בס"ד

אוניברסיטת אריאל

שם הקורס: הסתברות למדעי המחשב 1

מס' קורס: 7028410-2

מרצה: פרופ' דן חפץ

העורך: משה חנוקוגלו

תאריך: תשע"ז סמ' ק'

הגדרות בסיסיות

מרחב ההסתברות מורכב משלושה יסודות:

1. קבוצה המסומנת Ω של התוצאות האפשריות של הניסוי המתואר.
(גודל Ω יכול להיות אין סופי אך בן מניה)
2. אוסף של תתי קבוצות (מאורעות) של Ω להן ניתֵן הסתברות.
3. פונקציית הסתברות $p(\text{מאורע A}) \in [0,1]$ נותנת לכל מאורע A מספר בין 0 ל 1 (כולל 0 וכולל 1).

אקסיומות של פונקציית ההסתברות:

- $p(\emptyset) = 0\ ,\ p(\Omega) = 1$.
- אם $A_1, A_2, \ldots$ מאורעות זרים, כלומר שאין לאף זוג של תת קבוצות איבר משותף, אזי $p(\bigcup_i A_i) = \sum_i p(A_i)$.
- $p(A) = \sum_{w \in A} p(w)$

תכונות שנובעות מהאקסיומות:

- אם $A \subseteq B$ אז $p(A) \leq p(B)$
- $p(A^C) = 1 - p(A)$ כאשר A^C זה המשלים של A ‏($A^C = \Omega \backslash A$).
- באופן כללי, $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

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

משפט ההכלה והדחה

גרסה ראשונה למשפט: יהיו $A_1, A_2, \ldots, A_n$ מאורעות במרחב הסתברות (Ω,P) אזי:

$$p\Big(\bigcup_{1}^{n} A_i\Big) = \sum_{\emptyset \neq I \subseteq n}^{n} (-1)^{|I|-1} * p\Big(\bigcap_{i \in I} A_i\Big)$$

גרסה שניה למשפט: יהיו $A_1, A_2, \ldots, A_n$ מאורעות במרחב הסתברות (Ω,P) אזי:

$$p\Big(\bigcap_{1}^{n} {A_i}^c\Big) = \sum_{I \subseteq [n]}^{n} (-1)^{|I|} * p\left(\bigcap_{i \in I} A_i\right)$$

הסתברות מותנה

<u>הגדרה</u>: יהי (Ω,P) מרחב הסתברות ויהי $B \subseteq \Omega$ מאורע בעל הסתברות חיובית, כלומר $p(B) > 0$.

המרחב המותנה ב – B הוא מרחב ההסתברות עם מרחב מדגם Ω ופונקציית הסתברות:

$$P(w|B) = \begin{cases} 0 & w \notin B \\ \dfrac{p(w)}{p(B)} & w \in B \end{cases}$$

הערה: מרחב הסתברות מותנה הוא מרחב הסתברות, ולכן כל מה שהוכח על מרחב הסתברות כללי נכון גם במרחב הסתברות מותנה!

הסבר קצר: הכוונה בהסתברות מותנה היא שאנחנו מחפשים את ההסתברות של מאורע עם תנאי של קיום\אי קיום של מאורע אחר.

חישוב הסתברות מותנה: $P(A|B) = \frac{p(A \cap B)}{P(B)}$

יוצא מכך: $P(A|B) * P(B) = p(A \cap B)$. (לפעמים ההסתברות המותנת תהיה נתונה ונצטרך לחשב את החיתוך)

המשלים של:
$p(A|B) = r$
הוא: $p(A^c|B) = 1 - r$
<u>ולא</u> $p(A|B^c) = 1 - r$

כלל השרשרת: $A_1, A_2, \ldots, A_n$ מאורעות כאשר $p(\bigcap_1^{n-1} A_i) > 0$ אזי:

$$p(\bigcap_{1}^{n} A_i) = p(A_1) * p(A_1|A_2) * p(A_3|A_2 \cap A_1) *** p(A_n|\bigcap_{1}^{n-1} A_i)$$

נוסחת ההסתברות השלמה : יהיו $A, B_1, B_2, \dots, B_n$ מאורעות במרחב הסתברות (Ω,P) כך ש $\bigcap_1^n B_i = \Omega$ וכן $B_i \cap B_j = \emptyset$ לכל $1 \le i,j \le n$ שמקיימים $i \ne j$ אז מתקיים

$$p(A) = \sum_{i=1}^{n} p(B_i) * p(A|B_i)$$

(הערה : כדי להבין טוב את הנושא ממומלץ לראות את הסרטון ״מבוא להסתברות ח - הרצאה 03״ שנמצא ביוטיוב בערך מדקה 23 ועד סוף הסרטון.)

<u>נוסחת בייס :</u> יהיו A,B⊆ Ω מאורעות בעלי הסתברות חיובית, אזי מקיים: $p(A|B) = \frac{p(A)*p(B|A)}{p(B)}$.

אי תלות

<u>הגדרה :</u> נאמר שמאורעות A,B במרחב הסתברות הם בלתי תלויים אם $p(A) * p(B) = p(A \cap B)$.

טענה : מאורעות A,B הם בלתי תלויים ($p(B) > 0$) אם״ם $p(A|B) = p(A)$.

אי תלות מותנת:

<u>הגדרה :</u> יהיו A,B,C מאורעות במרחב הסתברות (Ω,P) נאמר ש B,A בלתי תלויים (ב״ת) בהינתן C אם :

$$p(A \cap B|C) = p(A|C) * p(B|C)$$

הערה : העובדה ש A,B הם ב״ת \ תלויים לא אומרת ב״ת \ תלויים בהינתן C!

הכללה של אי תלות מותנת : יהיו $A_1, A_2, \dots, A_n$ מאורעות במרחב הסתברות (Ω,P) נאמר ש $A_1, A_2, \dots, A_n$ ב״ת אם לכל $\{1, \dots, n\} \in$ I וכן $2 \le |I|$ מתקיים : $p(\bigcap_{i\in I} A_i) = \prod_{i\in I} p(A_i)$.

מקרה פרטי : A,B,C הם ב״ת אם מתקיים **כל** התאים :

1. $p(A) * p(B) = p(A \cap B)$
2. $p(A) * p(C) = p(A \cap C)$
3. $p(B) * p(C) = p(B \cap C)$
4. $p(A) * p(B) * p(C) = p(A \cap B \cap C)$

יהיו $A_1, A_2, \dots, A_n$ מאורעות במרחב הסתברות (Ω,P) נאמר ש $A_1, A_2, \dots, A_n$ ב״ת אם״ם מתקיים : $p(\bigcap_{i=1}^n B_i) = \prod_{i=1}^n p(B_i)$ כאשר ${A_i}^c$ או $B_i = A_i$ לכל $1 \le i \le n$.

משתנים מקריים והפלגותיהן

<u>הגדרה :</u> בהינתן מרחב הסתברות (Ω,P) וקבוצה S משתנה מקרי (מ״מ) הוא $X: \Omega \to S$ (פונקציה זו מקבלת תמיד איבר שלם ממרחב ההסתברות)

הערה : אם S לא כתובה באופן מפורש נניח $S = \mathbb{R}$.

<u>הגדרה :</u> התפלגות על קבוצה S צריכה לקיים 3 תנאים :

1. $\mu: S \to [0,1]$ היא פונקציה.
2. $\{x \in S: \mu(x) \ne 0\}$ (קבוצה זו תקרא התומך של μ) תהיה סופית או בת מניה.
3. $\sum_{x\in S} \mu(x) = 1$

<u>הגדרה :</u> בהינתן מרחב הסתברות (Ω,P) ומ״מ $X: \Omega \to S$ ההתפלגות של X שתסומן μ_X היא התפלגות על S המוגדרת ע״י $\mu_X(z)$ ולפעמיים יסומן $p(X = z)$.

הערה : הכוונה בכתיבת $X = z$ היא לקבוצה $\{\omega \in \Omega: X(\omega) = z\}$.

שים לב : אם x, y משתנים מקריים וכן $x = y$ אז $\mu_X = \mu_Y$ אבל להיפך לא נכון.

רשימת התפלגויות נפוצות:

מס.	שם	פרמטרים	קבוצה S	פונקציה על קבוצה S	סימון
1)	**ברנולי**	$0 \le p \le 1$	$\{0,1\}$	$\mu(1) = p$ $\mu(0) = 1 - p$	$x \sim Ber(p)$
2)	**אחידה**		A – סופית $A \neq \emptyset$	$\mu(t) = \frac{1}{\lvert A \rvert}$	$x \sim U(A)$[1]
3)	**גאומטרית**	$0 < p < 1$	$\mathbb{N}$ (בלי 0)	$\mu(t) = p(1-p)^{k-1}$	$x \sim Geom(p)$
4)	**בינומית**	n, p	$\{0, \dots, n\}$	$\mu(t) = \binom{n}{t} p^t (1-p)^{n-t}$	$x \sim Bin(n,p)$[2]
5)	**פואסון (Poisson)**	$\mathbb{R} \ni \lambda > 0$	$\mathbb{N} \cup \{0\}$	$\mu(t) = e^{-\lambda} * \frac{\lambda^t}{t!}$	$x \sim Geom(x)$
6)	**היפר-גאומטרית**	n, k, N	$\{0, \dots, n\}$[3]	$\mu(t) = \frac{\binom{k}{t}\binom{N-k}{n-t}}{\binom{N}{n}}$	$x \sim hyp(N,k,n)$
7)	**בינומית – שלילית**	$0 < p < 1$ $r \in \mathbb{N}$	$\mathbb{N}$ (בלי 0)[4]	$\mu(t) = \binom{t-1}{r-1} p^r (1-p)^{t-r}$	$x \sim NB(r,p)$[5]

הסברים על ההתפלגויות:

(ההסברים לא מכסים את כל המקרים ויכולים להיות עוד מקרים שימושיים)

1. מתאר ניסוי שהסיכוי להצלחה הוא p.
2. לכל המשתנים יש את אותה התפלגות.
3. נשתמש במקרה שבו יש הסתברות P שמקרה X יתרחש עד לפעם הראשונה שהוא מתרחש.
4. (לדוגמא: הסיכוי לקבל בהטלת מטבע 1 היא P עד לפעם הראשונה שיוצא 1).
5. נשתמש במקרה שנעשו n ניסויים והסיכוי להצליח בניסוי בודד הוא P, אז מה הסיכוי להצליח בk ניסויים מתוך n. (לדוגמא: ההסתברות שיקרו k הוצאות מוצלחות מתוך n "הוצאות" כאשר מה שהוצא **חוזר** ל"ערמה". וכל "הוצאה" יכולה להיות או מוצלחת או לא, בניגוד להסתברות היפר-גאומטרי).
6. ההסתברות שיקרה מאורע מסויים כאשר אנו יודעים שהממוצע להתרחשות מאורע זה הוא λ והוא גדול מ 0.
7. ההסתברות שיקרו k הוצאות מוצלחות מתוך n " הוצאות" כאשר מה שהוצא **לא חוזר** ל"ערמה" שמכילה N פרטים, וכל " הוצאה" יכולה להיות או מוצלחת או לא.
8. הסתברות זו מתארת רצף של נסויי ברנולי, **בלתי תלויים אחד בשני**, כאשר ידוע לנו שהתרחשו נסויים עד שיהיו r כשלונות, וההסתברות לכשלון בניסוי היא p.

<u>הגדרה:</u> משתנים מקריים $x, y: \Omega \to S$ יקראו בלתי תלויים אם לכל $z, w \in S$ מתקיים:

$$p(x = z, y = w) = p(x = z) * p(y = w)$$

התפלגות של משתנים מקריים רבים

<u>הגדרה:</u> בהינתן מרחב הסתברות (Ω, P) ומשתנים מקריים $X, Y: \Omega \to S$ נוכל להתייחס לזוג (X, Y) כמ"מ על S^2 המוגדר ע"י $(X,Y)(w_1, w_2) = (X(w_1), Y(w_2))$ לכן יש לו התפלגות

$\mu(x,y): S^2 \to [0,1]^2$, כמו כן יש את ההתפלגויות הבודדות הן של x והן של y שהן μ_x, μ_y.

$\mu(x,y)$ נקרא ההתפלגות המשותפת של x, y.

ול μ_x, μ_y נקרא ההתפלגויות השוליות של x, y.

---

[1] אם הקבוצה A היא רצף מספרים אזי נסמן לדוגמא, מספרים בין 1 ל n $x \sim U(1,n)$.
[2] $x \sim Ber(p) = x \sim Bin(1,p)$
[3] ובצורה מדויקת: $\{\max(0, n+k-N), \dots, \min(k, N)\}$
[4] ובצורה מדויקת: $\{\mathbb{N} - \{1, \dots, r-1\}\}$
[5] $x \sim Geom(p) = x \sim NB(1,p)$

לפעמים במקרה ש S היא סופית וקבוצה קטנה אז יהיה נוח לסרטט טבלה (דוגמא לקמן) שבא נוכל לראות את ההתפלגות של שני משתנים מקריים.

$y_2 \backslash y_1$	0	1	2	…	$p(y_2)$
0	$p(y_1 = 0, y_2 = 0)$	$p(y_1 = 1, y_2 = 0)$	$p(y_1 = 2, y_2 = 0)$		
1	$p(y_1 = 0, y_2 = 1)$	$p(y_1 = 1, y_2 = 1)$	$p(y_1 = 2, y_2 = 1)$		
2	$p(y_1 = 0, y_2 = 2)$	$p(y_1 = 1, y_2 = 2)$	$p(y_1 = 2, y_2 = 2)$		
⋮					
$p(y_1)$					

כאשר $S \ni 0,1,2,\ldots$.

כדי לקבל את $p(y_1 = i) = \sum_{k=0} p(y_1 = i, y_2 = k)$

וכמו כן אפשר הפוך $p(y_2 = i) = \sum_{k=0} p(y_1 = k, y_2 = i)$.

הערה: כאשר כותבים $p(y_q = i, y_k = j)$ הכוונה היא שמתקיים גם $y_q = i$ וגם $y_k = j$

תוחלת

<u>הגדרה ראשונית</u>: בהינתן מרחב הסתברות (Ω,P) ומ"מ $X: \Omega \to S$ אנו רוצים להגדיר את $\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega) * p(\omega)$

הערה: צריך לדאוג שלא יהיו לטור בעיות התכנסות. אם X הוא מ"מ אי-שלילי כלומר

$\forall\omega \in \Omega X(\omega) \geq 0$ אז הסכום הנ"ל או מתכנס או שווה לאין סוף. אך אין לו בעיות התכנסות.

<u>את כל ההוכחה עם X^+, X^- לא כתבתי. וצריך לעבור עליה.</u>

<u>משפט</u>: יהי מ"מ $X: \Omega \to S$ בעל תוחלת סופית אזי: $\mathbb{E}(X) = \sum_{z\in S} z * p(X = z)$.

הערה: כל שני משתנים מקריים שיש להם את אותה התפלגות יש להם את אותה תוחלת.

שם	תוחלת
ברנולי	p
אחידה	$\frac{a+b}{2}$
גאומטרית	$\frac{1}{p}$
בינומית	np
פואסון (Poisson)	λ
היפר-גאומטרית	$\frac{kn}{N}$
בינומית –שלילית	$\frac{r}{p}$

תכונות של התוחלת:

ליניאריות התוחלת: אם X,Y מ"מ על אותו מרחב הסתברות וכן $c \in \mathbb{R}$.

- $\mathbb{E}(cX) = c\mathbb{E}(X)$
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
- $\mathbb{E}(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i \mathbb{E}(X_i)$

מונוטוניות התוחלת: אם X,Y מ"מ על אותו מרחב הסתברות וכן $p(X \leq Y) = 1$: $\mathbb{E}(X) \leq \mathbb{E}(Y)$

יהי X מ"מ וכן $f: S \longrightarrow [0,\infty]$ אזי: $\mathbb{E}(f(X)) = \sum_{z\in S} f(z) * p(x=z)$.

שונות

<u>הגדרה</u>: יהי X מ"מ בעל תוחלת סופית, אזי השונות (variance) של X היא $Var(X) = \mathbb{E}\left[\left(x - \mathbb{E}(X)\right)^2\right]$.

הערה: השונות של מ"מ תמיד קיימת. (התכן והיא אין-סוף)

תכונות של שונות:

- $Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$.
- לכל $c \in R$ מתקיים: $Var(cX) = c^2 Var(X)$.
- לכל $c \in R$ מתקיים: $Var(X+c) = Var(X)$.
- $Var(ax+by) = a^2 Var(x) + 2abCov(x,y) + b^2 Var(y)$
- אם $X_1, \dots, X_n$ מ"מ ב"ת, אז $Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$
- לכל מ"מ בעל תוחלת סופית מתקיים $Var(X) \geq 0$

שם	שונות
ברנולי	$p*(1-p) \leq \frac{1}{4}$
אחידה	$\frac{(b-a+1)^2-1}{12}$
גאומטרית	$\frac{1-p}{p^2}$
בינומית	$np(1-p)$
פואסון (Poisson)	λ
היפר-גאומטרית	$\frac{kn(N-k)(N-n)}{N^2(N-1)}$
בינומית –שלילית	$\frac{r(1-p)}{p^2}$

אם $X_1, \dots, X_n$ מ"מ בלתי תלויים, אז $Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$.

שונות משותפת

<u>הגדרה</u>: יהיו Y,X מ"מ בעלי תוחלת סופית, אזי השונות שלהם היא:

$$Cov(X,Y) = \mathbb{E}\left[\left(X - \mathbb{E}(X)\right) * \left(Y - \mathbb{E}(Y)\right)\right]$$

הערות:

- אם ל Y,X שונות סופית אזי $Cov(X,Y)$ סופי.
- אם $Cov(X,Y) > 0$ אז נאמר ש Y,X מתואמים חיובית, אם $Cov(X,Y) < 0$ אז נאמר ש Y,X מתואמים שלילית, אחרת נאמר שהם לא מתואמים.

תכונות של שונות משותפת:

- $Cov(X,X) = Var(X)$
- $Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X) * \mathbb{E}(Y)$
- לכל $a,b \in \mathbb{R}$ מתקיים $Cov(aX,bY) = abCov(X,Y)$
- $Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$
- $Cov(X,Y) = Cov(Y,X)$
- אם: $X = \sum_{i=1}^n x_i$ אז: $Var(X) = \sum_{i=1}^n Var(x_i) + 2\sum_{1\leq i<j\leq n} Cov(x_i,x_j)$

אם X,Y מ"מ בלתי תלויים אז הם בלתי מתואמים ודבר זה גורר $Cov(X,Y) = 0$.

שים לב! אם שני מ"מ בלתי תואמים זה <u>לא</u> אומר שהם בלתי תלויים.

אי שוויון מרקוב

יהי X מ"מ אי שלילי, ויהי t מספר ממשי חיובי אזי $p(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$.

אי שוויון צ'בישב

יהי X מ"מ עם תוחלת ושונות סופיים, ויהי t מספר ממשי חיובי אזי $p(|X-\mathbb{E}(X)| \geq t) \leq \frac{Var(X)}{t^2}$.

מקדם המתאם

<u>הגדרה</u> : יהיו Y,X מ"מ בעלי שונות סופית שונה מאפס, אז מקדם המתאם שלהם הוא $\rho(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x)}\sqrt{Var(y)}}$.

תכונות מקדם המתאם :

- $\rho(x,y) = \rho(y,x)$
- $\rho(x+a,y) = \rho(x,y)$,$2a \in \mathbb{R}$
- $\rho(ax,y) = \frac{a}{|a|}\rho(x,y)$
- $\rho(x,x) = 1$
- $\rho(x,y) = 0$ אם"ם $Cov(x,y) = 0$
- $1 \leq \rho(x,y) \leq 1$
- אם $\rho(x,y) = 1$ אז $y = ax + b$ ו $a,b \in \mathbb{R}$ כאשר $a > 0$.
- אם $\rho(x,y) = -1$ אז $y = ax + b$ ו $a,b \in \mathbb{R}$ כאשר $a < 0$.

בהינתן מ"מ X בעל תוחלת ושונות סופיים והשונות שונה מאפס, למ"מ : $\overline{X} = \frac{X-\mathbb{E}(X)}{\sqrt{Var(X)}}$ יש <u>תוחלת 0 ושונות 1</u>.

נוסחת התוחלת השלמה

יהי X,Y מ"מ אזי מתקיים : $\mathbb{E}(X) = \mathbb{E}\big(\mathbb{E}(X|Y)\big) = \sum_y \mathbb{E}(X|Y=y) * p(Y=y)$.

שונות מותנה

עבור מ"מ X ומאורע A בעל הסתברות חיובית : $\text{Var}(X) = \mathbb{E}\big(Var(X|Y)\big) + Var\big(\mathbb{E}(X|Y)\big)$.

טבלת סיכום

מס.	שם	פרמטרים	קבוצה S	פונקציה על קבוצה S	סימון	תוחלת	שונות
1)	**ברנולי**	$0 \leq p \leq 1$	$\{0,1\}$	$\mu(1) = p$ $\mu(0) = 1-p$	$x \sim Ber(p)$	p	$p*(1-p) \leq \frac{1}{4}$
2)	**אחידה**		A – סופית $A \neq \emptyset$	$\mu(t) = \frac{1}{\lvert A \rvert}$	$x \sim U(A)$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
3)	**גאומטרית**	$0 < p < 1$	$\mathbb{N}$ (בלי 0)	$\mu(t) = p(1-p)^{k-1}$	$x \sim Geom(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
4)	**בינומית**	n,p	$\{0,\dots,n\}$	$\mu(t) = \binom{n}{t}p^t(1-p)^{n-t}$	$x \sim Bin(n,p)$	np	$np(1-p)$
5)	**פואסון (Poisson)**	$\mathbb{R} \ni \lambda > 0$	$\mathbb{N}\cup\{0\}$	$\mu(t) = e^{-\lambda} * \frac{\lambda^t}{t!}$	$x \sim Geom(x)$	λ	λ
6)	**היפר-גאומטרית**	n,k,N	$\{0,\dots,n\}$	$\mu(t) = \frac{\binom{k}{t}\binom{N-k}{n-t}}{\binom{N}{n}}$	$x \sim hyp(N,k,n)$	$\frac{kn}{N}$	$\frac{kn(N-k)(N-n)}{N^2(N-1)}$
7)	**בינומית – שלילית**	$0 < p < 1$ $r \in \mathbb{N}$	$\mathbb{N}$ (בלי 0)	$\mu(t) = \binom{t-1}{r-1}p^r(1-p)^{t-r}$	$x \sim NB(r,p)$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$