סיכום במבוא להסתברות

מאת: הלל בן חנוך

תוכן העניינים

I. מבואות וקומבינטוריקה

סטטיסטיקה תיאורית - מושגים

- **טיפוסי משתנים:**
  - **איכותי (שמי):** קטגוריות ללא סדר (צבע עיניים).
  - **אורדינלי (סודר):** קטגוריות עם סדר טבעי (דרגה).
  - **אינטרוולי (רווחי):** הפרשים בעלי משמעות, אפס שרירותי (טמפ' בצלזיוס).
  - **מנתי (יחס):** יחסים בעלי משמעות, אפס מוחלט (גובה, משקל).
  - **בדיד:** ערכים ספירים (מספר ילדים).
  - **רציף:** ערכים בקטע (טמפרטורה).
- **אוכלוסיה ומדגם:**
  - **אוכלוסיה** היא מכלול הישויות הנחקר.
  - **מדגם** הוא קבוצה חלקית מהאוכלוסיה המשמשת להסקה עליה.
- **מדדי מרכז:** ערך יחיד המתאר את מרכז ההתפלגות.
  - **ממוצע חשבוני:** $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.
  - **חציון (Median):** הערך האמצעי בנתונים ממוינים.
  - **שכיח (Mode):** הערך הנפוץ ביותר.
  - **אמצע הטווח:** $(\max(x_i) + \min(x_i))/2$.
  - **ממוצע גאומטרי:** $\sqrt[n]{x_1 \cdots x_n}$. **הרמוני:** $n/(\sum \frac{1}{x_i})$.
- **מדדי פיזור:** ערך יחיד המתאר את פיזור הנתונים סביב המרכז.
  - **שונות (Variance):** $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = (\frac{1}{n}\sum x_i^2) - (\bar{x})^2$.

- **סטיית תקן:** $s = \sqrt{s^2}$.
- **טווח (Range):** $\max(x_i) - \min(x_i)$.

- **פרדוקס סימפסון:** תופעה בה מגמה המופיעה במספר קבוצות נתונים מתהפכת או נעלמת כאשר הקבוצות משולבות.

קומבינטוריקה - תורת המניה

- **עיקרי הבחירה:** בחירת k פריטים מתוך n.

	עם חשיבות לסדר	ללא חשיבות לסדר
עם החזרה	n^k	$\binom{n+k-1}{k}$
ללא החזרה	$\frac{n!}{(n-k)!}$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

- **מקדם מולטינומי:** מספר הדרכים לחלק n איברים ל-t קבוצות מסומנות בגדלים $n_1, \ldots, n_t$ (כאשר $\sum n_i = n$) הוא:

$$\binom{n}{n_1, \ldots, n_t} = \frac{n!}{n_1! n_2! \cdots n_t!}$$

- **עקרון ההכלה וההדחה:**

$$|A \cup B| = |A| + |B| - |A \cap B|$$

$$|A \cup B \cup C| = |A| + |B| + |C| - (|A \cap B| + |A \cap C| + |B \cap C|) + |A \cap B \cap C|$$

- **הכללה:** לאיחוד של t קבוצות $A_1, \ldots, A_t$:

$$\left|\bigcup_{i=1}^{t} A_i\right| = \sum_{k=1}^{t} (-1)^{k-1} \sum_{|I|=k, I \subseteq \{1..t\}} \left|\bigcap_{j \in I} A_j\right|$$

II. מרחבי הסתברות בדידים

הגדרות יסוד

- **מרחב הסתברות בדיד:** זוג (Ω, P) כאשר Ω קבוצה בת-מניה (מרחב המדגם) ו-$P : \Omega \to [0,1]$ פונקציית הסתברות נקודתית המקיימת $\sum_{\omega\in\Omega} P(\omega) = 1$.
- **מאורע:** כל תת-קבוצה $A \subseteq \Omega$.
- **הסתברות של מאורע:** $P(A) = \sum_{\omega\in A} P(\omega)$.
- **מרחב אחיד:** אם Ω סופית, $|\Omega| = n$, וההתפלגות אחידה, אז לכל $\omega \in \Omega$, $P(\omega) = 1/n$, ולכל מאורע A, $P(A) = \frac{|A|}{n}$.

הסתברות מותנית ואי-תלות

- **הסתברות מותנית:** ההסתברות של מאורע A בהינתן שמאורע B התרחש (כאשר $P(B) > 0$) היא:

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

- **נוסחת ההסתברות השלמה:** אם $\{B_i\}$ היא חלוקה של Ω, אז לכל מאורע A:

$$P(A) = \sum_i P(A|B_i)P(B_i)$$

- **חוק בייס:** מאפשר "להפוך" את ההתניה.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

- **אי-תלות של שני מאורעות:** A, B בלתי תלויים אם $P(A\cap B) = P(A)P(B)$. זה שקול לכך ש-$P(A|B) = P(A)$ (אם $P(B) > 0$).

- **אי-תלות במשותף:** המאורעות $A_1, \ldots, A_n$ בלתי תלויים במשותף אם לכל תת-קבוצת אינדקסים $I \subseteq \{1..n\}$ מתקיים:

$$P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i)$$

III. משתנים מקריים בדידים

הגדרות

- **משתנה מקרי (בדיד):** פונקציה $X : \Omega \to \mathbb{R}$.
- **פונקציית התפלגות (מסת הסתברות):** $p_X(a) = P(X = a)$.
- **משתנה מציין (אינדיקטור):** עבור מאורע A, $X_A = 1$ אם A מתרחש ו-0 אחרת.
- **התפלגות משותפת:** של X, Y היא $p_{X,Y}(a, b) = P(X = a, Y = b)$.
- **התפלגות שולית:** $p_X(a) = \sum_b p_{X,Y}(a, b)$.
- **אי-תלות של משתנים מקריים:** X, Y בלתי תלויים אם $P(X = a, Y = b) = P(X = a)P(Y = b)$ לכל a, b.

תוחלת, שונות ומתאם

- **תוחלת:** $E[X] = \sum_a aP(X = a)$.
  - **לינאריות התוחלת (תמיד נכון):** $E[aX + bY] = aE[X] + bE[Y]$.
  - **LOTUS:** $E[g(X)] = \sum_a g(a)P(X = a)$.
  - **חוק התוחלת החוזרת:** $E[X] = E[E[X|Y]]$.

- **שונות (Variance):** $V(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$.
  - **תכונות:** $V(X) \geq 0$, $V(aX + b) = a^2V(X)$.
  - **סטיית תקן:** $\sigma_X = \sqrt{V(X)}$.
- **שונות משותפת (Covariance):** $\text{Cov}(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$.
  - **שונות של סכום:** $V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X,Y)$.
  - אם X, Y **בלתי מתואמים** ($\text{Cov}(X,Y) = 0$), אז $V(X + Y) = V(X) + V(Y)$.
  - אי-תלות גוררת חוסר מתאם. ההיפך לא תמיד נכון.
- **מקדם מתאם (של פירסון):** $\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$. מתקיים $-1 \leq \rho \leq 1$.

IV. התפלגויות בדידות חשובות

התפלגויות עיקריות

- **ברנולי** $b(p)$: תוצאת ניסוי יחיד עם הסתברות הצלחה p.
  - $P(X = 1) = p$, $P(X = 0) = 1 - p$.
  - $E[X] = p$, $V(X) = p(1 - p)$.
- **בינומית** $\text{Bin}(n,p)$: סופרת הצלחות ב-n ניסויי ברנולי ב"ת.
  - $P(X = k) = \binom{n}{k}p^k(1 - p)^{n-k}$.
  - $E[X] = np$, $V(X) = np(1 - p)$.
- **גאומטרית** $G(p)$: מספר הניסויים עד להצלחה הראשונה.
  - $P(X = k) = (1 - p)^{k-1}p$ עבור $k = 1, 2, \ldots$.

- $E[X]=1/p$, $V(X)=(1-p)/p^2$.
- **תכונת חוסר הזיכרון:** $P(X>k+m|X>m)=P(X>k)$.

- **פואסון** $\mathbf{Poi}(\lambda)$**:** סופרת אירועים נדירים במרחב/זמן. קירוב לבינומית עם n גדול ו-p קטן, כאשר $\lambda=np$.
  - $P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$ עבור $k=0,1,2,...$.
  - $E[X]=\lambda$, $V(X)=\lambda$.

- **היפרגאומטרית** $\mathbf{Hyp}(n;A,N)$**:** מספר ההצלחות (k) במדגם בגודל n ללא החזרה, מאוכלוסיה בגודל N עם A הצלחות.
  - $P(X=k)=\frac{\binom{A}{k}\binom{N-A}{n-k}}{\binom{N}{n}}$.
  - $E[X]=n\frac{A}{N}$, $V(X)=n\frac{A}{N}\frac{N-A}{N}\frac{N-n}{N-1}$.

- **בינומית שלילית** $\mathbf{NB}(r,p)$**:** מספר ה**כישלונות** (k) עד להצלחה ה-r.
  - $P(X=k)=\binom{k+r-1}{k}p^r(1-p)^k$.
  - $E[X]=r(1-p)/p$, $V(X)=r(1-p)/p^2$.

- **מולטינומית** $M(n;p_1,\ldots,p_m)$**:** וקטור $(X_1,\ldots,X_m)$ הסופר כמה פעמים התקבלה כל תוצאה ב-n ניסויים עם m תוצאות אפשריות.
  - $P(X_1=a_1,\ldots,X_m=a_m)=\binom{n}{a_1,\ldots,a_m}p_1^{a_1}\cdots p_m^{a_m}$ כאשר $\sum a_i=n$.

V. מרחב הסתברות כללי

סיגמא-אלגברה ואקסיומות קולמוגורוב

- **הבעייתיות במרחב רציף:** לא ניתן להגדיר הסתברות נקודתית חיובית שתסתכם ל-1, ואי אפשר להגדיר מידה על **כל** תת-הקבוצות שתהיה אינווריאנטית להזזה. הפתרון הוא להגדיר הסתברות רק על אוסף "נחמד" של מאורעות.

- **סיגמא-אלגברה:** אוסף $\mathcal{F}$ של תת-קבוצות של Ω (מאורעות) הסגור תחת:
  1. $\Omega \in \mathcal{F}$.
  2. סגירות למשלים: אם $A \in \mathcal{F}$, אז $A^c \in \mathcal{F}$.
  3. סגירות לאיחוד בן-מניה: אם $A_1, A_2, \dots \in \mathcal{F}$, אז $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

- **אלגברת בורל ($\mathcal{B}$):** הסיגמא-אלגברה הסטנדרטית על $\mathbb{R}$, הנוצרת על ידי כל הקטעים הפתוחים. היא מכילה את כל הקבוצות ה"סבירות" (קטעים, קרניים, נקודות, קבוצות בנות-מניה).

- **מרחב הסתברות כללי:** שלשה $(\Omega, \mathcal{F}, P)$ המקיימת את **אקסיומות קולמוגורוב**:
  1. $P(A) \geq 0$ לכל $A \in \mathcal{F}$.
  2. $P(\Omega) = 1$.
  3. **סיגמא-אדיטיביות:** אם A_i זרים בזוגות, $P(\cup A_i) = \sum P(A_i)$.

VI. משתנים מקריים רציפים

הגדרות למקרה הרציף

- **משתנה מקרי (כללי):** פונקציה $X : \Omega \to \mathbb{R}$ היא **מדידה** אם לכל קבוצת בורל $B \in \mathcal{B}$, התמונה ההפוכה $X^{-1}(B)$ היא מאורע ב-$\mathcal{F}$. (שקול לדרישה $\{X \leq x\} \in \mathcal{F}$ לכל x).

- **פונקציית הצטברות (CDF):** $F_X(x) = P(X \leq x)$.

- תכונות: מונוטונית לא יורדת, גבולות 0 ו-1 באינסוף, רציפה מימין.
- $P(a < X \le b) = F_X(b) - F_X(a)$.
- $P(X = x) = F_X(x) - F_X(x^-)$. אם F_X רציפה, $P(X = x) = 0$.

- **פונקציית צפיפות (PDF):** פונקציה $f_X(x)$ כך ש-$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$.
  - תכונות: $f_X(x) \ge 0$, $\int_{-\infty}^{\infty} f_X(x)dx = 1$.
  - קשר: $f_X(x) = F'_X(x)$ בנקודות גזירות.
  - הסתברות: $P(a \le X \le b) = \int_a^b f_X(x)dx$.

- **תוחלת ושונות (רציף):** מחליפים סכומים באינטגרלים.
  - $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.
  - $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$.
  - $V(X) = E[X^2] - (E[X])^2$.

- **טרנספורמציה של משתנים:** אם $Y = h(X)$ ו-h הפיכה, הצפיפות של Y היא:

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right| = \left. \frac{f_X(x)}{|h'(x)|} \right|_{x=h^{-1}(y)}$$

התפלגויות רציפות חשובות

- **אחידה $U[a, b]$:** בחירת נקודה אקראית בקטע.
  - $f(x) = \frac{1}{b-a}$ עבור $a \le x \le b$.
  - $E[X] = \frac{a+b}{2}$, $V(X) = \frac{(b-a)^2}{12}$.

- **מעריכית $\textbf{Exp}(\lambda)$:** זמן המתנה לאירוע בתהליך פואסוני.

- $f(x) = \lambda e^{-\lambda x}$ עבור $x \geq 0$.
- $E[X] = 1/\lambda$, $V(X) = 1/\lambda^2$.
- **חוסר זיכרון:** $P(X > s+t | X > s) = P(X > t)$.

- **נורמלית (גאוסיינית) $N(\mu, \sigma^2)$:**
  - $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.
  - $E[X] = \mu$, $V(X) = \sigma^2$.
  - **סטנדרטיזציה:** אם $X \sim N(\mu, \sigma^2)$, אז $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.
  - **סכום:** אם X, Y נורמליים וב"ת, גם סכומם נורמלי. $N(\mu_1, \sigma_1^2) + N(\mu_2, \sigma_2^2) \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

- **התפלגויות נגזרות מהנורמלית:** (עבור $Z, Z_i \sim N(0,1)$ iid)
  - **גמא $\Gamma(k, \lambda)$:** הכללה של מעריכית. סכום של k מעריכיות $\text{Exp}(\lambda)$ iid.
  - **חי-בריבוע χ_n^2:** התפלגות של $\sum_{i=1}^n Z_i^2$. מקרה פרטי של גמא.

VII. כלים אנליטיים וחסמים

פונקציות יוצרות

- **מומנטים:**
  - מומנט ה-n: $E[X^n]$.
  - מומנט מרכזי ה-n: $E[(X-\mu)^n]$.

המומנטים הגבוהים משמשים לתיאור צורת ההתפלגות.

- **צידוד (Skewness):** $\gamma_1 = E[(X-\mu)^3]/\sigma^3$. מודד א-סימטריה.

- **גבנוניות (Kurtosis):** $\gamma_2 = E[(X-\mu)^4]/\sigma^4 - 3$. מודד "כבדות זנבות".
- **פונקציה יוצרת מומנטים (MGF):** $M_X(t) = E[e^{tX}]$.
  - קיימת בסביבת 0 עבור משתנים "בעלי זנב קל".
  - **יוצרת מומנטים:** $E[X^n] = M_X^{(n)}(0)$.
  - **סכום של ב"ת:** $M_{X+Y}(t) = M_X(t)M_Y(t)$.
  - **קובעת התפלגות:** אם לשני משתנים אותו MGF אז יש להם אותה התפלגות.
- **פונקציה יוצרת הסתברות (PGF):** למשתנים שלמים אי-שליליים, $G_X(t) = E[t^X] = \sum P(X=k)t^k$.
  - $E[X] = G'_X(1)$, $E[X(X-1)] = G''_X(1)$, וכו'.
  - **סכום של ב"ת:** $G_{X+Y}(t) = G_X(t)G_Y(t)$.

אי-שוויונות ומשפטי גבול

- **אי-שוויון מרקוב:** עבור $X \geq 0$, $P(X \geq a) \leq \frac{E[X]}{a}$.
- **אי-שוויון צ'ביצ'ב:** $P(|X-\mu| \geq k) \leq \frac{V(X)}{k^2} = \frac{\sigma^2}{k^2}$.
- **הלמה של בורל-קנטלי:** עבור סדרת מאורעות A_n:
  - **למה ראשונה:** אם $\sum P(A_n) < \infty$, אז $P(A_n \text{ אינסוף פעמים}) = 0$. (לא דורש אי-תלות).
  - **למה שניה:** אם A_n **בלתי תלויים** ו-$\sum P(A_n) = \infty$, אז $P(A_n \text{ אינסוף פעמים}) = 1$.
- **החוק החלש של המספרים הגדולים (WLLN):** עבור X_i בלתי מתואמים עם תוחלת μ ושונות σ^2, הממוצע $\bar{X}_n = \frac{1}{n}\sum X_i$ מתכנס בהסתברות ל-μ.

$$P(|\bar{X}_n - \mu| \geq \epsilon) \xrightarrow[n\to\infty]{} 0$$

- **החוק החזק של המספרים הגדולים (SLLN):** עבור X_i בלתי תלויים ושווי התפלגות (iid) עם תוחלת μ, הממוצע $\bar{X}_n$ מתכנס כמעט בוודאות ל-μ.

$$P(\lim_{n\to\infty} \bar{X}_n = \mu) = 1$$

- **משפט הגבול המרכזי (CLT):** עבור X_i iid עם תוחלת μ ושונות σ^2, הסכום המנורמל מתכנס בהתפלגות להתפלגות נורמלית סטנדרטית.

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{D} N(0,1)$$

כאשר $S_n = \sum X_i$. זה מאפשר קירוב נורמלי להתפלגויות כמו בינומית ופואסון עבור n גדול.