מבוא להסתברות וסטטיסטיקה (0509-2801)

רון גולדמן

תוכן העניינים

זהו סיכום חלקי, שאינו מקיף את כל החומר בקורס.

מבוא להסתברות

מושגי יסוד של הסתברות

מרחבי הסתברות

הגדרה. מרחב מדגם Ω הוא קבוצת כל התוצאות האפשריות בניסוי.
תוצאה כלשהי במרחב המדגם נסמן על ידי $\omega \in \Omega$.

הגדרה. קבוצה חלקית של תוצאות ניסוי $A \subseteq \Omega$ תיקרא **מאורע** (Event).
אם תוצאת הניסוי היא $\omega \in A$, נאמר שהמאורע A התרחש.

הגדרה (מושגים בסיסיים מתורת הקבוצות). $e \in A$ - האיבר e שייך לקבוצה A.
$e \notin A$ - האיבר e אינו שייך לקבוצה A.
$A \subseteq B$ - הקבוצה A מוכלת בקבוצה B, כלומר לכל $a \in A$ מתקיים $a \in B$.
$A \subset B$ - הקבוצה A מוכלת חזק בקבוצה B, כלומר $A \subseteq B$ וגם $A \neq B$.
$A \cup B$ - איחוד, אוסף כל האיברים שנמצאים ב-A או ב-B או בשניהם.
$A \cap B$ - חיתוך, אוסף כל האיברים שנמצאים גם ב-A וגם ב-B.
A^c (או $\overline{A}$) - המשלים של A, כל האיברים ב-Ω שאינם ב-A.
$A \setminus B = A \cap B^c$ - כל האיברים ב-A שאינם ב-B.
קבוצות A, B זרות אם $A \cap B = \emptyset$.
מאורעות $A_1, \ldots, A_n$ זרים, אם $\bigcap_{i=1}^{n} A_i = \emptyset$.
$A_1, \ldots, A_n$ זרים בזוגות, אם לכל $i \neq j$ מתקיים $A_i \cap A_j = \emptyset$.

טענה. לכל A, B, C קבוצות מתקיים:

1. קומוטטיביות:

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

2. אסוציאטיביות:

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

3. דיסטריבוטיביות:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

4. $(A^c)^c = A$.

5. חוקי דה מורגן:

$$\begin{aligned}(A\cup B)^c &= A^c\cap B^c\\ (A\cap B)^c &= A^c\cup B^c\end{aligned}$$

הגדרה. יהי Ω מרחב מדגם. קבוצה $\mathcal{A}$ תיקרא **σ-אלגברה של Ω** אם ורק אם:

1. $\Omega\in\mathcal{A}$.
2. אם $A\in\mathcal{A}$, אזי גם $A^c\in\mathcal{A}$.
3. אם $\{A_n\}_{n=1}^{\infty}$ אוסף בן מניה של מאורעות ב-$\mathcal{A}$ אזי גם האיחוד $\bigcup_{n=1}^{\infty}A_n$ נמצא ב-$\mathcal{A}$.

הערה. כאשר Ω בדיד, $\mathcal{P}\left(\Omega\right)$ הוא σ-אלגברה על Ω ולכן ברוב המקרים נשתמש בו כאשר Ω בדיד.

הגדרה (אקסיומות קולמגורוב). יהי Ω מרחב מדגם, ותהי $\mathcal{A}$ σ-אלגברה על Ω. פונקציה $P:\mathcal{A}\to\mathbb{R}$ נקראת **פונקצית הסתברות** אם היא מקיימת את האקסיומות הבאות:

1. $P\left(A\right)\geq 0$ לכל $A\in\Omega$.
2. $P\left(\Omega\right)=1$.
3. לכל אוסף בן מניה $\{A_n\}_{n=1}^{\infty}$ של מאורעות הזרים בזוגות ב-$\mathcal{A}$ מתקיים

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P\left(A_n\right)$$

מרחב מדגם Ω יחד עם מרחב מאורעות $\mathcal{A}$ ופונקצית הסתברות P נקרא **מרחב הסתברות**.

משפט. לכל פונקצית הסתברות P מתקיים $P\left(\emptyset\right)$.

משפט (חוק החיבור). לכל אוסף סופי של מאורעות $\{A_n\}_{n=1}^{m}$ זרים בזוגות מתקיים כי

$$P\left(\bigcup_{n=1}^{m}A_n\right)=\sum_{n=1}^{m}P\left(A_n\right)$$

משפט (חוק המשלים). $P\left(A^c\right)=1-P\left(A\right)$.

משפט (מונוטוניות). אם $A\subseteq B$ אזי $P\left(A\right)\leq P\left(B\right)$.

משפט. לכל מאורע $A\in\mathcal{A}$ מתקיים כי $P\left(A\right)\leq 1$.

משפט (עקרון ההכלה וההדחה). אם $A,B\subseteq\Omega$ שני מאורעות אזי

$$P\left(A\cup B\right)=P\left(A\right)+P\left(B\right)-P\left(A\cap B\right)$$

מרחבי הסתברות סימטרים

הגדרה. מרחב הסתברות $(\Omega, \mathcal{A}, P)$ יקרא **סימטרי** אם $|\Omega|$ סופית וגם

$$\forall A \in \mathcal{A}. P(A) = \frac{|A|}{|\Omega|}$$

משפט (עקרון הכפל). אם מבצעים r בחירות בלתי תלויות כך שבכל בחירה יש m_r אפשרויות, אזי מספר האפשרויות הכולל ל-r הבחירות הוא $\prod_{k=1}^{r} m_k$.

מסקנה. מספר הסידורים של n פריטים בשורה הוא $n!$.

טענה. מספר הדרכים לבחור k איברים מתוך $n \geq k$ הוא

חשיבות לסדר/חזרות	כן	לא
אסור	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$
מותר	n^k	$\binom{k+n-1}{k}$

התפלגות מותנה ואי תלות

הגדרה. יהיו A, B מאורעות מעל אותו מרחב מדגם כך ש-$P(B) > 0$. ה**הסתברות מותנה של מאורע** A **בהינתן מאורע** B תסומן על ידי $\boldsymbol{P(A|B)}$ ומוגדרת על ידי

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

טענה. אם B מאורע כך ש-$P(B) > 0$ אז $P(\cdot|B)$ פונקציית הסתברות.

משפט (כלל השרשרת). יהיו $A_1, \ldots, A_n$ מאורעות כך שלכל $m \in \{1, \ldots, n\}$ מתקיים $P\left(\bigcap_{i=1}^{m} A_i\right) > 0$. אזי

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \prod_{i=2}^{n} P\left(A_i \middle| \bigcap_{j=1}^{i-1} A_j\right)$$

משפט (נוסחת ההסתברות השלמה). יהיו $B_1, \ldots, B_n$ מאורעות זרים בזוגות מעל אותו מרחב מדגם כך ש-$\biguplus_{i=1}^{n} B_i = \Omega$ וכן $P(B_i) > 0$ לכל $i \in \{1, \ldots, n\}$.
אז לכל $A \subseteq \Omega$ מתקיים

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

משפט (כלל בייז). יהיו $B_1, \ldots, B_n$ מאורעות זרים בזוגות מעל אותו מרחב מדגם כך ש-$\biguplus_{i=1}^{n} B_i = \Omega$, וכן לכל $i \in \{1, \ldots, n\}$ $P(B_i) > 0$. אזי לכל $A \subseteq \Omega$ כך ש-$P(A) > 0$ מתקיים

$$P(B_k|A) = \frac{P(A|B_k) P(B_k)}{\sum_{i=1}^{n} P(A|B_i) P(B_i)}$$

הגדרה. נאמר כי המאורעות $A, B \subseteq \Omega$ **בלתי תלויים** אם ורק אם מתקיים אחד מהתנאים הבאים:

1. $P\left(A\cap B\right)=P\left(A\right)P\left(B\right)$.
2. אם $P\left(B\right)>0$ אז $P\left(A|B\right)=P\left(A\right)$.
3. אם $P\left(A\right)>0$ אז $P\left(B|A\right)=P\left(B\right)$.

משפט. אם A ו-B מאורעות בלתי תלויים אזי גם הבאים בלתי תלויים:

- A ו-B^{c}.
- A^{c} ו-B.
- A^{c} ו-B^{c}.

משתנים מקריים חד-מימדיים

הגדרה. פונקציה $X:\Omega\rightarrow\mathbb{R}$ תיקרא **משתנה מקרי חד-מימדי** אם $\left\{ \omega\in\Omega\,|X\left(\omega\right)\leq k\right\}$ הינו מאורע לכל $k\in\mathbb{R}$.

סימון. אחר מעבר למשתנים מקריים, נסמן מאורעות על ידי אילוצים על המשתנים המקריים:

$$\left(X=k\right)\triangleq\left\{ \omega\in\Omega\,|X\left(\omega\right)=k\right\}$$

הגדרה. יהי X משתנה מקרי. **התומך** של X מוגדר להיות

$$\boldsymbol{S_{X}=}\mathrm{Supp}_{\boldsymbol{X}}\triangleq\mathrm{Im}\left(X\right)$$

הגדרה. פונקצית התפלגות מצטברת של משתנה מקרי X היא הפונקציה $F_{X}:\mathbb{R}\rightarrow\left[0,1\right]$ כך שלכל $k\in S_{X}$ מתקיים

$$F_{X}\left(k\right)=P\left(X\leq k\right)$$

טענה (**תכונות של פונקצית התפלגות מצטברת**). תהי $f:\mathbb{R}\rightarrow\left[0,1\right]$. קיים משתנה מקרי X כך ש-$f=F_{X}$ אם ורק אם כל הבאים מתקיימים:

1. $\lim_{k\rightarrow-\infty}F_{X}\left(k\right)=0$.
2. $\lim_{k\rightarrow\infty}F_{X}\left(k\right)=1$.
3. F_{X} מונוטונית עולה.
4. F_{X} רציפה מימין, כלומר לכל $k\in\mathbb{R}$ $F_{X}\left(k\right)=\lim_{h\rightarrow0^{+}}F_{X}\left(k+h\right)$.

משתנים מקריים בדידים

הגדרה. משתנה מקרי X יקרא **בדיד** אם קיימת $S\subset\mathbb{R}$ סופית או בת מניה כך שלכל $k\in S$ מתקיים כי $P\left(X=k\right)>0$ וגם $\sum_{k\in S}P\left(X=k\right)=1$.

הגדרה. יהי X משתנה מקרי ויהיו $n,m\in\mathbb{Z}$. נאמר כי X מתפלג **אחיד בדיד** בין n ל-m ונסמן $X\sim U\left[n,m\right]$ אם

$$S_X=\left\{k\in\mathbb{Z}\left|n\leq k\leq m\right.\right\}$$

ולכל $k\in S_X$ מתקיים

$$P\left(X=k\right)=\frac{1}{m-n+1}$$

הגדרה. יהיו X משתנה מקרי ו-$p\in\left[0,1\right]$. נאמר כי X מתפלג **ברנולי** או **אינדיקטור** עם הסתברות p אם

$$S_X=\left\{0,1\right\}$$

ולכל $k\in S_X$ מתקיים

$$P\left(X=k\right)=\begin{cases}1-p & k=0\\ p & k=1\end{cases}$$

הגדרה. יהיו X משתנה מקרי, $p\in\left[0,1\right]$ ו-$n\in\mathbb{N}$. נאמר כי X מתפלג **בינומי** ונסמן $X\sim\text{Bin}\left(n,p\right)$ אם הוא מתאר את מספר ההצלחות ב-n ניסויי ברנולי עם הסתברות p בלתי תלויים.

טענה. יהיו $n\in\mathbb{N}$, $p\in\left[0,1\right]$ כך ש-$X\sim\text{Bin}\left(n,p\right)$. אז $S_X=\left\{k\in\mathbb{N}\left|0\leq k\leq n\right.\right\}$ ולכל $k\in S_X$ מתקיים

$$P\left(X=k\right)=\binom{n}{k}p^k\left(1-p\right)^{n-k}$$

הגדרה. יהיו X משתנה מקרי ו-$p\in\left[0,1\right]$. נאמר כי X מתפלג **גיאומטרית** עם הסתברות p ונסמן $X\sim\text{Geo}\left(p\right)$ אם X מונה את מספר ניסויי הברנולי הבלתי תלויים עם הסתברות p עד להצלחה הראשונה (כולל) מתוך סדרת ניסויים אינסופית.

טענה. יהיו $p\in\left[0,1\right]$ ו-$X\sim\text{Geo}\left(p\right)$. אז $S_X=\mathbb{N}$ ולכל $k\in S_X$

$$P\left(X=k\right)=p\left(1-p\right)^{k-1}$$

טענה (**נוסחת הזנב**). יהיו $p\in\left[0,1\right]$ ו-$X\sim\text{Geo}\left(p\right)$. אז לכל $k\in S_X$ מתקיים:

$$P\left(X>k\right)=\left(1-p\right)^k$$

טענה (**תכונת חוסר הזיכרון**). יהיו $p\in\left[0,1\right]$ ו-$X\sim\text{Geo}\left(p\right)$. אז לכל $k\leq t\in\mathbb{N}$ מתקיים

$$P\left(X>t|X>k\right)=P\left(T>t-k\right)$$

הגדרה. יהיו $r\in\mathbb{N}^+$, $p\in\left[0,1\right]$ ו-X משתנה מקרי. נאמר כי X מתפלג **בינומית שלילית** עם פרמטרים r,p ונסמן $X\sim\text{NB}\left(r,p\right)$ אם X מונה את מספר ניסויי הברנולי עם הסתברות p עד ההצלחה ה-r (כולל) מתוך סדרת ניסויים אינסופית.

טענה. יהיו $r \in \mathbb{N}^+$, $p \in [0,1]$ ו-$X \sim \text{NB}(r,p)$. אז $S_X = \{k \in \mathbb{N} \,|\, r \le k\}$ ולכל $k \in S_X$ מתקיים:

$$P(X=k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

הגדרה. יהיו $n \in \mathbb{N}^+, D \le N \in \mathbb{N}^+$ ו-X משתנה מקרי. נאמר כי X מתפלג **היפר גאומטרית** עם פרמטרים N, D, n ונסמן $X \sim \text{HG}(N,D,n)$ אם מאוכלוסיה של N פריטים שבהם D "מיוחדים" ו-$N-D$ רגילים, מוצאים ללא החזרה n פריטים כך ש-X מספר הפריטים המיוחדים שהתקבלו.

טענה. יהיו $n \in \mathbb{N}^+, D \le N \in \mathbb{N}^+$ ו-$X \sim \text{HG}(N,D,n)$. אז $S_X = \{k \in \mathbb{N} \,|\, \max\{0, n-(N-D)\} \le k \le \min\{n, D\}\}$ ולכל $k \in S_X$ מתקיים:

$$P(X=k) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}$$

הגדרה. יהיו $0 < \lambda \in \mathbb{R}$ ו-X משתנה מקרי. נאמר כי X הוא **משתנה פואסוני** עם פרמטר λ ונסמן $X \sim \text{Poi}(\lambda)$ אם $S_X = \mathbb{N}$ ולכל $k \in S_X$ מתקיים:

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

טענה. יהיו $0 < n \in \mathbb{N}, \lambda \in \mathbb{R}$ ו-$X_n \sim \text{Bin}\left(n, \frac{\lambda}{n}\right)$. אז לכל $k \in S_{X_n}$ מתקיים

$$\lim_{n\to\infty} P(X_n = k) = \frac{\lambda e^{-\lambda}}{k!}$$

סימבולית:

$$\lim_{n\to\infty} X_n \sim \text{Poi}(\lambda)$$

טענה **(קירוב להתפלגות פואסונית)**. יהיו $0 < n \in \mathbb{N}, p \in [0,1]$ ו-$X \sim \text{Bin}(n,p)$. אם n גדול מאוד ו-p קטן מאוד אז לכל $k \in S_X$ מתקיים בקירוב:

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{(np)^k e^{-np}}{k!}$$

כלל אצבע: $n \ge 50$ ו-$np < 5$.

משתנים מקריים רציפים

הגדרה. משתנה מקרי X נקרא **רציף** אם קיימת **פונקצית צפיפות** $f_X : \mathbb{R} \to [0,\infty)$ רציפה למקוטעין כך שלכל $x \in \mathbb{R}$ מתקיים

$$F_X(x) \triangleq \int_{-\infty}^{x} f_X(t)\, dt$$

במקרה זה פונקציית ההתפלגות המצטברת רציפה ב-$\mathbb{R}$.

טענה. יהי X משתנה מקרי רציף ויהיו $a \le b \in \mathbb{R}$.

1. $P(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\, dx$.

2. $P\left(X=a\right)=\int_{a}^{a}f_{X}\left(x\right)dx=0$.
3. $P\left(X<a\right)=P\left(X\leq a\right)=\int_{-\infty}^{a}f_{X}\left(t\right)dt=F_{X}\left(a\right)$
4. $P\left(a\leq X\leq b\right)=P\left(a<X<b\right)=\int_{a}^{b}f_{X}\left(x\right)dx=F_{X}\left(b\right)-F_{X}\left(a\right)$.
5. $F_{X}\left(\infty\right)=\int_{\mathbb{R}}f\left(x\right)dx=1$.

משפט (המשפט היסודי של החשבון הדיפרנציאלי והאינטגרלי). יהא X מ"מ רציף. אז לכל $x\in\mathbb{R}$ מתקיים

$$\frac{d}{dx}F_{X}\left(x\right)=f_{X}\left(x\right)$$

הגדרה. תהא $f:\mathbb{R}\to\mathbb{R}$ רציפה למקוטעין. f תיקרא **פונקצית צפיפות** אם מתקיים:

1. $\forall x\in\mathbb{R}.f\left(x\right)\geq0$.
2. $\int_{\mathbb{R}}f\left(x\right)dx=1$.

הגדרה. יהיו $a<b\in\mathbb{R}$ ו-X משתנה מקרי רציף. נאמר כי X מתפלג **אחיד רציף** בקטע $[a,b]$ ונסמן $X\sim U\left(a,b\right)$ אם לכל $x\in\mathbb{R}$ מתקיים

$$f_{X}\left(x\right)=\begin{cases}\frac{1}{b-a} & x\in[a,b]\\0 & x\in\mathbb{R}\setminus[a,b]\end{cases}$$

טענה. יהיו $a<b\in\mathbb{R}$ ו-$X\sim U\left(a,b\right)$. אז לכל $x\in\mathbb{R}$ מתקיים

$$F_{X}\left(x\right)=\begin{cases}0 & x\in(-\infty,a]\\\frac{x-a}{b-a} & x\in[a,b]\\1 & x\in[b,\infty)\end{cases}$$

הגדרה. תהליך פואסון הינו תהליך אקראי $N\left(t\right)$ הסופר את מספר האירועים שקרו עד זמן t לכל $t\geq0$, כאשר התנאים הבאים מתקיימים:

1. $N\left(0\right)=0$.
2. לכל אינטרוול זמן $\left(s,s+t\right)$ מתקיים כי $N\left(s+t\right)-N\left(s\right)\sim\mathrm{Poi}\left(\lambda t\right)$ כאשר λ הינו ממוצע אירועים ליחידת זמן.
3. מספרי האירועים הקטעי זמן זרים הם בלתי תלויים.

הגדרה. יהיו $0<\lambda\in\mathbb{R}$ ו-X משתנה מקרי רציף. נאמר כי X מתפלג **מעריכית** עם פרמטר λ ונסמן $X\sim\exp\left(\lambda\right)$ אם

$$f_{X}\left(x\right)=\begin{cases}0 & x<0\\\lambda e^{-\lambda x} & x\geq0\end{cases}$$

טענה. יהיו $0<\lambda\in\mathbb{R}$ ו-$X\sim\exp\left(\lambda\right)$. אז לכל $x\in\mathbb{R}$ מתקיים

$$F_X\left(x\right)=\begin{cases}0 & x<0\\ 1-e^{-\lambda x} & x\geq0\end{cases}$$

משפט. יהי $N\left(t\right)$ תהליך פואסון עם פרמטר $0<\lambda\in\mathbb{R}$ מופעים ליחידת זמן. אזי הזמן באותן יחידות זמן עד המופע הראשון (וכן הזמן בין שני מופעים **עוקבים**) מתפלג מעריכית עם פרמטר λ.

טענה (**תכונת חוסרת הזיכרון של התפלגות מעריכית**). יהיו $0<\lambda\in\mathbb{R}$ ו-$X\sim\exp\left(\lambda\right)$. אז לכל $a\leq b\in\mathbb{R}$ מתקיים

$$P\left(X>b|X>a\right)=P\left(T>b-a\right)$$

הגדרה. יהיו $\mu\in\mathbb{R},0<\sigma\in\mathbb{R}$ ו-X מ"מ רציף. נאמר כי X מתפלג **נורמלית** עם פרמטרים μ,σ ונסמן $X\sim N\left(\mu,\sigma^2\right)$ אם לכל $x\in\mathbb{R}$ מתקיים

$$f_X\left(x\right)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\left(x-\mu\right)^2}{2\sigma^2}}$$

אם $\mu=0,\sigma=1$ נאמר כי X מתפלג **נורמלית סטנדרטית** ונסמנו ב-Z. נסמן $F_Z\triangleq\Phi$.

משפט. יהיו $\mu\in\mathbb{R},0<\sigma\in\mathbb{R}$ ו-$X\sim N\left(\mu,\sigma^2\right)$. אז לכל $a,b\in\mathbb{R}$ מתקיים

$$aX+b\sim N\left(a\mu+b,\left(a\sigma\right)^2\right)$$

מסקנה. יהיו $\mu\in\mathbb{R},0<\sigma\in\mathbb{R}$ ו-$X\sim N\left(\mu,\sigma^2\right)$. אזי $Z=\frac{X-\mu}{\sigma}\sim N\left(0,1\right)$, כלומר לכל $x\in\mathbb{R}$ מתקיים

$$F_X\left(x\right)=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

הערה. ההגדרה והטענה הבאה הם לשם השלמות ולא נלמדו בכיתה.

הגדרה. נגדיר את **טרנספורמציית התקנון** $T:\mathbb{R}\times\mathbb{R}^+\times\left(\Omega\to\mathbb{R}\right)\to\left(\Omega\to\mathbb{R}\right)$

$$T\left(X,\mu,\sigma\right)=\frac{X-\mu}{\sigma}$$

טענה. יהיו $\mu\in\mathbb{R},0<\sigma\in\mathbb{R}$ ו-$X\sim N\left(\mu,\sigma^2\right)$. אזי

$$T\left(X,\mu,\sigma\right)=Z$$

הערה. כאשר נתבקש למצוא ערך של התפלגות נורמלית נתקנן את המשתנה ונמצא את ההסתברות מתוך טבלת Z.

מומנטים ומדדים מרכזיים

הגדרה. יהי X מ"מ בעל פונקצית התפלגות מצטברת F_X **הפיכה**, ויהא $p\in[0,1]$. **האחוזון ה-p** של X הוא

$$x_p=\sup\left\{t\in\mathbb{R}\left|F_X\left(t\right)\leq p\right.\right\}$$

טענה. אם X מ"מ כך ש-F_X הפיכה אזי האחוזון ה-p $x_{(\cdot)}:[0,1]\to\mathbb{R}$ ההופכית שלה. כלומר לכל $x\in\mathbb{R}$ ולכל $p\in[0,1]$ מתקיים:

$$F_X\left(x_p\right)=p$$
$$x_{F_X(x)}=x$$

הגדרה (אחוזונים מיוחדים).

1. **החציון:** $x_{0.5}$
2. **רבעון תחתון:** $x_{0.25}$
3. **רבעון עליון:** $x_{0.75}$
4. **עשירון תחתון:** $x_{0.1}$
5. **עשירון עליון:** $x_{0.9}$

הגדרה. יהי X מ"מ בדיד. **התוחלת** (expectation) של X היא

$$\boldsymbol{E}\left[\boldsymbol{X}\right]\triangleq\sum_{k\in S_X}kP\left(X=k\right)$$

נאמר כי התוחלת $E\left[X\right]$ **קיימת** אם היא מתכנסת בהחלט (כטור). נאמר כי התוחלת $E\left[X\right]$ **סופית** (והמ"מ יקרא אינטגרבילי) אם הטור הינו סופי.

טענה. יהי X מ"מ בדיד, $p\in[0,1]$, $a\le b\in\mathbb{Z}$, $n,r\in\mathbb{N}^+$, $D\le N\in\mathbb{N}^+$, $0<\lambda\in\mathbb{R}$.

1. אם $X\sim U\left[a,b\right]$ אז $E\left[X\right]=\frac{a+b}{2}$.**טעות מסוג 2**
2. אם $X\sim\mathrm{Bin}\left(n,p\right)$ אז $E\left[X\right]=np$.
3. אם $X\sim\mathrm{Geo}\left(p\right)$ אז $E\left[X\right]=\frac{1}{p}$.
4. אם $X\sim\mathrm{NB}\left(r,p\right)$ אז $E\left[X\right]=\frac{r}{p}$.
5. אם $X\sim\mathrm{HG}\left(N,D,n\right)$ אז $E\left[X\right]=n\frac{D}{N}$.
6. אם $X\sim\mathrm{Poi}\left(\lambda\right)$ אז $E\left[X\right]=\lambda$.

הגדרה. יהי X מ"מ רציף. **התוחלת** של X היא

$$\boldsymbol{E}\left[\boldsymbol{X}\right]\triangleq\int\limits_{\mathbb{R}}xf_X\left(x\right)dx$$

נאמר כי התוחלת $E\left[X\right]$ **קיימת** אם היא מתכנסת בהחלט (כאינטגרל). נאמר כי התוחלת $E\left[X\right]$ **סופית** (והמ"מ יקרא אינטגרבילי) אם האינטגרל הינו סופי.

טענה. יהי X מ"מ רציף, $a<b\in\mathbb{R}$, $0<\lambda,\sigma\in\mathbb{R}$ ו-$\mu\in\mathbb{R}$.

1. אם $X\sim U\left(a,b\right)$ אז $E\left[X\right]=\frac{a+b}{2}$.

2. אם $X \sim \exp(\lambda)$ אז $E[X] = \frac{1}{\lambda}$.

3. אם $X \sim N(\mu, \sigma^2)$ אז $E[X] = \mu$.

הערה. לא לכל משתנה מקרי קיימת תוחלת.

טענה (**תכונות התוחלת**). יהיו X, Y מ"מ. אז

1. **ליניאריות:** לכל $a, b \in \mathbb{R}$ מתקיים $E[aX + b] = aE[X] + b$.
2. **סימטריה:** אם X סימטרי סביב $\mu \in \mathbb{R}$ (כלומר לכל $\varepsilon \in \mathbb{R}$ מתקיים כי $P(X \geq \mu + \varepsilon) = P(X \leq \mu - \varepsilon)$) אזי $E[X] = \mu$.
3. **אדיטיביות:** לכל $a, b \in \mathbb{R}$ מתקיים $E[X + Y] = E[X] + E[Y]$.
4. **תוחלת של פונקציה של משתנה מקרי:** לכל פונקציה g, אם X בדיד אז

$$E[g(X)] = \sum_{x \in S_X} g(x) P(X = x)$$

ואם X רציף אז

$$E[g(X)] = \int_{\mathbb{R}} g(x) f_X(x)\, dx$$

משפט (אי-שיוויון מרקוב). יהי X מ"מ אי-שלילי. אז לכל $a > 0$ מתקיים:

$$P(X \geq a) \leq \frac{E[X]}{a}$$

הגדרה. יהי X מ"מ. **השונות** (variance) של X היא

$$\mathrm{Var}\,(\boldsymbol{X}) \triangleq E\left[(X - E[X])^2\right]$$

טענה. יהי X מ"מ, $p \in [0,1]$, $a \leq b \in \mathbb{R}$, $n, r \in \mathbb{N}^+$, $D \leq N \in \mathbb{N}^+$, $0 < \lambda, \sigma \in \mathbb{R}$ ו-$\mu \in \mathbb{R}$.

1. אם $a, b \in \mathbb{Z}$ כך ש-$X \sim U[a, b]$ אז $\mathrm{Var}(X) = \frac{(b-a+1)^2 - 1}{12}$.
2. אם $X \sim \mathrm{Bin}(n, p)$ אז $\mathrm{Var}(X) = np(1-p)$.
3. אם $X \sim \mathrm{Geo}(p)$ אז $\mathrm{Var}(X) = \frac{1-p}{p^2}$.
4. אם $X \sim \mathrm{NB}(r, p)$ אז $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.
5. אם $X \sim \mathrm{HG}(N, D, n)$ אז $\mathrm{Var}(X) = n\frac{D(N-n)}{N(N-1)}\left(1 - \frac{D}{N}\right)$.
6. אם $X \sim \mathrm{Poi}(\lambda)$ אז $\mathrm{Var}(X) = \lambda$.
7. אם $a < b$ כך ש-$X \sim U(a, b)$ אז $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$.
8. אם $X \sim \exp(\lambda)$ אז $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

9. אם $X\sim N\left(\mu,\sigma^2\right)$ אז $\text{Var}\left(X\right)=\sigma^2$.

משפט. יהי X מ"מ חד-מימדי. אם $E^2\left[X\right]=\left(E\left[X\right]\right)^2$ אזי

$$\text{Var}\left(X\right)=E\left[X^2\right]-E^2\left[X\right]$$

הגדרה. יהי X מ"מ בעל שונות סופית. **סטיית התקן** (standard deviation) של X היא

$$\boldsymbol{\sigma_X}\triangleq\sqrt{\text{Var}\left(X\right)}$$

משפט. יהי X מ"מ בעל שונות סופית. אז לכל $\varepsilon\in\mathbb{R}$ מתקיים $\text{Var}\left(X+\varepsilon\right)=\text{Var}\left(X\right)$.

הגדרה. יהיו X,Y מ"מ. נאמר כי X,Y **בלתי תלויים** אם לכל $(x,y)\in S_X\times S_Y$ מתקיים כי המאורעות $\left(Y=y\right),\left(X=x\right)$ בלתי תלויים.

משפט. יהיו X,Y מ"מ ב"ת. אזי

$$\text{Var}\left(X+Y\right)=\text{Var}\left(X\right)+\text{Var}\left(Y\right)$$

משפט. יהי X מ"מ בעל שונות סופית. אז לכל $\alpha\in\mathbb{R}$ מתקיים:

$$\text{Var}\left(\alpha X\right)=\alpha^2\text{Var}\left(X\right)$$

בפרט:

$$\sigma_{\alpha X}=\alpha\sigma_X$$

משפט (אי-שיוויון צ'בישב). יהי X מ"מ. אז לכל $0<a\in\mathbb{R}$ מתקיים:

$$P\left(\left|X-E\left[X\right]\right|\geq a\right)\leq\frac{\text{Var}\left(X\right)}{a^2}$$

הגדרה. יהי X מ"מ ו-$k\in\mathbb{N}$. **המומנט ה-$\boldsymbol{k}$** של X הוא $m_k\triangleq E\left[X^k\right]$, בתנאי שהתוחלת מתכנסת בהחלט.

הגדרה. יהי X מ"מ. **הפונקציה יוצרת מומנטים** של X היא $M_X\left(t\right)\triangleq E\left[e^{tX}\right]$.

טענה. יהי X מ"מ, $p\in\left[0,1\right]$, $a\leq b\in\mathbb{R}$, $n,r\in\mathbb{N}^+$, $D\leq N\in\mathbb{N}^+$, $0<\lambda,\sigma\in\mathbb{R}$ ו-$\mu\in\mathbb{R}$.

1. אם $a,b\in\mathbb{Z}$ כך ש-$X\sim U\left[a,b\right]$ אז $M_X\left(t\right)=\frac{e^{at}\left(1-e^{(b-a+1)t}\right)}{(b-a+1)(1-e^t)}$.
2. אם $X\sim\text{Bin}\left(n,p\right)$ אז $M_X\left(t\right)=\left(1-p+pe^t\right)^n$.
3. אם $X\sim\text{Geo}\left(p\right)$ אז $M_X\left(t\right)=\frac{pe^t}{1-(1-p)e^t}$.
4. אם $X\sim\text{Poi}\left(\lambda\right)$ אז $M_X\left(t\right)=e^{\lambda\left(e^t-1\right)}$.
5. אם $a<b$ כך ש-$X\sim U\left(a,b\right)$ אז $M_X\left(t\right)=\frac{e^{tb}-e^{ta}}{t(b-a)}$.
6. אם $X\sim\exp\left(\lambda\right)$ אז $M_X\left(t\right)=\frac{\lambda}{\lambda-t}$.

7. אם $X\sim N\left(\mu,\sigma^{2}\right)$ אז $M_{X}\left(t\right)=e^{\mu t+\frac{1}{2}\sigma^{2}t^{2}}$.

טענה (**תכונות חשובות של פונקציה יוצרת מומנטים**). יהי X מ"מ.

1. $M_{X}\left(0\right)=1$.
2. לכל $k\in\mathbb{N}$, אם M_{X} גזירה k פעמים אז:
$$\frac{d^{k}M_{X}}{dt^{k}}\left(0\right)=m_{k}$$
3. לכל $a,b\in\mathbb{R}$ מתקיים $M_{aX+b}\left(t\right)=e^{tb}M_{X}\left(at\right)$
4. אם קיים $\varepsilon>0$ עבורו M_{X} קיימת ב-ε סביבה של 0 אז כל המומנטים קיימים וסופיים ומתקיים:
$$M_{X}\left(t\right)=\sum_{n=0}^{\infty}\frac{t^{n}E\left[x^{n}\right]}{n!}$$

משפט. קיימת התאמה חד-חד ערכית בין התפלגות לפונקציה יוצרת מומנטים.

משתנים מקריים דו מימדיים

התפלגות משותפת

הגדרה. יהיו X,Y מ"מ מעל אותו מרחב הסתברות $\left(\Omega,\mathcal{A},P\right)$. אזי $\left(X,Y\right)$ הוא **משתנה מקרי דו מימדי** $\left(X,Y\right):\Omega\to\mathbb{R}^{2}$ אם לכל $k,l\in\mathbb{R}$ מתקיים כי $\left\{ \omega\in\mathbb{R}\left|X\left(\omega\right)\le k\wedge Y\left(\omega\right)\le l\right.\right\}$ מאורע.

הגדרה. יהיו X,Y מ"מ בדידים מעל אותו מרחב. **פונקצית ההסתברות** של $\left(X,Y\right)$ היא הפונקציה $P_{X,Y}:S_{X,Y}\to\left[0,1\right]$ כך שלכל $\left(k,l\right)\in S_{X,Y}$ מתקיים
$$P_{X,Y}\left(k,l\right)\triangleq P\left(X=k,Y=l\right)$$

טענה. אם X,Y מ"מ בדידים ו-A מאורע אזי
$$P\left(A\right)=\sum_{\left(x,y\right)\in A}P_{X,Y}\left(x,y\right)$$

הגדרה. יהיו X,Y מ"מ בדידים מעל אותו מרחב הסתברות, **פונקצית ההסתברות השולית** של X היא $P_{X}:S_{X}\to\left[0,1\right]$ המוגדרת באופן הבא:
$$\forall k\in S_{X}.\ P_{X}\left(k\right)=\sum_{l\in S_{Y}}P_{X,Y}\left(k,l\right)$$

באותו אופן מגדירים את P_{Y}.

הגדרה. יהיו X,Y מ"מ מעל אותו מרחב הסתברות. **פונקצית ההתפלגות המצטברת** של $\left(X,Y\right)$ היא הפונקציה $F_{X,Y}:S_{X,Y}\to\left[0,1\right]$ כך שלכל $\left(x,y\right)\in S_{X,Y}$ מתקיים
$$F_{X,Y}\left(x,y\right)\triangleq P\left(X\le x,Y\le y\right)$$

טענה (**תכונות של פונקציית התפלגות מצטברת של מ"מ דו מימדי**). יהיו X,Y מ"מ מעל אותו מרחב הסתברות ויהיו $(k,l)\in S_{X,Y}$. אזי

1.

$$\lim_{k\to\infty} F_{X,Y}\left(k,l\right)=0$$
$$\lim_{l\to\infty} F_{X,Y}\left(k,l\right)=0$$
$$\lim_{(k,l)\to(\infty,\infty)} F_{X,Y}\left(k,l\right)=1$$

2. **מונוטוניות:** אם $k_1<k_2$ ו-$l_1<l_2$ אזי $F_{X,Y}\left(k_2,l_2\right)\geq F_{X,Y}\left(k_1,l_1\right)$ ובפרט מתקיים

$$P\left(k_1\leq X\leq k_2,l_1\leq Y\leq l_2\right)=F_{X,Y}\left(k_2,l_2\right)-F_{X,Y}\left(k_1,l_2\right)-F_{X,Y}\left(k_2,l_1\right)+F_{X,Y}\left(k_1,l_1\right)\geq 0$$

3. $F_{X,Y}$ **רציפה מימין בשני הרכיבים:**

$$\lim_{(h_1,h_2)\to(0^+,0^+)} F_{X,Y}\left(k+h_1,l+h_2\right)=F_{X,Y}\left(k,l\right)$$

הגדרה. מ"מ דו מימדי (X,Y) מעל מרחב הסתברות $(\Omega,\mathcal{A},P)$ הוא **משנה מקרי דו מימדי רציף** אם קיימת **פונקציית צפיפות משותפת** $f_{X,Y}:\mathbb{R}^2\to[0,\infty)$ רציפה למקוטעין (רציפה פרט לקבוצה ממידה אפס) כך שלכל $(k,l)\in S_{X,Y}$ מתקיים:

$$F_{X,Y}\left(k,l\right)=\int\limits_{-\infty}^{k}\int\limits_{-\infty}^{l} f_{X,Y}\left(u,v\right)dudv$$

טענה. יהי (X,Y) מ"מ דו-מימדי רציף. אז

1.

$$\iint\limits_{\mathbb{R}^2} f_{X,Y}\left(u,v\right)dudv=1$$

2. לכל $k,l\in\mathbb{R}$ מתקיים:

$$f_{X,Y}\left(k,l\right)=\frac{\partial^2 F_{X,Y}}{\partial k\partial l}\left(k,l\right)$$

הגדרה. יהי (X,Y) מ"מ דו-מימדי רציף. אז X,Y מ"מ רציפים יחד עם **פונקציית הצפיפות השולית:**

$$\forall u\in\mathbb{R}.f_X\left(u\right)=\int\limits_{\mathbb{R}} f_{X,Y}\left(u,v\right)dv$$
$$\forall v\in\mathbb{R}.f_Y\left(v\right)=\int\limits_{\mathbb{R}} f_{X,Y}\left(u,v\right)du$$

מדדים מרכזיים של משתנה מקרי דו-מימדי

הגדרה. יהי (X,Y) מ"מ דו-מימדי אז **התוחלת** של $g(X,Y):\mathbb{R}^2\to\mathbb{R}$ מוגדרת להיות:

במקרה הבדיד

$$\boldsymbol{E}\left[\boldsymbol{g}\left(\boldsymbol{X},\boldsymbol{Y}\right)\right]\triangleq\sum_{x\in S_X}\sum_{y\in S_Y}g\left(x,y\right)P_{X,Y}\left(x,y\right)$$

במקרה הרציף

$$\boldsymbol{E}\left[\boldsymbol{g}\left(\boldsymbol{X},\boldsymbol{Y}\right)\right]\triangleq\iint_{\mathbb{R}^2}g\left(x,y\right)f_{X,Y}\left(x,y\right)dxdy$$

הגדרה. יהיו X,Y מ"מ מעל אותו מרחב הסתברות. **השונות המשותפת** (covariance) של (X,Y) מוגדרת להיות

$$\mathbf{Cov}(\boldsymbol{X},\boldsymbol{Y})\triangleq E\left[\left(X-E\left[X\right]\right)\left(Y-E\left[Y\right]\right)\right]$$

טענה. יהיו X,Y מ"מ מעל אותו מרחב הסתברות.

- אם ככל ש-X גדל Y קטן (ולהיפך), אזי $\text{Cov}\left(X,Y\right)<0$.
- אם ככל ש-X גדל Y גדל (ולהיפך), אזי $\text{Cov}\left(X,Y\right)>0$.
- אם אף אחד מהקודמים לא מתקיים אז $\text{Cov}\left(X,Y\right)=0$.

הגדרה. יהיו X,Y מעל אותו מרחב הסתברות. נאמר כי X,Y **בלתי מתואמים** אם $\text{Cov}\left(X,Y\right)=0$.

טענה. לכל מ"מ X מתקיים $\text{Cov}\left(X,X\right)=\text{Var}\left(X\right)$.

משפט. יהיו X,Y מ"מ מעל אותו מרחב הסתברות. אזי

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

הגדרה. יהיו X,Y מ"מ מעל אותו מרחב הסתברות. **מקדם המתאם** (Pearson correlation coefficient) של (X,Y) מוגדר להיות

$$\boldsymbol{\rho_{X,Y}}\triangleq\frac{\text{Cov}\left(X,Y\right)}{\sigma_X\sigma_Y}$$

טענה (**תכונות מקדם המתאם**). יהיו X,Y מ"מ מעל אותו מרחב הסתברות. אז

1. $-1\le\rho_{X,Y}\le1$.
2. $\rho_{X,Y}=1$ אם ורק אם קיימים $a>0,b\in\mathbb{R}$ כך ש-$P\left(Y=aX+b\right)=1$.
3. $\rho_{X,Y}=-1$ אם ורק אם קיימים $a<0,b\in\mathbb{R}$ כך ש-$P\left(Y=aX+b\right)=1$.
4. $\rho_{X,Y}=0\iff\text{Cov}\left(X,Y\right)=0$.

התפלגות מותנה ואי תלות של משתנים מקריים

הגדרה. יהי (X,Y) מ"מ דו-מימדי בדיד מעל מרחב הסתברות $(\Omega,\mathcal{A},P)$, ויהי $u\in S_X$. אז **פונקצית ההסתברות המותנית** של Y בהינתן $X=u$ תסומן ע"י $P_{Y|X=u}$ כך שלכל $v\in S_Y$ מתקיים:

$$P_{Y|X=u}\left(v\right)=\frac{P_{X,Y}\left(u,v\right)}{P_X\left(u\right)}$$

בתנאי ש-$P_X\left(u\right)>0$. באותו אופן מגדירים את $P_{X|Y=v}$ עבור $v\in S_Y$.

הגדרה. יהי (X,Y) מ"מ דו-מימדי רציף מעל מרחב הסתברות $(\Omega,\mathcal{A},P)$, ויהי $u\in\mathbb{R}$. אז **פונקצית הצפיפות המותנית** של Y בהינתן $X=u$ תסומן ע"י $P_{Y|X=u}$ כך שלכל $v\in\mathbb{R}$ מתקיים:

$$f_{Y|X=u}\left(v\right)=\frac{f_{X,Y}\left(u,v\right)}{f_X\left(u\right)}$$

בתנאי ש-$f_X\left(u\right)>0$. באותו אופן מגדירים את $f_{X|Y=v}$ עבור $v\in\mathbb{R}$.

הגדרה. יהי (X,Y) מ"מ רציף דו-מימדי מעל מרחב הסתברות $(\Omega,\mathcal{A},P)$, ויהי $u\in\mathbb{R}$. אז **פונקצית ההתפלגות המצטברת** של Y בהינתן $X=u$ תסומן ע"י $F_{Y|X=u}$ כך שלכל $v\in\mathbb{R}$ מתקיים:

$$F_{Y|X=u}\left(v\right)=\int\limits_{-\infty}^{v}f_{Y|X=u}\left(v\right)dv$$

בתנאי ש-$f_X\left(u\right)>0$. באותו אופן מגדירים את $F_{X|Y=v}$ עבור $v\in\mathbb{R}$.

הגדרה. יהי (X,Y) מ"מ דו-מימדי. נאמר כי X,Y **בלתי תלויים** (סטוכסטית) אם לכל $(u,v)\in S_{X,Y}$ מתקיים $F_{X,Y}\left(u,v\right)=F_X\left(u\right)F_X\left(v\right)$.

טענה. יהי (X,Y) מ"מ דו-מימדי. אם קיימות $h,g:\mathbb{R}\to[0,\infty)$ כך שלכל $x,y\in\mathbb{R}$ מתקיים $f_{X,Y}\left(x,y\right)=h\left(x\right)g\left(x\right)$, אז X,Y ב"ת.

יתרה מזו, $\tilde{h}=\frac{h}{\|h\|_{L_1}},\tilde{g}=\frac{g}{\|g\|_{L_1}}$ הן פונקציות הצפיפויות השוליות של X,Y בהתאמה.

משפט. יהי (X,Y) מ"מ דו-מימדי. אם X,Y ב"ת אז לכל $f,g:\mathbb{R}\to\mathbb{R}$ גם $f\left(X\right),g\left(Y\right)$ ב"ת.

משפט. יהיו X,Y מ"מ מעל אותו מרחב הסתברות. אם X,Y ב"ת אז

$$E\left[XY\right]=E\left[X\right]E\left[Y\right]$$

מסקנה. אם X,Y ב"ת אזי $\mathrm{Var}\left(X+Y\right)=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)$.

מסקנה. אם X,Y ב"ת אז הם ב"מ.

הגדרה. יהי (X,Y) מ"מ דו-מימדי בדיד , ויהי $u\in S_X$. אז **התוחלת המותנה** של $g:\mathbb{R}^2\to\mathbb{R}$ בהינתן $X=u$ מוגדרת להיות

$$\boldsymbol{E\left[g\left(X,Y\right)|X=u\right]}\triangleq\sum_{v\in S_Y}g\left(u,v\right)P_{Y|X=u}\left(v\right)$$

הגדרה. יהי (X,Y) מ"מ דו-מימדי רציף , ויהי $u\in\mathbb{R}$. אז **התוחלת המותנה** של $g(X,Y):\mathbb{R}^2\to\mathbb{R}$ בהינתן $X=u$ מוגדרת להיות

$$\boldsymbol{E}\left[\boldsymbol{g}\left(\boldsymbol{X},\boldsymbol{Y}\right)|\boldsymbol{X}=\boldsymbol{u}\right]\triangleq\int\limits_{\mathbb{R}}g\left(u,v\right)f_{Y|X=u}\left(v\right)dv$$

משפט (התוחלת השלמה). יהי (X,Y) מ"מ דו-מימדי. אז לכל $g:\mathbb{R}\to\mathbb{R}$ מתקיים

$$E\left[g\left(Y\right)\right]=E\left[E\left[g\left(Y\right)|X\right]\right]$$

בפרט $E\left[Y\right]=E\left[E\left[Y|X\right]\right]$.

משפט. יהי (X,Y) מ"מ דו-מימדי. אז לכל $g_1,g_2:\mathbb{R}\to\mathbb{R}$ ו-$u\in\mathbb{R}$ מתקיים

$$E\left[g_1\left(Y\right)+g_2\left(Y\right)|X=u\right]=E\left[g_1\left(Y\right)|X=u\right]+E\left[g_2\left(Y\right)|X=u\right]$$
$$E\left[g_1\left(Y\right)g_2\left(X\right)|X=u\right]=g_2\left(u\right)E\left[g_1\left(Y\right)|X=u\right]$$

הסתברות רב-נורמלית

הגדרה. מ"מ (X,Y) מתפלג **דו-נורמלית** אם קיימים $\mu_X,\mu_Y\in\mathbb{R},\sigma_X,\sigma_Y,\rho>0$ כך שלכל $x,y\in\mathbb{R}$ מתקיים:

$$f_{X,Y}\left(x,y\right)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)+\left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}$$

טענה (**תכונות של התפלגות דו-נורמלית**). יהי (X,Y) מתפלג דו-נורמלית.

1. מתקיים $X\sim N\left(\mu_X,\sigma_X^2\right),Y\sim N\left(\mu_Y,\sigma_Y^2\right)$ ומקדם המתאם $\rho_{X,Y}=\rho$.
2. לכל $x\in\mathbb{R}$:
$$Y|X=x\sim N\left(\mu_Y+\frac{\sigma_Y}{\sigma_X}\rho\left(x-\mu_X\right),\left(1-\rho^2\right)\sigma_Y^2\right)$$
לכל $y\in\mathbb{R}$:
$$X|Y=y\sim N\left(\mu_X+\frac{\sigma_X}{\sigma_Y}\rho\left(y-\mu_Y\right),\left(1-\rho^2\right)\sigma_X^2\right)$$
3. $X\pm Y\sim N\left(\mu_X\pm\mu_Y,\sigma_X^2\pm2\rho\sigma_X\sigma_Y+\sigma_Y^2\right)$.
4. X,Y ב"ת אמ"מ X,Y ב"מ.

פוקנציות של משתנים מקריים ומשפט הגבול המרכזי

מינימום ומקסימום של משתנים מקריים

הגדרה. יהיו $X_1,\ldots,X_n$ מ"מ מעל אותו מרחב. נאמר כי $X_1,\ldots,X_n$ **בלתי תלויים** אם לכל $i\neq j\in[n]$ X_i,X_j בלתי תלויים.

הגדרה. יהיו $X_1,\dots,X_n$ מ"מ מעל אותו מרחב. נאמר כי $X_1,\dots,X_n$ **שווי התפלגות** אם לכל $i,j\in[n]$ מתקיים $F_{X_i}=F_{X_j}$.

טענה. יהיו $X_1,\dots,X_n$ מ"מ ב"ת ש"ה עם פונקציית התפלגות מצטברת F. אם $M=\max\{X_1,\dots,X_n\}$ אז לכל $m\in\mathrm{Dom}F$ מתקיים

$$F_M(m)=[F(m)]^n$$

אם $X_1,\dots,X_n$ רציפים עם פ"צ f אז:

$$f_M(m)=n[F(m)]^{n-1}f(m)$$

טענה. יהיו $X_1,\dots,X_n$ מ"מ ב"ת ש"ה עם פונקציית התפלגות מצטברת F. אם $Y=\min\{X_1,\dots,X_n\}$ אז לכל $y\in\mathrm{Dom}F$ מתקיים

$$F_Y(y)=1-[1-F(y)]^n$$

אם $X_1,\dots,X_n$ רציפים עם פ"צ f אז:

$$f_Y(y)=n[1-F(y)]^{n-1}f(y)$$

תוחלת וסכום של משתנים מקריים

משפט. יהיו X,Y מ"מ מעל אותו מרחב. אזי:

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)$$

משפט. יהיו $X_1,\dots,X_n$ מ"מ מעל אותו מרחב. אזי:

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n\mathrm{Var}(X_i)+2\sum_{i<j}\mathrm{Cov}(X_i,X_j)=\sum_{i=1}^n\mathrm{Var}(X_i)+\sum_{i\neq j}\mathrm{Cov}(X_i,X_j)$$

משפט. אם X,Y ב"ת, אזי $M_{X+Y}(t)=M_X(t)M_Y(t)$ לכל $t\in\mathbb{R}$.

משפט. יהיו $X_1,\dots,X_n$ מ"מ ב"ת ו-$\lambda_1,\dots,\lambda_n$ כך שלכל $i\in[n]$ מתקיים $X_i\sim\mathrm{Poi}(\lambda_i)$. אזי

$$S_n=\sum_{i=1}^n X_i\sim P\left(\sum_{i=1}^n\lambda_i\right)$$

משפט. אם (X,Y) מ"מ דו-מימדי. נסמן $W=X+Y$.

1. אם (X,Y) בדיד אז לכל $w\in S_X+S_Y$ מתקיים:

$$P_W(w)=\sum_{x=-\infty}^{\infty}P_{X,Y}(x,w-x)=\sum_{y=-\infty}^{\infty}P_{X,Y}(w-y,y)$$

אם בנוסף, X,Y ב"ת אז נקבל את **נוסחת הקונבולוציה הבדידה:**

$$P_W\left(w\right)=\left(P_X*P_Y\right)\left(w\right)=\left(P_Y*P_X\right)\left(w\right)=\sum_{x=-\infty}^{\infty}P_X\left(x\right)P_Y\left(w-x\right)=\sum_{y=-\infty}^{\infty}P_X\left(w-y\right)P\left(y\right)$$

2. אם $\left(X,Y\right)$ רציף אז לכל $w\in\mathbb{R}$ מתקיים:

$$f_W\left(w\right)=\int\limits_{\mathbb{R}}f\left(x,w-x\right)dx=\int\limits_{\mathbb{R}}f\left(w-y,y\right)dy$$

אם בנוסף, X,Y ב"ת אז נקבל את **נוסחת הקונבולוציה הרציפה:**

$$f_W\left(w\right)=\left(f_X*f_Y\right)\left(w\right)=\left(f_Y*f_X\right)\left(w\right)=\int\limits_{\mathbb{R}}f_X\left(x\right)f_Y\left(w-x\right)dx=\int\limits_{\mathbb{R}}f_X\left(w-y\right)f_Y\left(y\right)dy$$

החוק החלש של המספרים הגדולים ומשפט הגבול המרכזי

הגדרה. תהי $\left\{X_n\right\}_{n=1}^{\infty}$ סדרת תצפיות ב"ת ש"ה עם תוחלת סופית $E\left[X_n\right]=\mu\in\mathbb{R}$ לכל $n\in\mathbb{N}$.

- **הסכום** של $n\in\mathbb{N}$ תצפיות מוגדר להיות:

$$\boldsymbol{S_n}\triangleq\sum_{k=1}^{n}X_n$$

- **הממוצע** של $n\in\mathbb{N}$ תצפיות מוגדר להיות:

$$\overline{\boldsymbol{X_n}}\triangleq\frac{S_n}{n}$$

משפט (החוק החלש של המספרים הגדולים). תהי $\left\{X_n\right\}_{n=1}^{\infty}$ סדרת מ"מ ב"מ ש"ה עם תוחלת סופית $E\left[X_n\right]=\mu\in\mathbb{R}$ לכל $n\in\mathbb{N}$. אזי, הממוצע המדגמי מתכנס בהסתברות לתוחלת, כלומר לכל $\varepsilon>0$ מתקיים:

$$P\left(\left|\overline{X_n}-\mu\right|>\varepsilon\right)\underset{n\to\infty}{\longrightarrow}0$$

משפט (הגבול המרכזי). תהי $\left\{X_n\right\}_{n=1}^{\infty}$ סדרת מ"מ ב"ת ש"ה כך שלכל $n\in\mathbb{N}$ מתקיים $E\left[X_n\right]=\mu\in\mathbb{R},\mathrm{Var}\left(X_n\right)=\sigma^2\in\mathbb{R}$. אזי לכל $z\in\mathbb{R}$ מתקיים:

$$P\left(\frac{S_n-n\mu}{\sigma\sqrt{n}}\leq z\right)\underset{n\to\infty}{\longrightarrow}\Phi\left(z\right)$$

באופן מעשי, עבור n גדול מספיק ניתן להניח כי $\frac{S_n-n\mu}{\sigma\sqrt{n}}\sim N\left(0,1\right)$ וכן להניח כי $\overline{X_n}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$.

הערה. כאשר נרצה לקרה להתפלגות נורמלית:

1. עבור התפלגות סימטרית נדרוש $n\geq20$, עבור התפלגות מוטה נדרוש $n\geq100$.
2. היות והתצפיות הן ב"ת אז אם המ"מ מ"נ אז גם סכומן והממוצע שלהם מ"נ.

3. במקרה של משתנים בדידים המקבלים ערכים שלמים ועוקבים יש לבצע **תיקון רציפות.** לכל $a\in\mathbb{Z}$:

$$P\left(S_n\le a\right)=P\left(S_n<a-0.5\right)$$
$$P\left(S_n<a\right)=P\left(S_n<a+0.5\right)$$
$$P\left(S_n\ge a\right)=P\left(S_n>a-0.5\right)$$
$$P\left(S_n>a\right)=P\left(S_n>a+0.5\right)$$

הערה. ניתן להציג את המ"מ הבינומי כסכום של מ"מ ברנוליים ב"ת, ולהפעיל את משפט הגבול המרכזי כאשר $np\ge 5$ וגם $n\left(1-p\right)\ge 5$, כלומר n גדול מספיק ו-p "מרכזי" מספיק. זהו הקירוב המשלים לקירוב הפואסוני.
בהתפלגות הבינומית: $X=S_n\sim N\left(np,np\left(1-p\right)\right)$
והממוצע המדגמי: $\overline{X_n}=\frac{X}{n}\sim N\left(p,\frac{p(1-p)}{n}\right)$.

מבוא לסטטיסטיקה

הגדרה. **מדגם מקרי פשוט** הינו אוסף סופי של איברים שנבחרים בזה אחר זה מתוך אוכלוסייה באופן אקראי ועם החזרה, כך שכל התצפיות במדגם $X_1,\ldots,X_n$ בלתי תלויות ובעלות אותה התפלגות שהיא התפלגות האוכלוסייה.

אמידה

הגדרה. פונקציה של תצפיות המדגם, אשר אינה תלויה בשום פרמטר שאינו ידוע נקראת **סטטיסטי**.

הגדרה. סטטיסטי המשמש לאמידת פרמטר θ נקרא **אומד** עבור הפרמטר. הערך שמקבל העומד עבור מדגם ספציפי נקרא **אומדן**.

אמידה נקודתית

הגדרה. יהי $X_1,\ldots,X_n$ מדגם של n תצפיות באוכלוסייה בעלת התפלגות בדידה P_θ, התלויה בפרמטר θ בלתי ידוע. יהיו $x_1,\ldots,x_n$ תוצאות המדגם.
פונקצית הנראות מוגדרת על ידי:

$$L\left(\theta;x_1,\ldots,x_n\right)=L\left(\theta\right)=P_\theta\left(X_1=x_1,\ldots,X_n=x_n\right)$$

לכל ערך אפשרי של הפרמטר θ.

טענה. אם$X_1,\ldots,X_n$ מדגם של n תצפיות ב"ת באוכלוסייה בעלת התפלגות בדידה P_θ, התלויה בפרמטר θ בלתי ידוע. אז לכל $x_1,\ldots,x_n$ תוצאות מדגם מתקיים:

$$L\left(\theta;x_1,\ldots,x_n\right)=\prod_{k=1}^{n}P_\theta\left(X_k=x_k\right)$$

הגדרה. יהי $X_1,\ldots,X_n$ מדגם של n תצפיות באוכלוסייה בעלת התפלגות רציפה עם צפיפות f_θ, התלויה בפרמטר θ בלתי ידוע. יהיו $x_1,\ldots,x_n$ תוצאות המדגם.

פונקצית הנראות מוגדרת על ידי:

$$L\left(\theta;x_1,\ldots,x_n\right)=L\left(\theta\right)=f_\theta\left(x_1,\ldots,x_n\right)$$

לכל ערך אפשרי של הפרמטר θ.

טענה. אם $X_1,\ldots,X_n$ מדגם של n תצפיות באוכלוסייה בעלת התפלגות רציפה עם צפיפות f_θ, התלויה בפרמטר θ בלתי ידוע. אז לכל $x_1,\ldots,x_n$ תוצאות מדגם מתקיים:

$$L\left(\theta;x_1,\ldots,x_n\right)=\prod_{k=1}^{n}f_\theta\left(x_k\right)$$

הגדרה. יהי $X_1,\ldots,X_n$ מדגם של n תצפיות באוכלוסייה בעלי התפלגות התלויה בפרמטר θ בלתי ידוע. יהיו $x_1,\ldots,x_n$ תוצאות המדגם.
הערך $\hat{\theta}\left(x_1,\ldots,x_n\right)$ נקרא **אומד נראות מירבית** אם $\theta=\hat{\theta}\left(x_1,\ldots,x_n\right)$ מקסימום של $L\left(\theta\right)$.

טענה. יהי $X\sim\mathrm{Bin}\left(n,p\right)$ מדגם. אז לכל תוצאה x של המדגם, $\hat{p}\left(x\right)=\frac{x}{n}$ הוא אומד נראות מירבית של הפרמטר p.

טענה. יהי $X_1,\ldots,X_n\sim\exp\left(\lambda\right)$ מדגם מקרי ופשוט. אז לכל תוצאה $x_1,\ldots,x_n$ של המדגם, $\hat{\lambda}\left(x_1,\ldots,x_n\right)=\frac{1}{\overline{x_n}}$ (מוגדר היטב כי $P\left(x_1=\ldots=x_n=0\right)=0$) הוא אומד נראות מירבית של הפרמטר λ.

טענה. יהי $X_1,\ldots,X_n\sim U\left(0,\theta\right)$ מדגם מקרי ופשוט כאשר θ אינו ידוע. אז לכל תוצאה $x_1,\ldots,x_n$ של המדגם, $\hat{\theta}\left(x_1,\ldots,x_n\right)=\max\left\{x_1,\ldots,x_n\right\}$ הוא אומד נראות מירבית של הפרמטר θ.

הגדרה. אומד T הינו **אומד חסר הטיה** של פרמטר θ, אם $E\left[T\right]=\theta$ לכל ערך של θ.
ההפרש $E\left[T\right]-\theta$ נקרא **ההטיה של האומד** T.

משפט. יהי $X_1,\ldots,X_n$ מדגם של n תצפיות מאוכלוסייה בעלת תוחלת μ, אז לכל תוצאה $x_1,\ldots,x_n$ של המדגם, $\hat{\mu}\left(x_1,\ldots,x_n\right)=\overline{x_n}$ הוא אומד חסר הטיה עבור הפרמטר μ.

הגדרה. את טיב האומד T לפרמטר θ נמדוד ע"י **תוחלת ריבוע הטעות**:

$$\mathrm{MSE}\left(T\right)=E\left[\left(T-\theta\right)^2\right]$$

משפט. לכל אומד T עבור פרמטר θ מתקיים כי

$$\mathrm{MSE}\left(T\right)=\mathrm{Var}\left(T\right)+\left(E\left[T\right]-\theta\right)^2$$

מסקנה. אם T **אומד בלתי מוטה** עבור פרמטר θ מתקיים כי $\mathrm{MSE}\left(T\right)=\mathrm{Var}\left(T\right)$.

רווח סמך

הגדרה. יהא $X_1,\ldots,X_n$ מדגם מקרי מהתפלגות נורמלית $N\left(\mu,\sigma^2\right)$ כאשר σ^2 ידועה בהינתן שהערך הממוצע של מדגם הינו $\overline{x_n}$, אז **רווח סמך** עבור μ ברמת בטחון $1-\alpha$ הוא:

$$\left[\overline{x_n}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{x_n}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

הערה. אם n מספיק גדול הנוסחה נכונה לכל התפלגות, אם σ^2 ידועה.

טענה (**תכונות של רווח סמך**).

1. **רמת הסמך:**
$$P\left(\mu \in \left[\overline{x_n} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{x_n} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]\right) = 1-\alpha$$

2. **אורך רווח הסמך** הוא $L = 2z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$.

3. **הסטייה המקסימלית:**
$$P\left(z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \geq |\overline{x_n} - \mu|\right) = 1-\alpha$$

טענה (**תכונות אורך הרווח**).
1. ככל ש-σ קטנה, L קטן.
2. ככל ש-n גדל, L קטן.
3. ככל ש-$1-\alpha$ קטן, L קטן.

טענה. נתונה אוכלוסייה המפולגת נורמלית עם סטיית תקן ידועה σ, ותוחלת לא ידועה μ. אז בהינתן $d \geq 0$, אם $n \geq \left(z_{1-\frac{\alpha}{2}}\frac{\sigma}{d}\right)^2$ אז

$$P\left(d \geq z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

בדיקת השערות

מושגים בסיסיים

הגדרה. בבדיקת השערות מעלים תמיד שתי השערות המסומנות H_0 ו-H_1.
H_0 היא השערת המחקר או "ההשערה הסקפטית".
H_1 היא הטענה אותה רוצה הי החוקר להוכיח (ההנחה ה"חדשנית").
על סמך נתוני המדגם צריך להחליט אילו מההשערות נכונה.

הגדרה. בבדיקת השערה ישנן שתי החלטות אפשריות:

1. דחיית H_0 - זוהי ההחלטה האומרת שהשערת האפס איננה נכונה, זאת אומרת ש-H_1כנראה נכונה.
2. אי דחיית H_0 - זוהי ההחלטה האומרת שלא ניתון לדחות את השערת האפס והיא כנראה נכונה.

הגדרה. כלל שקובע מתי H_0 נכונה ומתי H_1 נכונה נקרא **כלל החלטה** או **מבחן**.

הגדרה. **אזור הדחייה** C הוא התחום של ערכי המדגם בו דוחים את H_0.
אזור הקבלה/אי הדחייה $\overline{C}$ הוא התחום שבו אין דוחים את H_1 והוא המשלים של C.

הגדרה. בכל מקרה של בדיקת השערות אנו עלולים להגיע להחלטה שגויה. ישנן שתי טעויות אפשריות ונרצה לדעת מה ההסתברות של כל אחת מהן.

מציאות\החלטה	H_0 נכונה	H_1 נכונה
H_0 נכונה	**רמת ביטחון**	**טעות מסוג 2**
H_1 נכונה	**טעות מסוג 1**	**עוצמה**

טענה. בבדיקת השערות:

- ההסתברות לטעות מסוג ראשון היא α.
- ההסתברות לטעות משני היא β.
- עוצמת המבחן היא $1-\beta$.
- רמת הביטחון היא $1-\alpha$.

בדיקת השערות על תוחלת כששונות ידועה

הגדרה. מובהקות התוצאה $p-\text{value}$ הינה ההסתברות לקבלת תוצאה קיצונית לפחות כמו זו שהתקבלה בניסוי, בהנחה שהשערת האפס נכונה. נסמן הסתברות זו ב-PV.

אם $\text{PV} \leq \alpha$ דוחים את H_0, אחרת לא דוחים את H_0.

משפט. כאשר השונות ידועה, ואנו רוצים לבדוק שתי השערות פשוטות על התוחלת μ:

$$H_0 : \mu = \mu_0 \;\; H_1 : \mu = \mu_1$$

אזי מבחן בעל עוצמה מקסימלית הוא המבחן הבא:

1. אם $\mu_1 > \mu_0$ אז אזור הדחיה הוא מהצורה:

$$C = \left\{X_1, \dots, X_n \left| \overline{X} > c \right.\right\}$$

2. אם $\mu_1 < \mu_0$ אז אזור הדחיה הוא מהצורה:

$$C = \left\{X_1, \dots, X_n \left| \overline{X} < c \right.\right\}$$

משפט (בדיקת השערות חד-צדדית). נתונות n תצפיות מהתפלגות נורמלית $X_1, \dots, X_n \sim N\left(\mu, \sigma^2\right)$ כאשר σ ידועה.

1. מבחן בעל עוצמה מקסימלית ברמה מובהקות α, לבדיקת השערת האפס $H_0 : \mu = \mu_0$ כנגד האלטרנטיבה $H_1 : \mu < \mu_0$, הנו המבחן C הדוחה את H_0 כאשר
$$\overline{X} < \mu_0 - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

2. מבחן בעל עוצמה מקסימלית ברמה מובהקות α, לבדיקת השערת האפס $H_0 : \mu = \mu_0$ כנגד האלטרנטיבה $H_1 : \mu > \mu_0$, הנו המבחן C הדוחה את H_0 כאשר
$$\overline{X} > \mu_0 + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

משפט (בדיקת השערות דו-צדדיות). נתונות n תצפיות מהתפלגות נורמלית $X_1,\dots,X_n\sim N\left(\mu,\sigma^2\right)$ כאשר σ ידועה.
אם $H_0:\mu=\mu_0$ ו-$H_1:\mu\neq\mu_0$ אז לא קיים מבחן בעל עוצמה מקסימלית לכל μ. אך נהוג לדחות את H_0 תחת אזור הדחיה:

$$C=\mathbb{R}\setminus\left[\overline{x_n}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{x_n}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$