פתרון <u>אפשרי</u> למבחן סיום לקורס "הסתברות וסטטיסטיקה":
תכנית רוטשילד ויצמן

1. אתה משתתף במשחק בו מוטלות שתי קוביות הוגנות (על כל אחת המספרים מ-1 ועד 6). אתה יכול לבחור בין ניחוש כי אחרי ההטלה סכום המספרים המופיעים בצד העליון של הקוביות הוא בין 2 לבין 6 או כי הסכום הוא בין 8 לבין 12. אם תצדק תקבל 20 שקלים. אם סכום המספרים המופיעים למעלה הוא 7 תשלם 50 שקלים. על מי מבין שתי האפשרויות כדאי לך להמר ומה תוחלת הרווח או ההפסד שלך?

נבחר כמרחב מדגם את זוגות המספרים (i,j) כאשר i ו- j הם בין 1 ל- 6, ולכל זוג הסתברות של 1 חלקי 36. ספירה פשוטה מראה כי ישנם 15 זוגות שסכום המספרים שלהם הוא בין 2 ל- 6 ו-15 זוגות שסכום המספרים עליהם הוא בין 8 ל- 12. לכן לא משנה אם אנו מהמרים על הסכום הראשון או השני. בכל מקרה תוחלת הסכום שנקבל היא 15 כפול 20 חלקי 36, כלומר 300 חלקי 36 שקלים. סכום של 7 קורה בהסתברות של 6 חלקי 36 ואז התשלום הוא 50 שקלים. לכן תוחלת התשלום היא 300 חלקי 36 שקלים. מכאן שתוחלת הרווח מהמשחק כולו היא 0.

2. בוחרים נקודה בעיגול היחידה במישור באופן אקראי, כאשר האקראיות היא לפי מידת השטח, כלומר, לשתי קבוצות להן שטחים שווים בעיגול, יש סיכוי זהה לכך שהנקודה תבחר מהן. מה תוחלת המרחק של הנקודה שנבחרה מהראשית?

נציג את העיגול בקואורדינטות פולריות, r בין 0 ל- 1 והזווית θ בין 0 ל 2π. אלמנט השטח הוא $rdrd\theta$. המרחק של נקודה מהראשית הוא הקואורדינטה r . שטח העיגול הוא π. לכן התוחלת המבוקשת מתקבלת על ידי האינטגרל של המשתנה r^2 לפי המשתנה dr בין 0 ל-1 והמשתנה $d\theta$ בין 0 ל 2π וכל זה מחולק ב- π . התשובה היא 2/3 .

3. "אין אפשרות להדגים (באופן לא טריוויאלי) את החוק החזק של המספרים הגדולים על מרחב מדגם סופי". הסבר מדוע.

החוק החזק של המספרים הגדולים מתייחס לסדרה של משתנים מקריים בלתי תלויים ושווי התפלגות. אם מרחב המדגם סופי וקבוצת הערכים (הסופית) נקבעת מראש על ידי ההתפלגות הזהה, הרי מספר המשתנים המקריים גם הוא סופי. לכן בסדרה אינסופית יהיו שניים שהם זהים. למעט המקרה הטריוויאלי בו כל המשתנים הם בעלי אותו ערך קבוע (ולכן בלתי תלויים) שני משתנים מקריים זהים תלויים זה בזה, ולכן אי אפשר יהיה להדגים (באופן לא טריוויאלי) את החוק החזק של המספרים הגדולים על מרחב כזה.

4. א. הצג דוגמא של משתנה מקרי עם תוחלת אינסופית. ב. הצג דוגמא (או לפחות הסבר כיצד לבנות דוגמא) למשתנה מקרי בעל תוחלת סופית אך שונות אינסופית. ג. האם יתכן משתנה מקרי עם תוחלת אינסופית אך שונות סופית?

נגדיר מרחב הסתברות שהוא האינטרוול $[-1/2, 1/2]$ עם מידת האורך. א. המשתנה המקרי הנותן לנקודה $x > 0$ את הערך x^{-1} ולשאר הנקודות את הערך 0 הוא עם תוחלת אינסופית, זאת כי התוחלת היא האינטגרל של הפונקציה לפי המידה והאינטגרל של x^{-1} על $(0,1/2]$ הוא אינסוף. ב. על אותו מרחב נגדיר משתנה אחר שמקבל את הערך $\sqrt{x^{-1}}$ על האינטרוול $(0,1/2]$ ואת הערך $-\sqrt{(-x)^{-1}}$ (כלומר ההשלמה האי זוגית של הפונקציה הקודמת) על האינטרוול $[-1/2, 0)$. התוחלת של המשתנה הזה היא סופית כי ערך האינטגרל של $\sqrt{x^{-1}}$ על $(0,1/2]$ הוא סופי. האי זוגיות גוררת כי האינטגרל של המשתנה החדש על כל האינטרוול הוא 0 ולכן התוחלת של המשתנה על כל האינטרוול היא אפס. השונות היא אינטגרל ריבוע ההפרש בין ערך המשתנה לבין התוחלת. לכן, במקרה שלנו השונות תהיה האינטגרל של x^{-1} על $(0,1/2]$ ועוד האינטגרל של $(-x)^{-1}$ על $[-1/2, 0)$. כבר ראינו כי האינטגרל הזה אינסופי. ג. בהגדרת השונות (שזה עתה כתבנו) צריך להחסיר את התוחלת מערך המשתנה. לכן, אם התוחלת אינסופית השונות כלל אינה מוגדרת. בפרט לא יתכן משתנה מקרי עם תוחלת אינסופית ושונות סופית.

5. במדינת מסוימת נהוג לעצור באופן אקראי נהגים לבדיקת נשיפה כדי לוודא שאינם נוהגים תחת אלכוהול, זאת מכיוון שידוע כי עשרים אחוזים של הנהגים שם אינן זהירים ונוהגים תחת השפעת אלכוהול. אלא שבדיקת הנשיפה אינה אמינה במאה אחוזים: : הבדיקה תאתר נהג תחת השפעה רק בהסתברות של 90 אחוזים, בעוד שבהסתברות של 5 אחוזים הבדיקה תצביע על צריכת אלכוהול מעל המותר לנהג שלא צרך קודם לכן אלכוהול מעל המותר. נהג נעצר והבדיקה הראתה כי צרך אלכוהול מעל המותר. מה הסיכוי שהנהג אכן נהג תחת השפעת אלכוהול?

אדם שהבדיקה הראתה כי הוא נוהג תחת השפעה יכול להיות בין 20 האחוזים שאכן נוהגים תחת השפעה, אך הסיכוי שהבדיקה תראה זאת הוא רק 90 אחוזים, או בין 80 האחוזים הנוהגים לפי החוק אך בכל אופן הבדיקה טוענת, בסיכוי של 5 אחוזים, כי שתו יותר מדי. לכן הסיכוי שהנהג שנעצר נהג תחת השפעת אלכוהול הוא 90 כפול 20 חלקי המספר המתקבל על ידי 90 כפול 20 ועוד 5 כפול 80 (זו נוסחת בייס למקרה הזה). חישוב פשוט מראה כי הסיכוי הוא 81.81 אחוזים.

6. מאפייה שהתחייבה לספק לחמניות לתושבי עיירה בת 40,000 נפש מוצאת כי ממוצע הביקוש ללחמניות הוא 10,000 ליום. בהנחה כי הדרישות היומיות ללחמנייה הן בלתי תלויות ושוות התפלגות בין תושב ותושב ובין יום ליום, כמה לחמניות צריכה המאפיה לאפות כל יום כדי שב-95 אחוז של הימים יהיו מספיק לחמניות לכל דורש? כמה לחמניות צריכה המאפיה להכין כל יום אם המטרה היא כי ב- 99 אחוז של הימים יהיו מספיק לחמניות לכל הדורשים? כמה לחמניות צריכה המאפיה להכין כל יום אם ברצונה לספק את הביקוש ב- 100% של הימים?

נגדיר משתנה המקרי המקבל את הערך1 אם תושב מבקש לחמנייה ביום מסוים ו- 0 אם אינו חפץ בלחמנייה. לפי הנתונים תוחלת המשתנה המקרי היא 0.25 וסטיית התקן היא השורש של 0.25 כפול 0.75 , כלומר 0.433. המשתנה המקרי שהוא מספר התושבים המבקשים לחמנייה ביום מסוים הוא סכום, על כל התושבים, של המשתנים המקריים אותם הגדרנו. סטיית התקן שלו היא $\sqrt{40.000}$ כפול 0.433, כלומר 86.6. מספר התושבים גדול מספיק כדי שנוכל להסיק ממשפט הגבול המרכזי כי התפלגות הביקוש ללחמניות ביום מסוים קרובה להתפלגות של משתנה נורמלי. במשתנה נורמלי 95% של הדגימות נופלות בטווח של 1.96 סטיות תקן מהממוצע (המספר ניתן בכיתה). לכן כדי שב 95% של הימים יסופק הביקוש במלואו מספיק כי מספר הלחמניות ביום יגיע ל10,000 ועוד 1.96 כפול 86.6 כלומר ל- 10,170 לחמניות. אם רוצים לספק את כל הדרישה ב-99% של הימים הטווח צריך לגדול ל- 2.324 סטיות תקן (המספר ניתן בכיתה). אם נכניס מספר זה לחישוב הקודם נקבל כי יספיקו 10,202 לחמניות ליום. אם המאפייה רוצה לספק את הביקוש בכל יום ויום עליה להכין 40.000 לחמניות כל יום (כי, אמנם בהסתברות נמוכה, ייתכן יום בו כל התושבים ירצו לחמנייה).