

ענה על כל השאלות. לרשותך שלוש שעות, אך קח בחשבון כי אין צורך להשתמש בכל הזמן המוקצב. **המבחן לא בודק רק את היכולת להציב נתונים בנוסחה.** תשובה מלאה אמורה להתייחס (בקיצור) **הן למודל שלפיו עובדים והן להנחות** המתבקשות לפני שמשמשים בנוסחה. אם ניסוח של שאלה או חלקה לא ברורים לך, נסח אינטרפרטציה סבירה לנוסח הלא ברור והמשך לפיה. אם השימוש בנוסחה מחייב ידיעה של פרמטר אותו אינך יודע, אתה יכול להשתמש בסימן מתמטי תוך הסבר איזה פרמטר חסר. בכל שאלה הסבר מה אתה עושה. תשובה לקונית, למשל, המספר הנכון בלבד או שימוש בנוסחה מבלי לציין מדוע אפשר להשתמש בה, **תזכה רק בחלק קטן של הניקוד.** וגם: התשובות חייבות לכלול רק חומר הרלבנטי לשאלות. הוספת חומר שאינו קשור ישירות לשאלה **תביא להורדה בניקוד.**

1. במיכל ענקי נמצאים כדורים לבנים ושחורים. הוציאו 100 כדורים באופן אקראי והיו ביניהם 55 כדורים שחורים. באיזו רמת מובהקות יודעים כי במיכל כולו יש בין 50% ל- 60% כדורים שחורים.

תשובה: נשתמש בפיתוח (שנעשה בין השאר בכיתה) האומר שההסתברות שהפרופורציה הנכונה של כדורים שחורים לכלל הכדורים תיפול ברווח בר סמך

$$[X - \alpha S n^{-1/2}, X + \alpha S n^{-1/2}]$$

שווה להסתברות שמשנתה נורמלי סטנדרטי יקבל ערכים קרובים יותר מאשר α סטיות תקן מהמוצע. כאן X הוא הממוצע במדגם, S היא סטיית התקן של המדגם, n גודל המדגם. מהנתונים עולה כי שורש n הוא 10. במונחים של פרופורציות מקבלים כי $X = 0.55$. כדי שהפרופורציה במיכל כולו תיפול בין 0.50 לבין 0.60 צריך כי הסטייה ברווח בר הסמך מהמוצע של המדגם תהיה קטנה מאשר 0.05. חישוב פשוט מראה כי $S = \sqrt{0.55 \cdot 0.45} = 0.497$. כדי לעמוד בתנאים צריך להתקיים $\alpha \cdot 0.497 \cdot 1/10 \leq 0.05$. בקירוב טוב מאוד $\alpha \leq 1$ כלומר סטיית תקן אחת. את ההסתברות שמשנתה נורמלי ייפול בתוך 1 סטיות תקן מהמוצע אפשר למצוא בטבלאות. בכיתה ניתן המספר 0.317 האומר מה הסיכוי ליפול מחוץ לסטיית תקן אחת, או 0.683 ליפול בתוך סטיית תקן 1. כלומר, ברמת המובהקות של קצת יותר מאשר 68 אחוזים, הפרופורציה הנכונה של הכדורים השחורים במיכל נמצאת בגבולות הנתונים. (אנו מסתפקים בתוצאה קרובה מכיוון שהנתונים עצמם מקורבים, למשל סטיית התקן S).

2. מכונית ובה שלושה אנשים: שני גברים ואישה, הייתה מעורבת בתאונה. למשטרה לא ידוע מי הנהג ולכן, מלכתחילה, מעריכה את ההסתברות של כל אחד מהגברים להיות הנהג כ- 0.4 בעוד שההסתברות כי האישה היא הנהגת מוערכת ב- 0.2. בדיקה נוספת של המשטרה העלתה ממצאים נוספים. התגלה דם על ההגה ומסיקים כי הנהג נפצע, וסוג הדם שלו הוא O. סטטיסטיקה מהימנה אומרת כי 10% מהגברים ו- 8% מהנשים הם בעלי סוג דם O אבל גם נודע כי הגבר הראשון מגיע ממשפחה בה הסיכוי לסוג דם O הוא 50%. אנא ייעץ למשטרה כיצד לעדכן את הערכת הסיכויים לגבי כל אחד מהשלושה להיות הנהג?

תשובה: במרחב מדגם מתאים נסמן ב- A , B ו- C את המאורעות: גבר א' נהג, גבר ב' נהג והאישה נהגה. ההערכה המקורית של המשטרה אומרת $p(A) = p(B) = 0.4$ בעוד $p(C) = 0.2$ את האינפורמציה הנוספת, כלומר את המאורע כי התגלה הדם על ההגה, נסמן ב- D . ההסתברויות המעודכנות שמחפשים הם $p(A|D)$, $p(B|D)$ ו- $p(C|D)$. נניח כי האינפורמציה שיש ושנודעת למשטרה משקפת את ההסתברויות האובייקטיביות. לכן כדי לחשב את ההסתברויות המעודכנות אפשר להשתמש בחוקי ההסתברות המותנית, ובפרט באחת מהגרסאות לחוק בייס.

לשם כך: הסיכוי שסוג הדם על ההגה הוא O ניתן על ידי

$$p(D | A)p(A) + p(D | B)p(B) + p(D | C)p(C)$$

וקל לחשב את המספר הזה: $0.5 \cdot 0.4 + 0.1 \cdot 0.4 + 0.08 \cdot 0.2 = 0.256$ בפרט, הסיכוי המעודכן שהגבר הראשון הוא הנהג ניתן על ידי

$$\frac{p(D | A)p(A)}{0.256} = p(A|D)$$

והתוצאה היא 0.78125, חישוב דומה לגבי ההסתברויות המותנות האחרות מוביל לכך שהסיכוי המעודכן לגבי הגבר השני הוא 0.15625, ולגבי האישה 0.0625.

3. א. נתון כי להתפלגות של משתנה מקרי מסוים X יש פונקציית צפיפות שנסמן ב- $f(x)$. מהי פונקציית הצפיפות של המשתנה המקרי $2X+3$?
- ב. נסח חוק גבול מרכזי.
- ג. בניסוח שבסעיף הקודם מופיע מושג ההתכנסות, או הקרבה, בין התפלגויות. הגדר באופן מדויק למה הכוונה הן לגבי התפלגויות כלליות והן לגבי אלה המופיעות במשפט שניסחת.
- ד. הראה כי אי אפשר להשמיט את הנחת אי-התלות בניסוח חוק הגבול המרכזי.

תשובה: א. נסמן את פונקציית הצפיפות של המשתנה החדש ב- h . אז $h(x) = f\left(\frac{x-3}{2}\right)$ (זו תוצאה של שינוי משתנה בפונקציה).

- ב. יהיו X_i עבור $i = 1, 2, 3, \dots$ סדרה של משתנים מקריים בלתי תלויים ושווי התפלגות. נסמן את התוחלת המשותפת שלהם ב- m . נסמן ב- $Z(n)$ את המשתנה המקרי שהוא סכום של X_i עבור $i = 1, 2, \dots, n$. אז ההתפלגות של

$$\frac{Z(n) - n \cdot m}{\sigma \sqrt{n}}$$

מתכנסת כאשר n שואף לאינסוף להתפלגות $N(0,1)$, כלומר להתפלגות נורמלית עם ממוצע 0 ושונות 1.

ג. סדרת ההתפלגויות, או החוקים, H_n שואפת להתפלגות H אם לכל מספר ממשי α בו פונקציית ההצטברות של H רציפה, מתקיים $H_n(x < \alpha) \rightarrow H(x < \alpha)$ כאשר H היא ההתפלגות הנורמלית המופיעה בניסוח המשפט ההתכנסות מתקיימת לכל α מכיוון שפונקציית ההצטברות של H רציפה אז בכל הישר.

- ד. נבחר את כל ה- X_i להיות שווים, למשל לפונקציה $h(x) = x$ על הקטע $[-1/2, 1/2]$. המשתנים המקריים האלה בוודאי שאינם בלתי תלויים. כל ההתפלגויות בביטוי שצריך לשאוף להתפלגות נורמלית הן כפולות של h ובפרט לא שואפות להתפלגות נורמלית.

4. נתונים שני משתנים מקריים X ו- Y שהם שווי התפלגות אבל לא בהכרח בלתי תלויים. נסמן את השונות המשותפת ב- σ^2 . הראה כי השונות של $X + Y$ יכולה לנוע רק בין 0 ובין $4\sigma^2$, ושני ערכים קיצוניים אלה אכן מתקבלים. (רמז: יתכן ותרצה להשתמש באי-שוויון קושי-שוורץ עבור פונקציות אינטגרליות האומר שריבוע האינטגרל של מכפלת שתי פונקציות כאלה הוא קטן או שווה למכפלת האינטגרלים של ריבוע הפונקציות, ובסימול מתמטי $(\int fg)^2 \leq \int f^2 \int g^2$).

תשובה: נסמן ב- m את הממוצע של כל אחד משני המשתנים. השונות של הסכום מוגדרת על ידי התוחלת $E((X+Y-2m)^2)$. ביטוי זה מציידו שווה ל-

$$E((X-m)^2) + E((Y-m)^2) + 2E((X-m)(Y-m))$$

שני הביטויים הראשונים שווים כל אחד ל- σ^2 (לפי הגדרה). את הביטוי הימני אפשר להעריך על ידי אי שוויון קושי-שוורץ על ידי $2\sqrt{E(X-m)^2 E(Y-m)^2}$ כלומר על ידי $2\sigma^2$. מכאן ההערכה של $4\sigma^2$ לגבי השונות של הסכום. ההערכה 0 מלמטה נובעת ישירות מההגדרה הערך 0 לגבי השונות של הסכום מתקבל, למשל, כאשר $Y = -X$. הערך $4\sigma^2$ מתקבל כאשר $X=Y$.

5. אחד משלושה אסירים נבחר באקראי ועתיד להיות מוצא להורג מחר. מנהל בית הסוהר מורה לאחד הסוהרים, שידוע מי אמור להיות מוצא להורג, לבחור בין אסיר ב' ואסיר ג' אחד שלא יוצא להורג ולספר זאת לאסיר א'. הסוהר אומר לאסיר א' כי ב' לא יוצא להורג. האינפורמציה החדשה עשויה לעדכן את סיכויי של א' להיות זה שיוצא להורג. הראה כי עדיין אי אפשר להסיק מסיפור המעשה מה הסיכוי המעודכן של א' להיות זה המוצא להורג למחרת. הראה זאת על ידי כך שתציג לפחות שתי אסטרטגיות פעולה של הסוהר המביאות למסקנות שונות בדבר הסיכוי הנזכר של א'.

תשובה: אסטרטגיה אפשרית של הסוהר היא שאם או ב' או ג' מוצאים להורג למחרת, הוא יצביע לאסיר א' על זה שלא יוצא להורג. אחרת יצביע על אסיר ג'. מכיוון שהצביע על אסיר ב' יכול אסיר א' להסיק ג' יוצא להורג ולא הוא. אסטרטגיה אפשרית אחרת היא שאם או ב' או ג' מוצאים להורג למחרת הסוהר יצביע לאסיר א' על זה שלא יוצא להורג. אחרת יצביע על אסיר ב'. מכיוון שהצביע על אסיר ב' יכול אסיר א' להסיק (מחוק בייס) כי הסיכוי שהוא יוצא להורג מחר היא 50%, את מכיוון שההצבעה על ב' יכלה לנבוע מכך ש-א' הוא המוצא להורג או מכך ש-ג' הוא המוצא להורג, ולשני המאורעות אותו סיכוי. יש כמובן עוד אסטרטגיות שמביאות לתוצאות אף שונות.

6. א. הגדר מהי התפלגות פואסונית.
 ב. הסבר (במדויק) את הקשר שבין מספר גדול של ניסיונות ברנולי שלכל אחד סיכוי הצלחה נמוך לבין התפלגות פואסונית.
 ג. מספר התביעות לשנה מחברת ביטוח הוא משתנה מקרי מפולג פואסונית עם תוחלת λ . גודל התביעה הממוצעת בכל שנה הוא עצמו משתנה מקרי X , שהוא פונקציה, כך אנו מניחים, של מספר התביעות, והניתנת על ידי הגודל a_k עבור $k = 0, 1, 2, 3, \dots$. רשום ביטוי המציג מה היא תוחלת הסכום שחברת הביטוח נתבעת לשלם כל שנה? מה תהיה תוחלת התשלום אם a_k אינו תלוי ב- k , נאמר הגודל קבוע וגדלו c ?
 ד. הסבר כיצד ייתכן כי התפלגות פואסונית יכולה להיות קרובה לנורמלית כאשר בהתפלגות נורמלית ההסתברות של כל אינטרוול $(-\infty, a)$ כאשר a שלילי, היא חיובית, בעוד שהתפלגות פואסון נותנת הסתברות 0 לכל אינטרוול כזה.

- תשובה:** א. התפלגות פואסונית עם פרמטר λ נותנת הסתברות חיובית רק למספרים הטבעיים $0, 1, 2, 3, \dots$ כאשר ההסתברות של המספר k היא $e^{-\lambda} \frac{\lambda^k}{k!}$.
 ב. נסתכל על משתנה מקרי המתקבל על ידי סיכום של n חזרות על ניסיון ברנולי של קבלת 1 ו-0 בהסתברות p ו- $(1-p)$ בהתאמה. כאשר n שואף לאינסוף ו- p שואף ל-0 אבל המכפלה np שואפת לקבוע λ – ההתפלגות של אותו משתנה מקרי שואפת להתפלגות פואסון עם אותו λ .
 ג. מכיוון שגודל התביעה a_k תלוי רק בפרמטר k הרי אם יש k תביעות, תוחלת סכום התביעות היא a_k .
 ד. ותוחלת סכום התביעות הכללי היא הסכום $\sum k a_k e^{-\lambda} \frac{\lambda^k}{k!}$ כאשר $k = 1, 2, 3, \dots$.
 הינו קבוע אפשר "להוציא את c מחוץ לסימן הסכימה" והסכום שנשאר היא התוחלת של המשתנה הפואסוני. כלומר, התוצאה המבוקשת היא $c\lambda$.
 ד. ההתפלגות הפואסונית קרובה לנורמלית כאשר λ גדול מספיק, נאמר גדול מ-25. אז האינטרוול $(-\infty, 0)$ כולו נמצא רחוק יותר מאשר 5 (שהוא שורש של 25) סטיות תקן מהמוצע שהוא λ . ההסתברות שלו אמנם חיובית אבל מסדר גודל של 10^{-7} שהוא מספר קטן, וההסתברות הזו תהיה אפילו קטנה בסדרי גודל כאשר λ גדול יותר. לכן ההתפלגות יכולה להיות קרובה לנורמלית.

בטוח שהצלחתם !