שיעור 1 - 4/3

א) מרחב הסתברות

הוא זוג סדור (Ω, P) כאשר Ω היא קבוצה סופית או בת מנייה הנקראת מרחב המדגם המכילה את כל התוצאות האפשריות של הניסוי. ו-P היא פונקציה $P:\Omega\to[0,1]$ ממרחב המדגם אל מספר ממשי בין 0 ל-1 הנקראת "פונקציית ההסתברות". כלומר פונקציית ההסתברות מקבלת תוצאה אפשרית ממרחב המדגם ומחזירה את ההסתברות להתרחשות תוצאה זו.

ב) מאורע

מאורע A הוא תת-קבוצה של מרחב המדגם $A\subseteq\Omega$, הנ"ל מספר כלשהו של מרחב המדגם.
כדי לחשב את ההסתברות שיתרחש מאורע A, ניקח את כל התוצאות האפשריות הכלולות ב-A ונציב בפונקציית ההסתברות. נסכום כל תוצאות ונקבל את ההסתברות להתרחשות מאורע A.

$$P(A)=\sum_{w\in A}P(w)$$

ג) תכונות בסיסיות למרחב הסתברות

1) ההסתברות של מאורע ריק היא 0, ואילו ההסתברות של כל המאורעות היא 1. $P(\emptyset)=0$, $P(\Omega)=1$

2) הסתברות של המאורע המשלים של A, המסומן A^c או $\bar{A}$, מ"ל את כל המקרים $A^c=\Omega\setminus A$
לכן ההסתברות שלו היא $P(A^c)=1-P(A)$.

3) אם A,B הם שני מאורעות כך ש- $A\subseteq B$ אז מתקיים: $P(A)\le P(B)$.

4) אם A,B הם מאורעות זרים אז האיחוד שלהם שווה לסכום ההסתברות שלהם: $P(A\cup B)=P(A)+P(B)$.

5) יהי A,B שני מאורעות שאינם בהכרח זרים אז: $P(A\cup B)=P(A)+P(B)-P(A\cap B)$.

6) משני התכונות הקודמות נגיע למסקנה שעבור שני מאורעות A,B תמיד מתקיים: $P(A\cup B)\le P(A)+P(B)$.
או באופן כללי יותר עבור ∞ מאורעות: $P\left(\bigcup_{n=1}^{\infty}A_n\right)\le\sum_{n=1}^{\infty}P(A_n)$.

שיעור 2 - 11/3

א) מרחב הסתברות אחיד

- הוא מרחב הסתברות שבו לכל תוצאה אפשרית $w\in\Omega$ יש את אותה ההסתברות, כך שמתקיים: $p(w)=\frac{1}{|\Omega|}$.
במרחב הסתברות אחיד, כל מאורע $A\subseteq\Omega$ מחושב כך: $P(A)=\frac{|A|}{|\Omega|}$.

שיעור 3 - 18/3

א) <u>עקרון הכלה-הדחה</u> - יהיו $A_1, A_2, \dots, A_n$ מאורעות במרחב הסתברות. האיחוד של כל מאורעות אלו שווה: $P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{\emptyset\neq I\subseteq\{1,\dots,n\}} (-1)^{|I|-1}\cdot P\left(\bigcap_{i\in I} A_i\right)$. כאשר I הוא כל תתי הקבוצה האפשריים של הקבוצה $\{1,2,\dots,n\}$. העיקרון במשפט זה הוא שכאשר אנו מאחדים מאורעות אנו יכולים להוציא ולהכניס איבר שנמצא בכמה מאורעות כמה פעמים שנרצה, כל עוד לבסוף אנו סופרים אותו פעם אחת בלבד.

ניסוח נוסף למשפט זה רק במאורעות משלימים הוא: $P\left(\bigcap_{i=1}^{n} A_i^c\right)=\sum_{I\subseteq\{1,\dots,n\}} (-1)^{|I|}\cdot P\left(\bigcap_{i\in I} A_i\right)$
לפעמים נעדיף להשתמש בניסוח השני. בנוסף נשים לב כי בניסוח השני כולל גם את הקבוצה הריקה $\emptyset$ בשונה מהניסוח הראשון.

ב) <u>דרך פתרון כללית על ידי הכלה-הדחה</u>

(1) קודם נגדיר מאורע A_i עבור $i\in\{1,2,\dots,n\}$, אנו מחפשים איחוד מסוים של מאורעות מסוג זה או את המשלים של חיתוך מסוים של מאורעות מסוג זה.

(2) נבנה את המשוואה שאנו צריכים לחשב. בדר"כ במרחב הסתברות אחיד נקבל:

$$P\left(\bigcap_{i\in I} A_i\right)=\frac{\bigcap_{i\in I} A_i}{|\Omega|}$$

$$P\left(\bigcup_{i\in I} A_i^c\right)=\frac{\bigcup_{i\in I} A_i}{|\Omega|}$$

(3) נשתמש במשפט הכלה-הדחה ובמתמטיקה בדידה כדי להגיע לתשובה הסופית.

שיעור 4 - 25/3 ושיעור 5 - 1/4

א) <u>מרחב הסתברות מותנה</u>

בהינתן מרחב הסתברות (Ω, P) ומאורע B שההסתברות להתרחשותו אינה 0 - $P(B)>0$. נגדיר מרחב הסתברות המותנה ב-B כזוג $(\Omega, P(\cdot|B))$. כלומר, מרחב הסתברות שכל התוצאות האפשריות בו מ-Ω, כאשר ההסתברות לכל תוצאה כזו היא $P(\cdot|B)$. הסימון "$\cdot$" מציין תוצאה שאינה מוגדרת, יכולה להיות תוצאה בודדת-ω או מאורע-A. אם לדוגמא התוצאה שנבדוק היא ω אזי $P(\omega|B)$ מציין את ההסתברות שיתרחש ω כאשר ידוע שהתרחש B. מתקיים: $P(\omega|B)=\begin{cases}0 & \omega\notin B\\ \frac{P(\omega)}{P(B)} & \omega\in B\end{cases}$.
כלומר אם ω לא ב-B אז בוודאי ההסתברות 0, אך אם אינו לא ב-B ההסתברות שיתרחש ω כאשר ידוע שהתרחש B היא $\frac{P(\omega)}{P(B)}$.

<u>הערה</u> - מרחב הסתברות מותנה הוא מרחב הסתברות לכל דבר, ולכן כל מה שלמדנו עד כה, כמו תכונות בסיסיות, אחידות ועקרון הכלה-הדחה חל גם על מרחב הסתברות מותנה. אלא שצריך להוסיף ההתניה.

ב) משפטים במרחב הסתברות מותנה

(1) יהי A,B מאורעות במרחב הסתברות (Ω,P) כך ש-$P(B)>0$, אזי $P(A|B)=\frac{P(A\cap B)}{P(B)}$

מפני שמכל המאורעות ב-A אנו מתעניינים רק באלו ששייכים גם ל-B, שהרי כל המאורעות שהם זרים ל-B ההסתברות שלהם היא 0 כאשר ידוע שהתרחש B.

(2) <u>כלל הכפל</u>: יהיו $A_1,A_2,\dots,A_n$ מאורעות במרחב הסתברות (Ω,P), כאשר $P(\bigcap_{i=1}^{n-1}A_i)>0$

אזי חיתוך כל המאורעות שווה: $P(\bigcap_{i=1}^{n}A_i)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)\cdot\ldots\cdot P(A_n|\bigcap_{i=1}^{n-1}A_i)$.

נשים לב כי אם $P(\bigcap_{i=1}^{n-1}A_i)>0$ אזי כל איחוד של $A_1,A_2,\dots,A_{n-1}$ הוא גדול מ-0 לפי חוק המונוטוניות האומר שאם $A\subseteq B$ אזי בהכרח $P(A)\le P(B)$. אנו צריכים להניח זאת שכל החיתוכים האפשריים האלו גדולים מ-0, מפני שלפי ההגדרה $P(A|B)=\frac{P(A\cap B)}{P(B)}$ חייב להיות ש-$P(B)>0$, ולכן נשים לב שבכל איבר במשוואה שלעיל צריך שהחיתוך יהיה גדול מ-0. $P(\bigcap_{i=1}^{n-1}A_i)>0$ כולל את כולם.

(3) <u>חוק ההסתברות השלמה</u>: יהי A,B מאורעות במרחב הסתברות (Ω,P), כך ש-$1>P(B)>0$,

אזי מתקיים: $P(A)=P(A|B)\cdot P(B)+P(A|B^c)\cdot P(B^c)$. צריך ש-$P(B)$ יהיה בין 0 ל-1 מפני שבמשוואה זו צריך לחלק ב-$P(B)$ וב-$P(B^c)$ שצריכים להיות שונים מ-0. אם $P(B)=1$ אזי $P(B^c)=0$.

חוק זה עוזר לנו כאשר אנו רוצים לחשב מאורע $P(A)$ אך קשה לנו לחשבו לבדו ולכן אנו עושים חלוקה למקרים. לדוגמא כאשר A הינו זוגי וכאשר A הינו אי-זוגי. כמו-כן ניתן להכליל לפי עיקרון חוק זה ליותר מ-2 מקרים, ואז נקבל את נוסחת ההסתברות השלמה עבור מקרה כללי.

המקרה הכללי: יהיו $A,B_1,B_2,\dots,B_n$ מאורעות במרחב הסתברות (Ω,P), כך ש-$P(B_i)>0$ לכל $1\le i\le n$, וגם $B_1,B_2,\dots,B_n$ זרים בזוגות כלומר $B_i\cap B_j=\emptyset$ לכל $i\neq j$, וגם מכסים את מרחב המדגם $\bigcup_{i=1}^{n}B_i=\Omega$, אזי $P(A)=P(A|B_1)\cdot P(B_1)+P(A|B_2)\cdot P(B_2)+\ldots+P(A|B_n)\cdot P(B_n)=\sum_{i=1}^{n}P(A|B_i)\cdot P(B_i)$

(4) <u>נוסחת ב"ס (Bayes Rule)</u>: יהיו A,B מאורעות במרחב הסתברות (Ω,P), כך ש-$P(A)>0$ וגם $P(B)>0$, אזי:

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

משתמשים בכלל זה כאשר נרצה לחשב את $P(B|A)$ אך מסיבה כלשהי זה יהיה קשה לנו, לכן נוכל לחשב את הצד ההפוך $P(A|B)$. במילים אחרות, כאשר יש נתון ואנו רוצים להוכיח מסקנה נוכל להשתמש בנוסחה זו ולהניח את המסקנה ולהסיק את הנתון. נוסחה זו מאוד שימושית במדעי המחשב.

ג) מאורעות בלתי תלויים

הגדרה: יהיו A ו-B מאורעות במרחב הסתברות (Ω,P), נאמר ש-A ו-B בלתי תלויים אם מתקיים: $P(A\cap B)=P(A)\cdot P(B)$. רק בדרך זו ניתן להוכיח שמאורעות בלתי תלויים.

הגדרה 2: יהיו A ו-B מאורעות במרחב הסתברות (Ω,P) כך ש-$P(B)>0$, אזי A ו-B יקראו בלתי תלויים אם מתקיים: $P(A)=P(A|B)$. כלומר התרחשות B אינה משפיעה על התרחשות A.
זו הגדרה פחות פרקטית.

הגדרה 3: מאורעות A ו-B במרחב הסתברות (Ω,P) הם בלתי תלויים אם ורק אם כל השוויונות הבאים מתקיימים:

$$P(A\cap B)=P(A)\cdot P(B),$$
$$P(A\cap B^c)=P(A)\cdot P(B^c).$$
$$P(A^c\cap B)=P(A^c)\cdot P(B).$$
$$P(A^c\cap B^c)=P(A^c)\cap P(B^c).$$

הגדרה זו באה ללמד אותנו שעיקרון הגדרה 1 חל גם על מאורעות משלימים.

ד) אי-תלות עבור יותר מ-2 מאורעות

נגדיר שתי הגדרות שקולות לאי-תלות עבור 3 מאורעות ומעלה:

(1) הגדרה 1: מאורעות $A_1,A_2,\dots,A_n$ יקראו בלתי תלויים (mutually independent) אם לכל תת קבוצה $I\subseteq\{1,2,\dots,n\}$ מגודל לפחות 2 $(|I|\geq 2)$, מתקיים: $P\left(\bigcap_{i\in I}A_i\right)=\prod_{i\in I}P(A_i)$

הגדרה זו נקבל 2^n-n-1 משוואות.

דוגמה עבור $n=3$: מאורעות A,B,C יקראו בלתי תלויים לפי הגדרה זו אם מתקיימים כל ארבעת התנאים:

(1) $P(A\cap B)=P(A)\cdot P(B)$
(2) $P(A\cap C)=P(A)\cdot P(C)$
(3) $P(B\cap C)=P(B)\cdot P(C)$
(4) $P(A\cap B\cap C)=P(A)\cdot P(B)\cdot P(C)$

(2) הגדרה 2: מאורעות $A_1,A_2,\dots,A_n$ יקראו בלתי תלויים אם כל האפשרויות לחיתוך כל n המאורעות עם המשלימים שווה למכפלת המאורעות, כלומר שמתקיים: $P\left(\bigcap_{i=1}^{n}B_i\right)=\prod_{i=1}^{n}P(B_i)$ לכל בחירה של $B_i\in\{A_i,A_i^c\}$.

הגדרה זו נקבל 2^n משוואות.

דוגמה עבור $n=3$: מאורעות A,B,C יקראו בלתי תלויים לפי הגדרה זו אם מתקיימים כל שמונת התנאים:

(1) $P(A\cap B\cap C)=P(A)\cdot P(B)\cdot P(C)$
(2) $P(A\cap B\cap C^c)=P(A)\cdot P(B)\cdot P(C^c)$
(3) $P(A\cap B^c\cap C)=P(A)\cdot P(B^c)\cdot P(C)$
(4) $P(A^c\cap B\cap C)=P(A^c)\cdot P(B)\cdot P(C)$
(5) $P(A\cap B^c\cap C^c)=P(A)\cdot P(B^c)\cdot P(C^c)$
(6) $P(A^c\cap B\cap C^c)=P(A^c)\cdot P(B)\cdot P(C^c)$
(7) $P(A^c\cap B^c\cap C)=P(A^c)\cdot P(B^c)\cdot P(C)$
(8) $P(A^c\cap B^c\cap C^c)=P(A^c)\cdot P(B^c)\cdot P(C^c)$

* שתי ההגדרות שקולות, הוכחה בדפי סיכומים עבור $n=3$.

ה) א׳-תלות מותנה

בהינתן מאורעות A,B,C במרחב הסתברות (Ω,P) כך ש-$0<P(C)$. נוכל לשאול האם A ו-B תלו"מ או בלתי תלו"מ כאשר ידוע שהתרחש C, כלומר במרחב המותנה $(\Omega,P(\cdot|C))$? באופן שקול, נשאל האם מתקיים: $P(A\cap B|C)=P(A|C)\cdot P(B|C)$?

באופן עקרוני לא נוכל להסיק מהתלות או א׳-התלות בין A ל-B במרחב המקורי לשאלה שלנו האם A ו-B תלו"מ או לא תלו"מ במרחב הסתברות המותנה. בסיכומים באנגלית יש שתי דוגמאות שבאחת יש מאורעות תלו"מ במרחב הסתברות המקורי אך לא תלו"מ במרחב הסתברות המותנה. ובדוגמא השניה ההיפך.

מסקנה: התניה במאורע C יכולה להשפיע על התלות/א׳-תלות של A ו-B.

שיעור 6 - 8/4

א) משתנים מקריים

הגדרה: יהי מרחב הסתברות (Ω,P) ותהי S קבוצה. משתנה מקרי הוא פונקציה ממרחב המדגם Ω ל-S, $X:\Omega\to S$. כלומר משתנה מקרי הוא התאמה של אירוע אפשרי במרחב הסתברות לערך מספרי. נסמן משתנה מקרי בדרך כלל באחת האותיות X,Y,Z.

הערה - לרוב לא נכתוב כלל את S ואז נניח ש-$S=\mathbb{R}$ כל המספרים הממשיים. בדרך כלל רק חלק מהמספרים אלו יהיו רלוונטים וכל שאר המספרים ההסתברות להם תהיה 0 או שואפת ל-0.

דוגמאות: (1) משתנה מקרי יכול לציין סכום של שתי זריקות בקוביה.

(2) מספר ההצלחות בניסוי שההסתברות להצלחה בניסוי אחד הוא P.

ב) התפלגות

התפלגות של קבוצה S היא פונקציה $\mu:S\to[0,1]$, המתאימה לכל ערך מהקבוצה S ערך ממשי בין $[0,1]$. הגדרה כללית התפלגות על קבוצה.

התפלגות זו צריכה לקיים שתי תכונות: הערה: גם הדרישה שהתמונה של μ היה בין $[0,1]$ זו תכונה נוספת.

(1) הקבוצה $\{x\in S:\mu(x)\neq 0\}$ היא קבוצה סופית או בת מניה. הקבוצה זו נקראת "התומך של μ".

(2) מתקיים: $\sum_x \mu(x) = 1$.

הגדרת התפלגות משתנה מקרי.

התפלגות של משתנה מקרי $X:\Omega\to S$ היא בעצם פונקציה שמתאימה לכל ערך של המשתנה המקרי את ההסתברות להתרחשות מקרה זה. נסמן ערך כלשהו של המשתנה המקרי X כך: x. x הוא בעצם המאורע $\{\omega\in\Omega:X(\omega)=x\}$. נסמן התפלגות של מ"מ X בערך x בשתי דרכים: $\mu_X(x)=P(X=x)$*.

*נשתמש לרוב בסימון השני

הערה: עבור שתי משתנים מקריים X,Y נאמר כי ל-X ו-Y יש את אותה ההתפלגות רק אם יש ביניהם אותה התאמה מ-Ω ל-S, כלומר $p(X=x)=p(Y=x)$.

ג) התפלגות משותפת

יהי (Ω,p) מרחב הסתברות ויהיו $X,Y:\Omega\to S$ שני משתנים מקריים. נוכל להתייחס לזוג (X,Y) כמשתנה מקרי אחד האקראי $(X,Y):\Omega\to S^2$, כאשר S^2 הוא יחס דו-מקומי על S שאיבריו הם כל הזוגות הסדורים האפשריים של S, כלומר $(X,Y)(w)=(X(w),Y(w))\in S^2$.

ההתפלגות של משתנה מקרי זה מוגדרת כך: $M_{(X,Y)}(x,y)=P(X=x,Y=y)$. הסימן פסיק (,) מציין כאן "וגם", כלומר גם $P(X=x)$ וגם $P(Y=y)$. התפלגות זו נקראת "ההתפלגות המשותפת של X ו-Y" (Joint Distribution), ואילו ההתפלגות של X ו-Y כל אחד בנפרד נקראות "התפלגויות שוליות" (Marginal Distribution).

X\Y	0	1	2
0			
1			
2			

כאשר התומך של ההתפלגות המשותפת $supp(M_{(X,Y)})$ הוא מספר נמוך מאוד ניתן לייצג את ההתפלגות המשותפת בטבלה, כמו הדוגמה משמאל בה התומך $\{0,1,2\}$.

<u>הערה חשובה</u> - כאשר נתון שתי התפלגויות שוליות של שני משתנים מקריים, אי אפשר לקבוע מכך ישירות מה תהיה ההתפלגות המשותפת שלהם מתוך כך, אלא צריך לחשב את כל מקרה לגופו, בעזרת טבלה לדוגמה. בהרצאה ראינו דוגמה שבה Y_2 ו-Y_3 יש את אותה התפלגות $Y_2\sim Y_3$, אך $(Y_1,Y_3)\nsim(Y_1,Y_2)$, כי Y_1 ו-Y_2 תלויים ואילו Y_1 ו-Y_3 לא.

לעומת זאת, מתוך ההתפלגות המשותפת כן ניתן לקבוע את ההתפלגויות השוליות, בכך שעבור כל ערך של אחד מהמשתנים המקריים נסכם את כל האפשרויות במשתנה המקרי השני. כלומר,

$$P(X=x)=\sum_{y\in S}P(X=x,Y=y) \quad . \quad P(Y=y)=\sum_{x\in S}P(X=x,Y=y)$$

לדוגמה, בטבלה לעיל נחשב את $P(X=1)$ ע"י סכימת כל הערכים <u>בשורה</u> של 1, ואת $P(Y=0)$ ע"י סכימת כל הערכים <u>בטור</u> של 0.

בס"ד

שיעורים 7-8 - 15/4, 22/4

יש כל מיני התפלגויות נפוצות שאנו צריך לעבוד איתם, נלמד את המרכזיים שבהם.
חשוב לזכור שלא כל התפלגות היא התפלגות נפוצה, אלא יש התפלגויות כלליות כמו שלמדנו בשבוע שעבר.

א) התפלגות אחידה (Uniform)

תהי S קבוצה סופית כלשהי. נאמר שמשתנה מקרי X מתפלג אחיד על S אם מתקיים:
$P(X=x)=\frac{1}{|S|}$ לכל $x\in S$, כמובן שמתקיים $p(X=x)=0$ לכל $x\notin S$. נסמן ש-X מתפלג אחיד על S כך $X\sim U(S)$.

לרוב לא נתעסק עם קבוצה כללית S, אלא תהיה נתונה קבוצה $S=\{a,a+1,...,b\}$ של שלמים עוקבים, ואז נוכל לסמן באופן פשוט יותר ש-X מתפלג אחיד על קבוצה זו כך $X\sim U(a,b)$.
ואז לכל $x\in S$ מתקיים: $P(X=x)=\frac{1}{b-a}$. באופן כללי ההתפלגות היא: $f_X(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & \text{אחרת}\end{cases}$

דוגמה: יהי X תוצאת הטלה של קוביה הוגנת, אזי $X\sim U(1,6)$.

ב) התפלגות ברנולי (Bernoulli)

X מתפלג ברנולי עם פרמטר p ונסמן $X\sim Ber(p)$ אם מתקיים: $p(X=1)=p$ ו- $p(X=0)=1-p$.
כלומר בהתפלגות זו יש שתי תוצאות אפשריות, שאחת מהם תקבל את הערך 1 עם הסתברות P, והשנייה את הערך 0 עם ההסתברות $1-P$.

דוגמה: יהי A מאורע כך ש-$P(A)=p$, אזי האינדיקטור של A, המסומן 1_A, המקבל את הערך 1 אם A התרחש ואת הערך 0 אם לא התרחש, מתפלג ברנולי. $1_A\sim Ber(p)$.

ג) התפלגות בינומית (Binomial)

יהי n שלם חיובי ויהי $0\le p\le 1$ ממש. נאמר ש-X מתפלג בינומית עם פרמטרים n ו-p, ונסמן $X\sim Bin(n,p)$, אם לכל $0\le k\le n$ מתקיים: $P(X=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}$. התפלגות זו בעצם מתארת מקרה של n ניסויים בלתי תלויים, שההסתברות להצלחה בניסוי בודד היא P, והמשתנה המקרי X סופר את מספר ההצלחות $1\le K\le n$. נוכיח שהתפלגות זו אכן התפלגות:

(1) n הוא מספר סופי לכן גם התומך הוא סופי.

(2) סכום ההסתברויות שווה ל-1 - ניתן לראות שמתקיים לפי נוסחת הבינום: $\sum_{k=0}^{n}\binom{n}{k}\cdot p^k\cdot p^{n-k}=[p+(1-p)]^n=1$

(3) כל הסתברות היא בין $[0,1]$ - ניתן לראות שכל הסתברות $p(X=k)\ge 0$ לכל k, כי כל האיברים במכפלה הם חיוביים: ומכיוון שסכום כל ההסתברויות שווה ל-1, כל הסתברות חייבת להיות בין 0 ל-1.

הקשר בין התפלגות בינומי להתפלגות ברנולי - יהיו $X_1,\dots,X_n$ משתנים מקריים בלת״ם כך שלכל $1\le i\le n$ X_i מתפלג ברנולי $X_i\sim Ber(p)$, אזי הסכום $S_n=X_1+\dots+X_n$ מתפלג בינומי $S_n\sim Bin(n,p)$.

לפי זה נשים לב כי ההתפלגות $Bin(1,p)$ היא פשוט $Ber(p)$.

3) התפלגות גיאומטרית

יהי $0<p<1$ מספר ממשי. נאמר שמשתנה מקרי X מתפלג גאומטרית עם פרמטר p, ונסמן $X\sim Geom(p)$, אם לכל $k\in\mathbb{N}$ (בלי 0) מתקיים: $P(X=k)=p\cdot(1-p)^{k-1}$. התפלגות גאומטרית בעצם מתארת מקרה שבו נערכים ניסויים ברנולים בלת״ם, שההסתברות להצלחה בניסוי בודד היא p, והמשתנה מקרי X סופר את מספר הניסויים עד הצלחה ראשונה כולל. קיימת תכונת חוסר זיכרון כשידוע שכבר היו k כישלונות.

ה) התפלגות היפר-גאומטרית

יהיו n,D,N שלמים חיוביים. נאמר שמשתנה מקרי X מתפלג היפר גאומטרית עם פרמטרים n,D,N ונסמן $X\sim Hyp(N,D,n)$, אם לכל $0\le k\le n$ מתקיים:
$$P(X=k)=\frac{\binom{D}{k}\cdot\binom{N-D}{n-k}}{\binom{N}{n}}$$

הערה: פורמלית התומך הוא $\{\max\{0,n+D-N\},\dots,\min\{n,D\}\}$, אמנם בפועל נתייחס לתומך כאל $0\le k\le n$ מכיוון שגם עבור ערכים שאינם בקבוצה הפורמלית נקבל הסתברות 0, שאכן מורה שמקרה זה לא יכול לקרות ולכן לא נתחשב בקבוצה הפורמלית.

התפלגות היפר גאומטרית בעצם מתארת מקרה שישנה אוכלוסיה של N פריטים המחולקים לשתי קבוצות: קבוצה D וקבוצה $N-D$. בוחרים מדגם אקראי ללא חזרה של n פריטים מתוך ה-N, כלומר הבחירות תלויות, המשתנה המקרי סופר כמה פריטים הם מקבוצה D מתוך ה-n במדגם.

הערה - אם המדגם של n נבחר עם חזרה, כך שהבחירות אינן תלויות, אזי בכל אחת מ-n הבחירות ההסתברות להוציא פריט מסוג D היא $\frac{D}{N}$. נקבל שמקרה זה מתפלג בינומי - $X\sim Bin(n,\frac{D}{N})$.

אם n קטן מאוד ביחס ל-N ו-D אזי ההתפלגות הבינומית $Bin(n,\frac{D}{N})$ היא קירוב טוב להתפלגות היפר גאומטרית $Hyp(N,D,n)$, כי ההסתברות שנבחר את אותו פריט פעמיים היא נמוכה. ולכן במקום לחשב את נוסחת ההתפלגות ההיפר גאומטרית נוכל לחשב את נוסחת התפלגות הבינומית שהיא נוחה יותר.

1) התפלגות פואסון

יהי $\lambda>0$ מספר ממשי. נאמר שמשתנה מקרי X מתפלג פואסון עם פרמטר λ, ונסמן $X\sim Poi(\lambda)$ אם לכל $k\in\mathbb{N}\cup\{0\}$ מתקיים: $P(X=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$.

התפלגות פואסון שימושית כאשר יש הרבה מאורעות בלתי תלויים (או כמעט בלתי תלויים, בקורס זה אין התייחסות למושג זה), נסמן את כמות המאורעות n. ידוע לנו בפי' כלשהו (תוחלת) שמספר מהאירועים האלו יתקיים, נסמנו ב-λ. אזי התפלגות מספר המאורעות שיקרו מתקרבת להתפלגות פואסון ככל שמספר המאורעות n יגדל. אם n שואף ל-∞ נקבל בדיוק התפלגות פואסון. זהו בדיוק

<u>משפט הגבול הפואסוני</u> - תהי $\{p_n\}_{n=1}^{\infty}$ סדרה של הסתברויות (מספרים בין 0 ל-1), כך ש- $\lim_{n\to\infty} n\cdot p_n=\lambda$ לאיזשהו מספר $\lambda>0$ ממשי. אזי ההתפלגות הבינומית $Bin(n,p_n)$ שואפת להתפלגות $Pois(\lambda)$ כש-n שואף ל-∞. כלומר, לכל $k\in\mathbb{N}\cup\{0\}$ מתקיים:

$$\lim_{n\to\infty}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}=e^{-\lambda}\cdot\frac{\lambda^k}{k!}.$$

אנו נשתמש בהתפלגות פואסון כאשר נרצה לבחון התפלגות של התרחשות מספר אירועים ביחידת זמן, האירועים בלתי תלויים, ואנו יודעים מהו קצב ההתרחשויות הממוצע ביחידת זמן, שהוא λ. במקרה זה נאמר כי $X\sim Pois(\lambda)$.

2) התפלגות בינומית שלילית

יהי r מספר טבעי ויהי $0\le p\le 1$ ממשי. נאמר שמשתנה מקרי X מתפלג בינומית שלילית עם פרמטרים r ו-p, ונסמן $X\sim NB(r,p)$ אם לכל $r\ge n$ טבעי מתקיים: $P(X=n)=\binom{n-1}{r-1}\cdot p^r\cdot(1-p)^{n-r}$

התפלגות זו מתאימה למקרה שבו יש מספר כלשהו של ניסויים ברנולי (הצלחה וכישלון) בלתי תלויים, ואנו מסמנים X משתנה מקרי הסופר את מספר הניסויים עד r הצלחות כולל. במילים אחרות $P(X=n)$ הוא ההסתברות שנדרשו n ניסיונות כדי להצליח r פעמים. נשים לב שבניסוי ה-n בוודאי הייתה הצלחה, וב-$n-1$ הניסויים הקודמים היו $r-1$ הצלחות ו-$n-r$ כישלונות. לכן נקבל את הנוסחא: $P(X=n)=\binom{n-1}{r-1}\cdot p^r\cdot(1-p)^{n-r}$.

עוד נשים לב שעבור $r=1$, כלומר עד הצלחה ראשונה נקבל התפלגות גיאומטרית. ניתן להסתכל על התפלגות בינומית שלילית כאוסף של r משתנים מקריים גיאומטריים $X_1,X_2,\dots,X_r$ כאשר לכל $1\le i\le r$ המשתנה המקרי $X_i\sim G(p)$ מציין את מספר הניסיונות מאז ההצלחה ה-$i-1$ (לא כולל) עד ההצלחה ה-i (כולל).

<u>טענה:</u> יהי r מספר טבעי ויהי $0<p<1$ ממשי, ויהיו $X_1,X_2,\dots,X_r$ משתנים מקריים בלתי תלויים כך ש $X_i\sim G(p)$ לכל $1\le i\le n$, ומתקיים: $X=X_1+X_2+\dots+X_r$, אזי $X\sim NB(r,p)$. כלומר X מתפלג בינומית שלילית עם פרמטרים r,p.

שיעור 9 - 29/4

א) תוחלת

יהי $X:\Omega\to S$ משתנה מקרי אי-שלילי, כלומר $X(\omega)\geq 0$ לכל $\omega\in\Omega$ כך שכל ערכי S גדולים או שווים ל-0, אזי "התוחלת" של X, המסומנת $E(X)$, תחושב כך: $E(X)=\sum_{x\in S} x\cdot P(X=x)$. התוחלת עשויה להיות גם שווה ל-∞. במילים אחרות, מכפילים כל ערך במשתנה המקרי בהסתברות שלו ואז סוכמים את כל הערכים. התוחלת היא בעצם ממוצע הערכים אותם יכול המשתנה המקרי לקבל, ופירושה הוא ערך נצפה (Expected Value), משום שאם נבדוק אותו ניסוי שהמשתנה המקרי מייצג פעמים רבות ממוצע הערכים יהיה התוחלת.

הערות: (1) התוחלת לא חייבת להיות אחת מן התוצאות האפשריות, לדוגמא תוחלת של קוביה היא 3.5.

(2) ניתן לראות מהנוסחא שהתוחלת של משתנה מקרי תלויה אך ורק בהתפלגות שלו, כלומר בפונקציה $\mu: S\to[0,1]$, ולא בכל ערכי המשתנה המקרי $X(\omega)$ כאשר $\omega\in\Omega$, מפני שישנם ערכי ω עבורם ההסתברות היא 0 ($P(\omega)=0$), ואז אינם משפיעים על התוחלת. המסקנה היא שאם נדע כיצד מ"מ מתפלג בלבד נוכל לחשב את התוחלת שלו, לכן נוכל לחשב נוסחא עבור כל מ"מ שמתפלג כמו אחת משפחות ההתפלגויות המיוחדות.

ב) תוחלת התפלגויות מיוחדות

(1) התפלגות אחידה - יהי משתנה מקרי $X\sim U(a,b)$ עבור a,b שלמים אי שליליים. אזי: $E(X)=\frac{a+b}{2}$.

(2) התפלגות ברנולי - יהי משתנה מקרי $X\sim Ber(P)$, אזי $E(X)=P$. המסקנה העולה מכך היא שלכל מאורע A, האינדיקטור של A, המסומן 1_A ומתפלג ברנולי $1_A\sim Ber(P(A))$, התוחלת שלו הוא: $E(1_A)=P(A)$. נשתמש בכך בהמשך.

(3) התפלגות בינומית - יהי משתנה מקרי $X\sim Bin(n,P)$, אזי $E(X)=n\cdot P$.

(4) התפלגות גיאומטרית - יהי משתנה מקרי $X\sim Geom(P)$, אזי $E(X)=\frac{1}{P}$.

(5) התפלגות היפר-גיאומטרית - יהי משתנה מקרי $X\sim HyP(N,D,n)$, אזי $E(X)=n\cdot\frac{D}{N}$.

(6) התפלגות פואסון - יהי משתנה מקרי $X\sim Pois(\lambda)$, אזי $E(X)=\lambda$.

(7) התפלגות בינומית שלילית - יהי משתנה מקרי $X\sim NB(r,P)$, אזי $E(X)=\frac{r}{P}$.

ג) ליניאריות התוחלת

תכונות חשובות של תוחלת:

יהיו $X,Y:\Omega\to\mathbb{R}$ שני משתנים מקריים אי-שליליים בעלי תוחלת סופית, ויהי $c>0$ מספר ממשי, אזי מתקיים:

(1) הומוגניות - $E(c\cdot X)=c\cdot E(X)$. תוחלת של מ"מ $c\cdot X$, שהוא מכפלה של מ"מ X, הוא: $c\cdot E(X)$.

(2) ליניאריות - $E(X+Y)=E(X)+E(Y)$. תוחלת של מ"מ $X+Y$ שהוא חיבור בין שני מ"מ X,Y הוא: $E(X)+E(Y)$.

כללים אלו נכונים גם כשמחברים משתנים מקריים רבים, לרוב נשתמש בהם בדרך זו.
כלומר אם יש לנו משתנה מקרי X שמאוד קשה לנו לחשב את התוחלת שלו, אם ניתן לפרק אותו ל- $X_1, X_2, \dots, X_n$ משתנים מקריים פשוטים יותר כך שמתקיים $X=\sum_{i=1}^{n} X_i$. אזי נוכל לחשב את התוחלת של X לפי משפט שלמדנו בדרך הבאה: $E(X)=E\left(\sum_{i=1}^{n} X_i\right)=\sum_{i=1}^{n} E(X_i)$, כלומר סכום התוחלות

ג) מונוטוניות התוחלות

יהיו $X,Y:\Omega\to S$ שני משתנים מקריים אי-שליליים. אזי
(1) אם מתקיים $P(X\le Y)=1$, כלומר לכל ערך $w\in\Omega$ שנדגיש בהם תמיד $X(w)\le Y(w)$, אזי $E(X)\le E(Y)$.
(2) מתקיים $E(X)=E(Y)$. אם ורק אם $P(X=Y)=1$, כלומר לכל $w\in\Omega$ שנדגיש בהם מתקיים $X(w)=Y(w)$.

ה) תוחלת של פונקציה של משתנה מקרי

פונקציה של משתנה מקרי- יהי משתנה מקרי $X:\Omega\to S$, ותהי פונקציה $f:S\to\mathbb{R}$, אזי $f(X)$ גם הוא משתנה מקרי $f(X):\Omega\to\mathbb{R}$. נרצה לחשב תוחלת של משתנה מקרי כזה כאשר ידוע לנו את התוחלת וההתפלגות של X. נחלק לשני מקרים:
(1) אם f היא פונקציה לינארית, כלומר מוגדרת $f(X)=a\cdot X+b$ כאשר $a,b\in\mathbb{R}$. אזי התוחלת של $f(X)$ תחושב כך: $E(f(X))=f(E(X))=a\cdot E(X)+b$. טענה זו נובעת מתכונות התוחלת שלמדנו בסעיף ג.
(2) אם f היא פונקציה אי-שלילית $f:S\to[0,\infty)$, אפילו אינה לינארית, אזי נוכל לחשב את התוחלת של $f(X)$ תחושב לפי הנוסחה: $E(f(X))=\sum_{x} f(x)\cdot P(X=x)$. כלומר עוברים על כל הערכים בתחום של X וסוכמים את המכפלה של $f(x)$ בהסתברות שמ"מ X יקבל את הערך x.

שיעור 10 - 6/5

ו) שונות (Variance)

יהי משתנה מקרי $X:\Omega\to S$ בעל תוחלת סופית $E(X)$, השונות של X, המסומנת $Var(X)$ מוגדרת להיות: $Var(X)=E\left((X-E(X))^2\right)=\sum_{x\in S}(x-E(X))^2\cdot P(X=x)$. השונות היא ערך חיובי שנותן לנו מידע על פיזור הערכים של המשתנה המקרי X. אנו מחשבים זאת על ידי סכימת המרחק של כל ערכי המשתנה המקרי מהתוחלת, אמנם סכום כזה תמיד יצא 0 כי יש איברים בסכום שהם חיוביים ויש שליליים, לכן מעלים כל איבר בריבוע כדי שכולם יהיו חיוביים. אם השונות גדולה מראה על פיזור גדול, ואם קטנה מראה על פיזור קטן.

השונות מוגדרת לכל משתנה מקרי עם תוחלת סופית, אמנם השונות עצמה יכולה להיות אינסופית.

ב) סט"ת תקן

יהי X משתנה מקרי עם שונות $Var(X)$, אזי הסט"ת תקן (Standard deviation) של X הוא השורש של השונות $\sigma_X=\sqrt{Var(X)}$. בסט"ת תקן אנו מחזירים את המרחק הממוצע של ערכי המשתנה המקרי מהתוחלת שחישבנו בשונות על ידי העלאה בריבוע, אל המרחק לגודלו המקורי על ידי פעולת שורש על השונות.

ג) תכונות בסיסיות של שונות

יהי X משתנה מקרי עם תוחלת סופית שנסמנה ב-μ. תכונות השונות הן:

(1) השונות תמיד גדולה או שווה ל-0, $Var(X)\geq 0$.

(2) השונות שווה ל-0 אם ורק אם כל ערכי המשתנה המקרי שווים לתוחלת μ, $P(X=\mu)=1$.

(3) לכל $a\in\mathbb{R}$ מתקיים: $Var(a\cdot X)=a^2\cdot Var(X)$. שונות של מ"מ $a\cdot X$ שהוא מכפלה של מ"מ X שווה ל-$a^2\cdot Var(X)$.

(4) לכל $b\in\mathbb{R}$ מתקיים: $Var(X+b)=Var(X)$. שונות של מ"מ $X+b$ שווה לשונות של X, מפני שכשמעלים את כל ערכי המשתנה המקרי ב-b גם התוחלת עולה ב-b, ואז המרחק מהתוחלת נשאר כשהיה.

מסקנה: $Var(aX+b)=a^2Var(X)$

(5) מתקיים: $Var(X)=E(X^2)-(E(X))^2$, כאשר $E(X^2)=\sum_{x\in S}x^2\cdot P(X=x)$. כמעט תמיד נשתמש בתכונה זו לחישוב השונות מפני שהיא מהירה הרבה יותר. נשים לב כי $E(X)$ ח"ב להיות סופי אך $E(X^2)$ יכול להיות אינסופי.

ד) שונויות של התפלגויות מיוחדות

(1) התפלגות ברנולי - יהי משתנה מקרי $X\sim Ber(p)$, אזי $Var(X)=p(1-p)$.

(2) התפלגות אחידה - יהי משתנה מקרי $X\sim U(a,b)$, אזי $Var(X)=\frac{(b-a+1)^2-1}{12}$.

(3) התפלגות בינומית - יהי משתנה מקרי $X\sim Bin(n,p)$, אזי $Var(X)=np(1-p)$.

(4) התפלגות גיאומטרית - יהי משתנה מקרי $X\sim Geom(p)$, אזי $Var(X)=\frac{1-p}{p^2}$

(5) התפלגות היפר גאומטרית - יהי משתנה מקרי $X\sim HyP(N,D,n)$, אזי $Var(X)=\frac{D\cdot n(N-D)(N-n)}{N^2(N-1)}$.

(6) התפלגות בינומית שלילית - יהי משתנה מקרי $X\sim NB(r,p)$, אזי $Var(X)=\frac{r(1-p)}{p^2}$.

(7) התפלגות פואסון - יהי משתנה מקרי $X\sim Pois(\lambda)$, אזי $Var(X)=\lambda$.

שיעור 11 - 13/5

א) שונות משותפת -- Covariance

<u>הגדרה</u>: יהיו X ו-Y משתנים מקריים בעלי תוחלות סופיות, השונות המשותפת של X ו-Y, המסומנת $Cov(X,Y)$ מוגדרת להיות: $Cov(X,Y)=E[(X-E(X))\cdot(Y-E(Y))]$. נשים לב כי $Cov(X,X)=Var(X)$.

<u>טענה</u>: יהיו X ו-Y משתנים מקריים בעלי שונויות סופיות (ולכן גם בעלי תוחלות סופיות), אז גם השונות המשותפת שלהם $Cov(X,Y)$ סופי ומתקיים: $Cov(X,Y)=E(X\cdot Y)-E(X)\cdot E(Y)$. זוהי בעצם עוד דרך נוחה יותר לחישוב השונות המשותפת.

השונות המשותפת היא מדד לקשר בין שני משתנים מקריים X ו-Y בעלי תוחלות סופיות. אם השונות המשותפת חיובית $Cov(X,Y)>0$ נאמר ש-X ו-Y מתואמים חיובית (Positively correlated), מכיוון שערכי המשתנים המקריים נוטים מהתוחלת באותו כיוון (שניהם מעליו או שניהם מתחתיו באותו מאורע). אם השונות המשותפת שלילית $Cov(X,Y)<0$ נאמר ש-X ו-Y מתואמים שלילית, מכיוון שערכי המשתנים המקריים נוטים מהתוחלת בכיוונים מנוגדים. אם השונות המשותפת שווה ל-0 $Cov(X,Y)=0$, נאמר ש-X ו-Y בלתי מתואמים (uncorrelated). אין איזשהו סדר ביניהם בסט"ה מהתוחלת.

<u>טענה</u>: אם X ו-Y בלתי תלויים, אזי הם בלתי מתואמים $Cov(X,Y)=0$. אמנם ההיפך לא נכון, יכול להיות ששני משתנים מקריים בלתי מתואמים אך הם תלויים.

ב) תכונות שונות משותפת

יהיו X,Y,Z משתנים מקריים, ויהיו a,b מספרים ממשיים $a,b\in\mathbb{R}$. אזי מתקיים:

(1) $Cov(X,Y)=Cov(Y,X)$.

(2) $Cov(a\cdot X,\ b\cdot Y)=ab\cdot Cov(X,Y)$.

(3) $Cov(X+Z,Y)=Cov(X,Y)+Cov(Z,Y)$.

ג) שונות של סכום משתנים מקריים

<u>משפט עבור שני מ"מ</u>: יהיו X ו-Y משתנים מקריים עם שונויות סופיות, אזי שונות הסכום שלהם תחושב: $Var(X+Y)=Var(X)+Var(Y)+2\cdot Cov(X,Y)$.

<u>משפט עבור n מ"מ</u>: יהיו $X_1,X_2,\dots,X_n$ משתנים מקריים בעלי שונויות סופיות, אזי שונות הסכום של כל n המשתנים המקריים תחושב: $Var\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}Var(X_i)+2\cdot\sum_{1\le i<j\le n}Cov(X_i,X_j)$. כלומר סוכמים את שונות כל המשתנים המקריים ועוד פעמיים סכום השונות המשותפת של כל זוג משתנים מקריים מתוך ה-n.

המסקנה היא שאם $X_1,X_2,\dots,X_n$ בלתי תלויים בזוגות, אזי $Var\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}Var(X_i)$, מכיוון שהשונות המשותפת של כל זוג מ"מ שווה ל-0.

שיעור 12 - 20/5

א) מקדם המתאם - Correlation Coefficient

הגדרה: יהיו X ו-Y שני משתנים מקריים בעלי שונות סופיות ושונות מ-0 (כלומר ערכי המ״מ אינם קבועים). אזי מקדם המתאם של X ו-Y הוא $\rho(X,Y)=\frac{cov(X,Y)}{\sqrt{Var(X)\cdot Var(Y)}}$.

מקדם המתאם הוא מדד לקשר ליניארי בין שני משתנים מקריים. ערכו של מקדם המתאם הוא בין 1 ל-1-, אם המקדם קרוב לערכים הקיצוניים 1 או 1- זהו אות לכך שיש קשר ליניארי חזק בין שני המשתנים המקריים, אך אם קרוב ל-0 הקשר הליניארי ביניהם חלש, אי שאין קשר כלל כאשר מקדם המתאם שווה ל-0, כלומר X ו-Y בלתי מתואמים.

ב) תכונות מקדם המתאם

(1) $\rho(X,Y)=\rho(Y,X)$

(2) $\rho(X+a,Y)=\rho(X,Y)$, עבור כל $a\in\mathbb{R}$.

(3) $\rho(a\cdot X,Y)=\frac{a}{|a|}\cdot\rho(X,Y)$, עבור כל $a\in\mathbb{R}\setminus\{0\}$.

(4) $\rho(X,X)=1$.

(5) $\rho(X,Y)=0$ אם ורק אם $cov(X,Y)=0$. כלומר X ו-Y בלתי מתואמים.

(6) $|\rho(X,Y)|\le 1$, כלומר תמיד מקדם המתאם בין 1 ל-1-.

(7) מתקיים $\rho(X,Y)=1$ אם ורק אם קיימים שני מספרים ממשיים $a,b\in\mathbb{R}$ ו-$a>0$, כך ש-$Y=aX+b$ כלומר קיים קשר ליניארי מושלם בין X ל-Y. דרך נוספת להסתכל על היחס בין X ל-Y: $P(Y=aX+b)=1$

(8) מתקיים $\rho(X,Y)=-1$ אם ורק אם קיימים שני מספרים ממשיים $a,b\in\mathbb{R}$ ו-$a<0$ כך ש-$Y=aX+b$ כך ששוב קיים קשר ליניארי מושלם בין X ל-Y. דרך נוספת להסתכל על היחס בין X ל-Y: $P(Y=aX+b)=1$.

ג) מבוא לאי-שוויוני ריכוז

כאשר לא ידוע לנו ההתפלגות של משתנה מקרי אך כן ידוע לנו התוחלת והשונות שלו, נוכל פחות או יותר לדעת כיצד נראית ההתפלגות של משתנה מקרי זה בעזרת המשפטים הבאים:

משפט אי-שוויון מרקוב - יהי X משתנה מקרי אי-שלילי, אזי לכל $0<t$ ממשי מתקיים: $P(X\ge t)\le\frac{E(X)}{t}$.

כשמשתמשים במשפט זה צריך לדעת ש-X אי-שלילי.

באופן כללי החסם המתקבל מאי-שוויון זה הוא הדוק, כלומר שהוא יחסית קרוב להסתברות הנכונה.

משפט אי-שוויון צ'בישב - יהי X משתנה מקרי בעל שונות סופית, אזי לכל $0<t$ ממשי מתקיים: $P(|X-E(X)|\ge t)\le\frac{Var(X)}{t^2}$. החיסרון במשפט זה שצריך לחשב שונות, היתרון שהחסם משמעותי יותר בגלל שה-t בריבוע.

אם נגדיר את t להיות מכפלה כלשהי של סטיית התקן $\alpha\cdot\sigma=t$ עבור $\alpha>0$ ממשי, אזי לפי משפט צ'בישב נקבל $P(|X-E(X)|\ge\alpha\sigma)=\frac{1}{\alpha^2}$. כלומר ככל שמתרחקים מהתוחלת במכפלות של סטיית תקן ההסתברות הולכת וקטנה.

כי שיותר סביר שהערך של המ״מ יהיה קרוב לתוחלת, ככל שמתרחקים ההסתברות קטנה.

שיעור 13 - 27/5

א) תוחלת מותנה ושונות מותנה

יהי משתנה מקרי X במרחב ההסתברות (Ω,P), ויהי A מאורע כך ש- $P(A)>0$, אזי:

(1) התוחלת של המשתנה המקרי X המותנה ב-A היא: $E(X|A)=\sum_{\omega\in\Omega} X(\omega)\cdot P(\omega|A)=\sum_k k\cdot P(X=k|A)$,
כלומר סוכמים את המכפלות של כל הערכים שהוא יכול להתקבל בהסתברות שהם יקרו כאשר ידוע שיתרחש A.

(2) השונות של המשתנה המקרי X המותנה ב-A הוא: $Var(X|A)=E((X-E(X|A))^2|A)=E(X^2|A)-(E(X|A))^2$
כאשר $E(X^2|A)=\sum_k k^2\cdot P(X=k|A)$.

הערה - כל התכונות של התוחלת והשונות שלמדנו חלים גם על משתנה מקרי המותנה במאורע.

ב) משתנה מקרי שהוא תוחלת של משתנים מקריים מותנים

יהי X ו-Y שני משתנים מקריים במרחב ההסתברות (Ω,P), אזי נוכל להגדיר משתנה מקרי $E(X|Y)$, שהוא התוחלת של משתנה מקרי שהוא X מותנה ב-Y. משתנה מקרי זה מוגדר כך: $E(X|Y)(\omega)=E(X|Y=Y(\omega))$,

חישוב התוחלת של משתנה מקרי זה הוא: $E(E(X|Y))=\sum_y E(X|Y=y)\cdot P(Y=y)$.

משפט חוק התוחלת השלמה: יהי X ו-Y שתי משתנים מקריים במרחב ההסתברות (Ω,P) עם תוחלות סופיות, אזי מתקיים: $E(X)=E(E(X|Y))$.

ג) משתנה מקרי שהוא השונות של משתנים מקריים מותנים

יהי X ו-Y שני משתנים מקריים במרחב הסתברות (Ω,P), אזי נוכל להגדיר משתנה מקרי $Var(X|Y)$, שהוא השונות של משתנה מקרי X המותנה ב-Y. משתנה מקרי זה מוגדר כך: $Var(X|Y)(\omega)=Var(X|Y=Y(\omega))$.

משפט חוק השונות השלמה: יהי X ו-Y שני משתנים מקריים במרחב ההסתברות (Ω,P) עם תוחלות סופיות, אזי מתקיים: $Var(X)=E(Var(X|Y))+Var(E(X|Y))$.