

מבוא להסתברות וסטטיסטיקה

תרגיל 2-מדדי מרכז, מדדי פיזור, עקרון המשלים, עקרון ההכלה וההדחה

15 במרץ 2016

שכיחות	ימים
10	[5 – 7)
18	[7 – 9)
34	[9 – 11)
22	[11 – 13)
16	[13 – 15)
10	[15 – 17)

1. חיי מדף של מוצר מסוים (בימים) היו כדלהלן: מצאו:

(א) ממוצע, חציון, שכיח.

(ב) ציירו היסטוגרמה.

(ג) מה ישתנה בתשובתכם לסעיף א' אם השורה התחתונה היא $+15$ (15 יום או יותר)?

2. סיגל מקבלת מאביה מתנה לחתונה. הוא מציג לה שתי קופסאות זהות: באחת שני יהלומים יקרים ואבן אחת פשוטה, השניה עם שתי אבנים פשוטות ויהלום יקר אחד. סיגל איננה יודעת מה נמצא באיזו קופסה. למען ההגינות, מרשה האב לבתו להוציא אבן אחת מקופסה שנבחרה על ידה באקראי, ואז להחליט אם להשאיר בידה אותה קופסה או לקחת את השניה (אחרי שהחזירה את האבן שבחרה לקופסה שלה). סיגל נוקטת במדיניות הבאה: אם האבן שהוציאה היא יהלום, היא משאירה את הקופסה, אחרת מחליפה אותה. מה הסיכוי שסיגל תקבל 2 יהלומים כמתנת חתונה?

3. בבית המשפט המחוזי בתל אביב עתרו אירגוני זכויות הנשים כנגד אוניברסיטת בר-אילן. ארגוני הזכויות הציגו נתוני קבלה לשנת הלימודי המראים שאחוזי הקבלה של נשים לאוניברסיטה קטנים בעשרות אחוזים מאלו של גברים. אי לכך טענו, שהאוניברסיטה מפלה לרעה נשים. מטעם ההגנה טענו נציגי האוניברסיטה שהדבר משולל יסוד. הם הציגו נתונים המראים בבירור שלכל מחלקה באוניברסיטה אחוז הקבלה של נשים מסך כל הנשים שהגישו מועמדות לקבלה לפקולטה היו גבוהים משמעותית מאלו של הגברים. ידוע ששני הצדדים הציגו נתוני אמת. האם יתכן הדבר? ואם כן, כיצד הדבר יתכן? נמקו והביאו דוגמא פשוטה (רצוי עם מספרים חד ספרתיים) הממחישה את טענתכם.

4. בעולם האמיתי נתונים לעיתים קרובות מכילים "רעש" דהיינו חלק מהדגימות אינן משקפות את הנתונים האמיתיים, לדוגמא, בגלל תקלה במכשיר המדידה.

(א) לפעמים מתעוררת בעיה כאשר משתמשים בממוצע או בשונות כמדדים במקרה שכזה, מדוע? (האם אתם יכולים לחסום את גודל ההטיה מהממוצע או השונות "האמיתית" של הנתונים בהינתן שמדדת דגימת "רעש").

(ב) נגדיר: $MED = Med\{|x_k - Med\{x_k\}|\}$ (חציון ההפרשים המוחלטים מהחציון).

האם הוא מדד לפיזור (מקיים את האקסיומות) דהיינו מקיים סימטריות, הומוגניות ואינוריאנטיות? (הוכיחו טענותיכם).

(ג) נגדיר $T = \frac{1}{n-2m} \sum_{i=m+1}^{n-m} x_i$ כאשר $x_1 \leq x_2 \leq \dots \leq x_n$ מסודרים (דהיינו מוציאים את m הדגימות הקטנות ביותר והגדולות ביותר וממצעים את הנותרים). האם זהו מדד מרכז? הוכיחו טענותיכם.

5. הוכיחו את השוויון הבאה: $\frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 = \left(\frac{1}{n} \sum_{k=1}^n x_k^2 \right) - \bar{x}^2$ כאשר $\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k$

6. כלל ההכלה וההדחה בהסתברות אומר: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ נא להוכיח את הכלל לאיחוד של 3 קבוצות. רמז: חוק הקיבוץ.

7. כמה עובדות מתורת המספרים:

(א) כל מספר טבעי $n \in \mathbb{N}$ גדול מ-1 ניתן להצגה באופן יחיד כמכפלה של מספרים ראשוניים $n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ כאשר $k_i \in \mathbb{N}$ החזקה של המספר הראשוני p_i .

(ב) שני טבעיים נקראים זרים אם אין טבעי שונה מ-1 שמחלק את שניהם. במילים אחרות, אם אין להם גורם ראשוני משותף בפירוק שלהם למכפלות ראשוניים.

עבור $n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ טבעי, חשבו כמה טבעיים קטנים מ- n הם זרים ל- n (התוצאה הינה ביטוי בפרמטרים (m, p_i, n))

רמז: השתמשו בעקרון המשלים ואח"כ בעיקרון ההכלה וההדחה .