}=\frac{2}{3}\cdot\frac{2}{3}$ נגיע ל 3 לפני ביקור ראשון ב 0. לכן בהסתברות $\frac{5}{9}$ נגיע ל 0 לפני שנגיע ל 3. בהמשך אסור לקבל 3 הצלחות רצופות. הסיכוי שלא יתקבלו 3 הצלחות רצופות הוא $1-\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}=\frac{19}{27}$.

לכן, הפתרון הוא $\frac{5}{9}\cdot\frac{19}{27}$.

<u>שאלה 6</u>

<u>הסבר בדרך ראשונה</u>

ניתן לחזור ממצב 5 למצב 5 ב 6 צעדים (על-ידי מעבר מ 5 ל 0 ואחר-כך 5 זכיות רצופות).

ניתן לחזור ממצב 5 למצב 5 ב 7 צעדים (למשל על-ידי מעבר מ 5 ל 0 ואחר-כך הפסד שאחריו יבואו 5 נצחונות).

המחלק המשותף המכסימלי של 6 ו 7 הוא 1. לכן בודאי המחלק המשותף המכסימלי של כל זמני החזרה האפשריים למצב 5 הוא 1.

<u>הסבר בדרך שניה</u>

המחזור של מצב 0 הוא 1 (כי ניתן לחזור מ 0 ל 0 בצעד אחד).

5 ו 0 הם מצבים באותה מחלקה בלתי פריקה. מכיון שאי מחזוריות היא תכונה מחלקתית, אז גם מצב 5 הוא לא מחזורי.

<u>שאלה 7</u>

מתקיים $\lim_{t\to\infty}\frac{E(S_t)}{t}=10\pi_9-\pi_0$.

מתקיים $\pi_0=\sum_{k=0}^{9}\pi_k P_{k,0}=\sum_{k=0}^{9}\pi_k\cdot\frac{1}{3}=\frac{1}{3}$.

עבור כל $1\le i\le 8$ מתקיים $\pi_i=\sum_{k=0}^{9}\pi_k P_{k,i}=\frac{2}{3}\pi_{i-1}$.

לכן $0.97 < \frac{1}{3}\left[\frac{1}{1-\frac{2}{3}} - \frac{\left(\frac{2}{3}\right)^9}{1-\frac{2}{3}}\right] = \frac{1}{3}\left[\sum_{i=0}^{\infty}\left(\frac{2}{3}\right)^i - \sum_{i=9}^{\infty}\left(\frac{2}{3}\right)^i\right] = \sum_{i=0}^{8}\frac{1}{3}\cdot\left(\frac{2}{3}\right)^i = \sum_{i=0}^{8}\pi_i$.

מתקיים $0.03 = 1-0.97 > \pi_9 = 1 - \sum_{i=0}^{8}\pi_i$.

נקבל $0 > 10\cdot 0.03 - \frac{1}{3} > 10\pi_9 - \pi_0$.

<u>הערה</u>

אפשר היה גם לחשב את π_9: מתקיים $\pi_9 = \frac{2}{3}\pi_8 + \frac{2}{3}\pi_9$.

לכן, $\pi_9 = 2\pi_8 = 2\cdot\pi_0\left(\frac{2}{3}\right)^8 = 2\cdot\left(\frac{2}{3}\right)^9$.

שאלה 8

תוחלת סכום תמיד שווה לסכום התוחלות. עבור כל i, j מתקיים הודות לאי תלות:
$E(X_i X_j) = E(X_i)E(X_j) = 0$.

שאלה 9

שונות סכום המכפלות שווה לסכום השונויות שלהם בתוספת סכום השונויות המשותפות של המכפלות השונות.

יש $\binom{n}{2} = \frac{n(n-1)}{2}$ זוגות של מכפלות. כל המכפלות של זוגות הם שווי התפלגות.

מתקיים $V(X_1X_2) = E\left((X_1X_2)^2\right) - \left(E(X_1X_2)\right)^2 = 1 - 0^2 = 1$.

לכן סכום השונויות של המכפלות היא $\frac{n(n-1)}{2}$.

נראה שהשונויות המשותפות בין המכפלות השונות הן אפסים.

עבור מכפלות שאין להן איברים משותפים, יש אי תלות ולכן השונות המשותפת היא אפס.

נראה שהשונות המשותפת בין X_1X_2 ל X_1X_3 היא אפס.

$$E\left((X_1X_2)(X_1X_3)\right) - E(X_1X_2)E(X_1X_3) = E\left(X_1^{\,2}X_2X_3\right) - 0 = 0$$

לכן שונות הסכום שווה לסכום השונויות של המכפלות.

שאלה 10

נראה שהשונות המשותפת בין S ו Q היא אפס. מכאן מתקבל שמקדם המתאם הוא אפס.

$$Cov(S,Q) = Cov\left(\left(\sum_{k=1}^{n} X_i\right), \left(\sum_{i=1}^{n}\sum_{j=i+1}^{n} X_i X_j\right)\right) = \sum_{k=1}^{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n} Cov(X_k, X_i X_j)$$

נעשה חלוקה למקרים:

אם $k \neq i, k \neq j$, אז בגלל אי תלות השונות המשותפת היא אפס.
נניח בלי הגבלת הכלליות ש $k = i$: נקבל

$$Cov(X_i, X_i X_j) = E((X_i)^2 X_j) - E(X_i^2)E(X_j) =$$
$$= E(X_i^2)E(X_j) - E(X_i^2)E(X_j) = 0 - 0 = 0$$

<u>שאלה 11</u>

S^2 הוא סכום של ריבועי משתנים ומכפלות של זוגות של משתנים. ריבועי המשתנים הם קבועים ששווים ל 1. סכום המכפלות האחרות הן פונקציה לינארית עולה של המשתנה Q.
נראה זאת גם על-ידי חישוב:

$$S^2 = (X_1 + X_2 + \cdots + X_n)^2 = \sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_i X_j = n + 2Q$$

לכן יש קשר לינארי עולה בין שני המשתנים. קשר לינארי עולה הוא תנאי מספיק והכרחי למתאם של $+1$.

<u>שאלה 12</u>

יהי Y משתנה שמודד את מספר המשתנים X_i שמקבלים ערך חיובי. $(Y = y)$ אם y משתנים מקבלים ערך $+1$. $E(Y) = n/2$, $\sqrt{Var(Y)} = \sqrt{n \cdot 0.5 \cdot 0.5} = 0.5\sqrt{n}$

המשתנה Q מקבל ערך אי שלילי אם"ם $\binom{y}{2} + \binom{n-y}{2} \geq 0.5\binom{n}{2}$ או אם $y(n-y) \leq 0.5\binom{n}{2}$. שיוויון בין שני האגפים מתקבל בנקודות $\frac{n \pm \sqrt{n}}{2}$.

נראה גם בגישה נוספת:
$Q = (S^2 - n)/2$ ולכן המאורע $(Q \geq 0)$ שקול למאורע $(S^2 \geq n)$ או ל $(|S| \geq \sqrt{n})$.
נעשה שימוש במשפט הגבול המרכזי.

שלומי