הסתברות למדעים: דף נוסחאות

גרסה 1.6, ינואר 2011

ברק שושני
baraksh@gmail.com | http://baraksh.co.il/

1 הגדרות בסיסיות

1.1 מרחבי הסתברות

מרחב הסתברות $(\Omega,\mathcal{F},P)$ מורכב מ**מרחב מדגם** Ω שהוא קבוצת כל התוצאות האפשריות של הניסוי, **שדה מאורעות** $\mathcal{F}$ שהוא σ־אלגברה ומייצג את כל השאלות שאפשר לשאול על הניסוי, ו**מידת הסתברות** P.

σ־אלגברה היא קבוצה לא ריקה $\mathcal{F}$ של קבוצות חלקיות ל־Ω כך שמתקיים:
$\emptyset,\Omega\in\mathcal{F}$, $A\in\mathcal{F}\Longrightarrow A^C\in\mathcal{F}$ (סגירות ביחס ללקיחת משלים של קבוצה),
$\{A_n\}\subseteq\mathcal{F}\Longrightarrow\bigcup A_n\in\mathcal{F}$ (סגירות ביחס לאיחוד בן־מניה של קבוצות).

מידת הסתברות היא פונקציה $P:\mathcal{F}\to\mathbb{R}$ כך שמתקיים: $P(\Omega)=1$, אם $A\in\mathcal{F}$ אז $P(A)\geq 0$, אם $\{X_n\}$ סדרה בת־מנייה של קבוצות זרות בזוגות, אז $P\left(\bigcup X_n\right)=\sum P(X_n)$.

1.2 קבוצות ופונקציות בורל

כל קבוצה שניתן ליצור אותה ממספר סופי או בן־מנייה של קטעים (סגורים, פתוחים, חסומים או אינסופיים) באמצעות מספר סופי או בן־מנייה של פעולות איחוד, חיתוך והפרש (או משלים) של קבוצות, היא **קבוצת בורל**. פורמלית, זוהי כל קבוצה שנמצאת ב־**σ־אלגברה של בורל**, שהיא ה־σ־אלגברה המינימלית שמכילה את כל הקטעים הפתוחים (או הסגורים).

פונקציית בורל היא פונקציה $\varphi:\mathbb{R}\to\mathbb{R}$ כך שלכל קבוצת בורל $B\subseteq\mathbb{R}$ הקבוצה:
$$\varphi^{-1}(B)=\{x\in\mathbb{R}:\varphi(x)\in B\}$$
גם היא קבוצת בורל. בפרט, פונקציות רציפות ופונקציות מונוטוניות הן פונקציות בורל. אם $X:\Omega\to\mathbb{R}$ הוא משתנה מקרי אז גם $\varphi(X)$ הוא משתנה מקרי.

1.3 $\limsup$ ו־$\liminf$

עבור סדרות של מספרים, מגדירים:
$$\limsup x_n=\lim_{n\to\infty}\left(\sup_{m\geq n}x_m\right),\qquad \liminf x_n=\lim_{n\to\infty}\left(\inf_{m\geq n}x_m\right)$$
בדומה, עבור סדרות של מאורעות נגדיר:
$$\limsup A_n=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m,\qquad \liminf A_n=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}A_m$$
במלים, $\limsup A_n$ היא קבוצת האיברים $\omega\in\Omega$ עבורם לכל $n\in\mathbb{N}$ קיים $m\geq n$ כך ש־$\omega\in A_m$, או האיברים המופיעים באינסוף קבוצות A_m. (לעתים רושמים i.o. - infinitely often). בדומה, $\liminf A_n$ היא קבוצת האיברים $\omega\in\Omega$ עבורם קיים $n\in\mathbb{N}$ כך שלכל $m\geq n$ מתקיים $\omega\in A_m$, או האיברים המופיעים בכל הקבוצות A_m החל ממקום מסוים. (לעתים רושמים eventually). נשים לב כי $\left(\liminf A_n\right)^C=\limsup A_n^C$.

1.4 גבול של סדרות עולות ויורדות של קבוצות

$$A_n\subseteq A_{n+1}\Longrightarrow\lim_{n\to\infty}A_n=\bigcup_{n=1}^{\infty}A_n,\quad A_n\supseteq A_{n+1}\Longrightarrow\lim_{n\to\infty}A_n=\bigcap_{n=1}^{\infty}A_n$$

1.5 אינדיקטורים

תהי $\mathcal{I}_A:\Omega\to\mathbb{R}$ **פונקציית אינדיקטור** של המאורע $A\subseteq\Omega$:
$$\mathcal{I}_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$$
אם $\{A_n\}_n$ היא סדרה של תת־קבוצות של Ω אז מתקיים:
$$\mathcal{I}_{\liminf A_n}=\liminf\mathcal{I}_{A_n},\qquad \mathcal{I}_{\limsup A_n}=\limsup\mathcal{I}_{A_n}$$
נציין גם את התכונות הבאות של אינדיקטורים:
$$\mathcal{I}_{A\cap B}=\min\{\mathcal{I}_A,\mathcal{I}_B\},\qquad \mathcal{I}_{A\cup B}=\max\{\mathcal{I}_A,\mathcal{I}_B\}$$
$$\mathcal{I}_{\bigcap A_n}=\inf\mathcal{I}_{A_n},\qquad \mathcal{I}_{\bigcup A_n}=\sup\mathcal{I}_{A_n}$$

1.6 כמה תוצאות מתורת המידה

משפט ההתכנסות הדומיננטית של לבג: תהי $\{f_n\}$ סדרת פונקציות מדידות אשר מתכנסת נקודתית לפונקציה f כמעט בכל מקום. אם קיימת פונקציה אינטגרבילית g כך ש־$|f_n|\leq g$ לכל n, אז גם f אינטגרבילית ומתקיים:
$$\int f\,\mathrm{d}\mu=\lim_{n\to\infty}\int f_n\,\mathrm{d}\mu$$
קבוצת קנטור היא קבוצה אשר עוצמתה היא עוצמת הרצף, אך מידת לבג שלה היא אפס. כדי ליצור אותה מתחילים עם הקטע $[0,1]$, מסירים את השליש האמצעי שלו, מסירים את השליש האמצעי של שני הקטעים שנוצרו, וממשיכים בתהליך עד אינסוף.

* טענה ממבחן (1): תהי P פונקציית הסתברות לבג המוגדרת על $[0,1]$. אזי לכל $\varepsilon>0$ ולכל קבוצת בורל A יש B שהוא איחוד סופי של קטעים סגורים כך ש־$P\left((A\backslash B)\cup(B\backslash A)\right)<\varepsilon$.
* טענה ממבחן (2): תהי P פונקציית הסתברות המוגדרת על ה־σ־שדה $\mathcal{F}$. ידוע ש־$\mathcal{F}_0$ הוא שדה שיוצר את $\mathcal{F}$. אזי לכל $\varepsilon>0$ ולכל $A\in\mathcal{F}$ יש $B\in\mathcal{F}_0$ כך ש־$P\left((A\backslash B)\cup(B\backslash A)\right)<\varepsilon$.

2 משתנים מקריים

2.1 הגדרות

משתנה מקרי הוא פונקציה $X:\Omega\to\mathbb{R}$ כך ש:
$$\forall x\in\mathbb{R}:\{\omega\in\Omega:X(\omega)\leq x\}\in\mathcal{F}$$
או במלים, לכל מספר ממשי x המאורע $X\leq x$ נמצא בשדה $\mathcal{F}$. **ה־σ־אלגברה הנוצרת ע"י** X היא:
$$\mathcal{F}_X=\sigma(X)=X^{-1}(\mathcal{B})=\left\{X^{-1}(B):B\in\mathcal{B}\right\}$$
כאשר $\mathcal{B}$ היא ה־σ־אלגברה של בורל.

פונקציית ההתפלגות המצטברת $F_X:\mathbb{R}\to[0,1]$ של המשתנה המקרי X מוגדרת כך:
$$F_X(x)=P(X\leq x)=P\left(\{\omega\in\Omega:X(\omega)\leq x\}\right)$$
היא מונוטונית לא־יורדת, שואפת ל־0 ב־$-\infty$ ול־1 ב־$+\infty$, ורציפה מימין בכל נקודה:
$$F_X(x+)=\lim_{t\downarrow x}F_X(t)=F_X(x)$$
$$F_X(x-)=\lim_{t\uparrow x}F_X(t)=P(X<x)$$
פונקציית הצפיפות $f_X:\mathbb{R}\to[0,\infty)$ של המשתנה המקרי X היא פונקציה המקיימת עבור $x,a,b\in\mathbb{R}$ וקבוצת בורל $A\subseteq\mathbb{R}$:
$$F_X(x)=\int_{-\infty}^{x}f_X(t)\,\mathrm{d}t,\qquad P(x\in A)=\int_A f_X(x)\,\mathrm{d}x$$
$$P(a\leq X\leq b)=F_X(b)-F_X(a)=\int_a^b f_X(x)\,\mathrm{d}x$$
על הפונקציה להיות מנורמלת:
$$\int_{-\infty}^{+\infty}f_X(x)\,\mathrm{d}x=1$$
אם קיימת פונקציית צפיפות רציפה למקוטעין, אז פונקציית ההתפלגות F_X גזירה ברציפות למקוטעין, ומתקיים $f_X(x)=F_X'(x)$.

אטום של התפלגות (לא רציפה) של משתנה מקרי X הוא מספר $x\in\mathbb{R}$ המקיים:
$$P(X=x)=F(x)-F(x-)>0$$
מספר $x\in\mathbb{R}$ נקרא **אחוזון** p של משתנה מקרי X אם:
$$F_X(x-)=P(X<x)\leq p\leq P(X\leq x)=F_X(x)$$
האחוזון $p=\frac{1}{2}$ נקרא **חציון**.
פונקציה $X^*:(0,1)\to\mathbb{R}$ נקראת **פונקציית אחוזון** של משתנה מקרי X אם $X^*(p)$ הוא אחוזון p של X לכל $p\in(0,1)$. אם F_X הפיכה, אז $X^*=F_X^{-1}$.

2.2 טרנספורמציות

תהי $\varphi:\mathbb{R}\to\mathbb{R}$ פונקציה <u>חד־חד־ערכית</u> ונגדיר $Y=\varphi(X)$. אז מתקיים:
$$f_Y(y)=f_X\left(\varphi^{-1}(y)\right)\left|\frac{\mathrm{d}}{\mathrm{d}y}\varphi^{-1}(y)\right|$$
כמו כן, אם φ מונוטונית עולה $F_Y(y)=F_X\left(\varphi^{-1}(y)\right)$, ואם היא מונוטונית יורדת $F_Y(y)=1-F_X\left(\varphi^{-1}(y)\right)$. אם φ איננה חד־חד־ערכית, אין נוסחה פשוטה. ניתן דוגמה עבור $Y=\varphi(X)=X^2$. פונקציית ההתפלגות המצטברת היא:
$$F_Y(y)=P\left(X^2\leq y\right)=F_X(\sqrt{y})-F_X(-\sqrt{y})$$
לכן, מכלל השרשרת:
$$f_Y(y)=F_Y'(y)=\frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y})+f_X(-\sqrt{y})\right)$$

3 תוחלת ושונות

3.1 תוחלת

יהי X משתנה מקרי. במקרה הבדיד, **התוחלת** של X היא:
$$\mathrm{E}(X)=\sum_k kP(X=k)$$
והיא קיימת רק אם הטור <u>מתכנס בהחלט</u>. במקרה הרציף, התוחלת היא:
$$\mathrm{E}(X)=\int_{-\infty}^{+\infty}xf_X(x)\,\mathrm{d}x=\int_0^1X^*(p)\,\mathrm{d}p$$
$$=\int_0^{+\infty}(1-F_X(x))\,\mathrm{d}x-\int_{-\infty}^0F_X(x)\,\mathrm{d}x$$
$$=\int_0^{+\infty}P(X>x)\,\mathrm{d}x-\int_{-\infty}^0P(X\leq x)\,\mathrm{d}x$$
והיא קיימת רק אם האינטגרל $\int_{-\infty}^{+\infty}xf_X(x)\,\mathrm{d}x$ <u>מתכנס בהחלט</u>.
התוחלת היא מונוטונית: אם $Y\leq X$ אז $\mathrm{E}(X)\leq\mathrm{E}(Y)$, ולינארית: $\mathrm{E}(aX+bY)=a\,\mathrm{E}(X)+b\,\mathrm{E}(Y)$.
אם $\varphi:\mathbb{R}\to\mathbb{R}$ פונקציית בורל אז:
$$\mathrm{E}(\varphi(X))=\int_0^1\varphi(X^*(p))\,\mathrm{d}p=\int_{-\infty}^{+\infty}\varphi(x)f_X(x)\,\mathrm{d}x$$
אם X סימטרי ביחס למספר $a\in\mathbb{R}$, כלומר $X^*(p+)=2a-X^*((1-p)-)$ או $F_x(x)=1-F_X((2a-x)-)$ או פשוט $f_X(x)=f_X(2a-x)$, <u>והתוחלת של X קיימת</u>, אז $\mathrm{E}(X)=a$.

3.2 שונות

השונות של X היא:
$$\mathrm{Var}(X)=\mathrm{E}\left((X-\mathrm{E}(X))^2\right)=\mathrm{E}\left(X^2\right)-(\mathrm{E}(X))^2$$
השונות מעלה קבועים בריבוע, ואינה מושפעת מהזזות: $\mathrm{Var}(aX+b)=a^2\,\mathrm{Var}(X)$. **סטיית התקן** היא השורש של השונות: $\sigma_X=\sqrt{\mathrm{Var}(X)}$.

3.3 פונקציה יוצרת מומנטים

באופן כללי הביטוי $\mathrm{E}(X^n)$ נקרא "**המומנט** ה־n של X"; התוחלת היא המומנט הראשון. אם התוחלת $\mathrm{E}\left(\mathrm{e}^{tX}\right)$ קיימת בסביבה כלשהי של $t=0$, אז **הפונקציה יוצרת המומנטים** היא:
$$M_X(t)=\mathrm{E}\left(\mathrm{e}^{tX}\right)=\int_{-\infty}^{+\infty}\mathrm{e}^{tx}f_X(x)\,\mathrm{d}x$$
מתקיים:
$$M_X(0)=1,\qquad M_X'(0)=\mathrm{E}(X),\qquad M_X''(0)=\mathrm{E}\left(X^2\right)$$
ועבור צירוף לינארי של משתנים מקריים בלתי־תלויים (אך לא בהכרח שווי־התפלגות) מתקיים $M_{aX+bY}(t)=M_X(at)M_Y(bt)$.

3.4 תוחלת ושונות של אינדיקטור

יהי $A\in\mathcal{F}$ מאורע, ונגדיר אינדיקטור $\mathcal{I}_A:\Omega\to\mathbb{R}$:
$$\mathcal{I}_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$$
האינדיקטור הוא בעצם משתנה מקרי, ומתקיים:
$$\mathrm{E}(\mathcal{I}_A)=\int_\Omega\mathcal{I}_A(\omega)\,\mathrm{d}P(\omega)=\int_A\mathrm{d}P(\omega)=P(A)$$
$$\mathrm{Var}(\mathcal{I}_A)=P(A)(1-P(A))$$
בפרט, אם $A\subseteq\mathbb{R}$ קבוצת בורל ו־X משתנה מקרי, נגדיר אינדיקטור $\mathcal{I}_A:\mathbb{R}\to\mathbb{R}$:
$$\mathcal{I}_A(x)=\begin{cases}1 & x\in A\\ 0 & x\notin A\end{cases}$$
ואז:
$$\mathrm{E}(\mathcal{I}_A(X))=\int_{\mathbb{R}}\mathcal{I}_A(x)f_X(x)\,\mathrm{d}x=\int_Af_X(x)\,\mathrm{d}x=P(X\in A)$$

4 מספר משתנים מקריים

4.1 הגדרות

משתנה מקרי דו־מימדי הוא פונקציה $(X,Y):\Omega\to\mathbb{R}^2$ כך ש:
$$\forall(x,y)\in\mathbb{R}^2:\{\omega\in\Omega:X(\omega)\leq x,Y(\omega)\leq y\}\in\mathcal{F}$$
או במלים, לכל וקטור (x,y) המאורע שבו $X\leq x$ וגם $Y\leq y$ נמצא בשדה $\mathcal{F}$.
פונקציית ההתפלגות המצטברת המשותפת $F_{X,Y}:\mathbb{R}^2\to[0,1]$ מוגדרת כך:
$$F_{X,Y}(x,y)=P(X\leq x,Y\leq y)$$
גבולותיה מקיימים:
$$F_{X,Y}(x,y)\to\begin{cases}0 & x\to-\infty\text{ or }y\to-\infty\\ 1 & x\to+\infty\text{ and }y\to+\infty\end{cases}$$
היא רציפה מימין: $F_{X,Y}(x+,y+)=F_{X,Y}(x,y)$, ועבור $x_1<x_2,y_1<y_2$ מתקיים:
$$P(x_1<x<x_2,y_1<y<y_2)=$$
$$=F_{X,Y}(x_2,y_2)+F_{X,Y}(x_1,y_1)-F_{X,Y}(x_2,y_1)-F_{X,Y}(x_1,y_2)$$
פונקציית הצפיפות המשותפת $f_{X,Y}:\mathbb{R}^2\to[0,\infty)$ היא פונקציה המקיימת:
$$P((X,Y)\in B)=\iint_Bf_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$
לכל קבוצת בורל $B\subseteq\mathbb{R}^2$. על הפונקציה להיות מנורמלת:
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y=1$$
ומתקיים הקשר:
$$F_{X,Y}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f_{X,Y}(u,v)\,\mathrm{d}u\,\mathrm{d}v$$
$$f_{X,Y}(x,y)=\frac{\partial^2}{\partial x\partial y}F_{X,Y}(x,y)$$
אטום של התפלגות (לא רציפה) של משתנה מקרי דו־מימדי (X,Y) הוא וקטור $(x,y)\in\mathbb{R}^2$ המקיים $P((X,Y)=(x,y))>0$.
פונקציות ההתפלגות המצטברת השוליות הן:
$$F_X(x)=\lim_{y\to\infty}F_{X,Y}(x,y),\qquad F_Y(y)=\lim_{x\to\infty}F_{X,Y}(x,y)$$
פונקציות הצפיפות השוליות הן:
$$f_X(x)=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\,\mathrm{d}y,\qquad f_Y(y)=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\,\mathrm{d}x$$

4.2 תלות ואי־תלות

התנאים הבאים שקולים:
* שני משתנים מקריים $X,Y:\Omega\to\mathbb{R}$ הם בלתי־תלויים,
* $P(X\in A,Y\in B)=P(X\in A)P(Y\in B)$ לכל שתי קבוצות בורל $A,B\subseteq\mathbb{R}$,
* $F_{X,Y}(x,y)=F_X(x)F_Y(y)$,
* $f_{X,Y}(x,y)=f_X(x)f_Y(y)$ (כמעט לכל x,y).

אם X,Y בלתי־תלויים אז:
* $\mathrm{E}(XY)=\mathrm{E}(X)\,\mathrm{E}(Y)$,
* $\mathrm{E}(g(X)h(Y))=\mathrm{E}(g(X))\,\mathrm{E}(h(Y))$ לכל שתי פונקציות בורל חסומות $g,h:\mathbb{R}\to\mathbb{R}$,
* $g(X),h(Y)$ בלתי־תלויים לכל שתי פונקציות בורל $g,h:\mathbb{R}\to\mathbb{R}$.

4.3 שונות משותפת

יהיו X,Y משתנים מקריים. אז השונות המשותפת שלהם היא:
$$\mathrm{Cov}(X,Y)=\mathrm{E}((X-\mathrm{E}(X))(Y-\mathrm{E}(Y)))=\mathrm{E}(XY)-\mathrm{E}(X)\,\mathrm{E}(Y)$$
מתקיים:
$$\mathrm{Cov}(X,a)=0,\quad\mathrm{Cov}(X,X)=\mathrm{Var}(X),\quad\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$$
$$\mathrm{Cov}(aX,bY)=ab\,\mathrm{Cov}(X,Y)$$
$$\mathrm{Cov}(X+Y,Z)=\mathrm{Cov}(X,Z)+\mathrm{Cov}(Y,Z)$$
$$\mathrm{Cov}\left(\sum_{i=1}^nX_i,\sum_{j=1}^nY_j\right)=\sum_{i=1}^n\sum_{j=1}^n\mathrm{Cov}(X_i,Y_j)$$
אם המשתנים X,Y הם בלתי־תלויים אז $\mathrm{Cov}(X,Y)=0$, אבל ההפך אינו בהכרח נכון. כמו כן מתקיים הקשר:
$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\,\mathrm{Cov}(X,Y)$$
ובאופן כללי:
$$\mathrm{Var}\left(\sum_{i=1}^nX_i\right)=\sum_{i=1}^n\mathrm{Var}(X_i)+2\sum_{i<j}\mathrm{Cov}(X_i,X_j)$$
בפרט, אם המשתנים X_i בלתי־תלויים מתקיים $\mathrm{Var}\left(\sum X_i\right)=\sum\mathrm{Var}(X_i)$.

4.4 מקדם מתאם

מקדם המתאם של X,Y הוא:
$$\rho_{X,Y}=\frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$
מתקיים תמיד $-1\leq\rho_{X,Y}\leq 1$, כאשר 1 משמעותו שמתקיים קשר לינארי חיובי בין המשתנים, -1 משמעותו שקיים קשר לינארי שלילי ו־0 משמעותו שאין קשר לינארי.

4.5 קונבולוציה

יהיו U,V משתנים מקריים בלתי־תלויים בעלי פונקציות צפיפות f_U,f_V. אז פונקציית הצפיפות של $U+V$ היא **הקונבולוציה**:
$$f_{U+V}(x)=\int_{-\infty}^{+\infty}f_U(y)f_V(x-y)\,\mathrm{d}y$$

5 התניה

5.1 התניה של מאורעות

יהיו A,B מאורעות. אם $P(B)>0$ אז **ההסתברות המותנית** של A בהינתן B היא:
$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(B|A)P(A)}{P(B)}$$
(השוויון הימני נקרא **נוסחת בייס**).
אם $\{A_i\}$ היא קבוצה סופית או בת־מנייה של מאורעות זרים בזוגות אשר האיחוד שלהם הוא כל מרחב המדגם, אז מתקיימת **נוסחת ההסתברות השלמה**:
$$P(B)=\sum_iP(B\cap A_i)=\sum_iP(B|A_i)P(A_i)$$
הנוסחה עובדת גם במקרה של התניה "כפולה":
$$P(B|C)=\sum_iP(B\cap A_i|C)$$
המאורעות A,B הם בלתי־תלויים אם ורק אם $P(A|B)=P(A)$ ו־$P(B|A)=P(B)$.

5.2 התניה של משתנים מקריים

יהיו X,Y משתנים מקריים בעלי צפיפות משותפת $f_{X,Y}$. **פונקציית הצפיפות המותנית** של Y בהינתן $X=x$ היא:
$$f_{Y|X=x}(y)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{f_{X|Y=y}(x)f_Y(y)}{f_X(x)}$$
פונקציית ההתפלגות המותנית היא:
$$F_{Y|X=x}(y)=P(Y\leq y|X=x)=\int_{-\infty}^{y}f_{Y|X=x}(v)\,\mathrm{d}v$$
התוחלת המותנית היא:
$$\mathrm{E}(Y|X=x)=\int_{-\infty}^{+\infty}yf_{Y|X=x}(y)\,\mathrm{d}y$$
והשונות המותנית היא:
$$\mathrm{Var}(Y|X=x)=\mathrm{E}\left((Y-\mathrm{E}(Y|X=x))^2\Big|X=x\right)$$
$$=\mathrm{E}\left(Y^2\big|X=x\right)-(\mathrm{E}(Y|X=x))^2$$
כל אלה הן <u>פונקציות של מספר כלשהו, x</u>. ניתן גם להחליף את $X=x$ ב־X ולקבל <u>משתנים מקריים</u> $f_{Y|X},F_{Y|X},\mathrm{E}(Y|X),\mathrm{Var}(Y|X)$ אשר תלויים במשתנה המקרי X.

5.3 התניה בטרנספורמציה של משתנה מקרי

תהי f פונקציה ממשית חד־חד־ערכית. אז מתקיים:
$$\mathrm{E}(X|Y)(\omega)=\mathrm{E}(X|Y=Y(\omega))=\mathrm{E}(X|f(Y)=f(Y(\omega)))$$
$$=\mathrm{E}(X|f(Y)=f(Y)(\omega))=\mathrm{E}(X|f(Y))(\omega)$$
לפיכך $\mathrm{E}(X|Y)=\mathrm{E}(X|f(Y))$.

5.4 נוסחת התוחלת השלמה

נוסחת התוחלת השלמה היא:
$$\mathrm{E}(Y)=\mathrm{E}(\mathrm{E}(Y|X))=\int_{-\infty}^{+\infty}\mathrm{E}(Y|X=x)f_X(x)\,\mathrm{d}x$$
כאשר X יכול להיות כל משתנה מקרי. אם Y הוא פונקציה של X, אז ניתן להוציא את X מחוץ לתוחלת הפנימית - ברגע שאנו יודעים את X, הוא הופך מבחינתנו לקבוע. לכן למשל:
$$\mathrm{E}(XY)=\mathrm{E}(\mathrm{E}(XY|X))=\mathrm{E}(X\,\mathrm{E}(Y|X))$$
כמו כן, עבור אינדיקטור $\mathcal{I}_A$ מתקיים:
$$P(A)=\mathrm{E}(\mathcal{I}_A)=\mathrm{E}(\mathrm{E}(\mathcal{I}_A|X))=\mathrm{E}(P(A|X))$$
לכן, עבור $A=\{Y\leq y\}$:
$$P(Y\leq y)=\int_{-\infty}^{+\infty}P(Y\leq y|X=x)f_X(x)\,\mathrm{d}x$$
באותו אופן, עבור שונות מותנית מתקיימת **נוסחת השונות השלמה**:
$$\mathrm{Var}(X)=\mathrm{E}(\mathrm{Var}(X|Y))+\mathrm{Var}(\mathrm{E}(X|Y))$$
נזכיר גם את **נוסחת השונות המשותפת השלמה**:
$$\mathrm{Cov}(X,Y)=\mathrm{E}(\mathrm{Cov}(X,Y|Z))+\mathrm{Cov}(\mathrm{E}(X|Z),\mathrm{E}(Y|Z))$$

5.5 שילוב של משתנים רציפים ובדידים

אם X רציף ו־Y בדיד, מתקיימים הקשרים:
$$P(Y=y|X=x)=\frac{f_{X|Y=y}(x)P(Y=y)}{f_X(x)}$$
$$P(Y=y)=\int_{-\infty}^{+\infty}P(Y=y|X=x)f_X(x)\,\mathrm{d}x$$
$$f_X(x)=\sum_yP(Y=y)f_{X|Y=y}(x)$$
$$\mathrm{E}(Y|X=x)=\sum_yyP(Y=y|X=x)$$

5.6 סכום משתנים באורך מקרי

יהי N משתנה מקרי המקבל ערכים שלמים אי־שליליים ויהיו X_i משתנים מקריים בלתי־תלויים, שווי־התפלגות ובלתי־תלויים ב־N, אז מתקיים:

$$\mathrm{E}\left(\sum_{i=1}^{N} X_i\right) = \mathrm{E}(N)\,\mathrm{E}(X_i)$$

$$\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) = \mathrm{E}(N)\,\mathrm{Var}(X_i) + [\mathrm{E}(X_i)]^2\,\mathrm{Var}(N)$$

6 אי־שוויונים וקירובים

6.1 אי־שוויון מרקוב

יהי X משתנה מקרי <u>אי־שלילי</u>, אז לכל $a > 0$ מתקיים:

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

ניתן לבצע הזזה של המשתנה, כלומר $P(X - b \geq a - b)$, כדי לשפר את החסם - כל עוד $X - b$ אי־שלילי. כמו כן, במקרה של מספרים שלמים (בהתפלגות בדידה), $P(X > a) = P(X \geq a + 1)$.

6.2 אי־שוויון צ'בישב

יהי X משתנה מקרי בעל שונות $\mathrm{Var}(X) = \sigma^2$, אז לכל $k > 0$ מתקיים:

$$P(|X - \mathrm{E}(X)| \geq k) \leq \frac{\sigma^2}{k^2}$$

קיימת גם **גרסה חד־צדדית** של שוויון זה:

$$P(X - \mathrm{E}(X) \geq k) \leq \frac{\sigma^2}{\sigma^2 + k^2}$$

אם משתנה מקרי הוא **סימטרי סביב התוחלת**, כלומר $P(X \leq \mathrm{E}(X) + x) = P(X \geq \mathrm{E}(X) - x)$ לכל $x \in \mathbb{R}$, אז ניתן לרשום:

$$P(X - \mathrm{E}(X) \geq k) = \frac{1}{2} P(|X - \mathrm{E}(X)| \geq k)$$

6.3 אי־שוויון צ'רנוף

יהי X משתנה מקרי ותהי $M_X(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right)$ הפונקציה יוצרת המומנטים שלו. אז מתקיים (מאי־שוויון מרקוב):

$$P(X \geq a) \leq \mathrm{e}^{-ta} M_X(t), \qquad t > 0$$

$$P(X \leq a) \leq \mathrm{e}^{-ta} M_X(t), \qquad t < 0$$

באמצעות מציאת ה־t (החיובי או השלילי לפי המקרה) שנותן את הערך הנמוך ביותר לפונקציה $\mathrm{e}^{-ta} M_X(t)$, נוכל למצוא את החסם הטוב ביותר.

6.4 אי־שוויון ינסן

אם $\varphi : \mathbb{R} \to \mathbb{R}$ פונקציה קמורה, כלומר $\varphi''(x) \geq 0$ לכל $x \in \mathbb{R}$, אז מתקיים:

$$\mathrm{E}(\varphi(X)) \geq \varphi(\mathrm{E}(X))$$

6.5 קירוב לינארי

יהיו X, Y משתנים מקריים. התחזית הלינארית האופטימלית, $\hat{Y}$, של Y לפי X ניתנת באמצעות הנוסחה:

$$\frac{\hat{Y} - \mathrm{E}(Y)}{\sigma_Y} = \rho_{X,Y} \frac{X - \mathrm{E}(X)}{\sigma_X}$$

ביטוי מפורש ל־$\hat{Y}$ הוא $\hat{Y} = aX + b$ כאשר:

$$a = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}, \qquad b = \mathrm{E}(Y) - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} \mathrm{E}(X)$$

השגיאה הריבועית הממוצעת (MSE) של התחזית $\hat{Y}$ היא:

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

והיא מספקת מדד למידת הדיוק של הקירוב הלינארי. אם $\mathrm{MSE}\left(\hat{Y}\right) = 0$ הקירוב הוא מדויק.

6.6 אי־שוויון המקסימום של קולמוגורוב

יהי $X_1, \ldots, X_n$ אוסף של משתנים מקריים בלתי־תלויים בעלי תוחלת 0 ושונות סופית. אז מתקיים:

$$P\left(\max_{1 \leq k \leq n} \left|\sum_{i=1}^{k} X_i\right| \geq a\right) \leq \frac{1}{a^2} \sum_{i=1}^{k} \mathrm{Var}(X_i)$$

כמקרה פרטי נקבל עבור $n = 1$ את אי־שוויון צ'בישב.

7 סדרות של משתנים מקריים

7.1 סוגי התכנסות

תהי $\{X_n\}$ סדרת משתנים מקריים. אז:

* X_n **מתכנסת כמעט בוודאות** (או **כמעט תמיד**) ל־X אם:

$$P\left(X_n \xrightarrow[n\to\infty]{} X\right) = 1$$

* X_n **מתכנסת בהסתברות** (או **מתכנסת חלש**) ל־X אם לכל $\varepsilon > 0$:

$$P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0$$

* X_n **מתכנסת בהתפלגות** ל־X אם לכל $t \in \mathbb{R}$ שהוא נקודת רציפות של F_X:

$$F_{X_n}(t) \xrightarrow[n\to\infty]{} F_X(t)$$

התכנסות כמעט בוודאות גוררת התכנסות בהסתברות, אך לא להפך. התכנסות בהסתברות גוררת התכנסות בהתפלגות, אך לא להפך.
ניתן להראות כי התנאי הבא שקול להתכנסות כמעט בוודאות של X_n ל־X:

$$P(\limsup \{|X_n - X| > \varepsilon\}) = 0$$

יהיו $X_1, X_2, \ldots$ בלתי־תלויים ושווי־התפלגות המקיימים $\mathrm{E}(|X_n|) < \infty$, אז מתקיים:

$$P\left(\frac{X_n}{n} \xrightarrow[n\to\infty]{} 0\right) = 1$$

להוכחת הטענה יש להשתמש בטענה הבאה: אם $Y \geq 0$ אז $\mathrm{E}(Y) < \infty$ אם ורק אם $\sum_{n=0}^{\infty} P(Y > n) < \infty$.

7.2 הלמות של בורל וקנטלי

למה ראשונה - לכל סדרת מאורעות $\{A_n\}$, לא בהכרח בלתי־תלויים, מתקיים:

$$\sum_{n=1}^{\infty} P(A_n) < \infty \quad \Longrightarrow \quad P\left(\limsup_{n\to\infty} A_n\right) = 0$$

למה שנייה - לכל סדרת מאורעות בלתי־תלויים $\{A_n\}$ מתקיים:

$$\sum_{n=1}^{\infty} P(A_n) = \infty \quad \Longrightarrow \quad P\left(\limsup_{n\to\infty} A_n\right) = 1$$

7.3 חוק המספרים הגדולים

תהי $\{X_n\}$ סדרת משתנים מקריים בלתי־תלויים ושווי־התפלגות בעלי תוחלת μ סופית. אז מתקיים:

$$P\left(\left|\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$$

(התכנסות בהסתברות). זהו **החוק החלש** של המספרים הגדולים. **החוק החזק** זהה, אך ההתכנסות היא כמעט בוודאות, כלומר:

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{} \mu\right) = 1$$

למעשה, החוק החזק פועל גם על סדרת משתנים בלתי־תלויים <u>שאינם שווי־התפלגות</u> אך מקיימים את **תנאי קולמוגורוב**:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \mathrm{Var}(X_n) < \infty$$

במקרה זה מתקיים:

$$P\left(\frac{1}{n}\sum_{i=1}^{n} (X_i - \mathrm{E}(X_i)) \xrightarrow[n\to\infty]{} 0\right) = 1$$

7.4 משפט הגבול המרכזי

תהי $\{X_n\}$ סדרת משתנים מקריים בלתי־תלויים ושווי־התפלגות בעלי תוחלת μ ושונות σ^2 סופית ושונה מאפס. אז לכל $a \in \mathbb{R}$ מתקיים:

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{\mathrm{d}} \mathcal{N}(0,1)$$

(התכנסות בהתפלגות), כלומר:

$$P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq a\right) \xrightarrow[n\to\infty]{} \Phi(a)$$

כמו כן, אם $\{a_n\} \subseteq \mathbb{R}$ היא סדרה כך ש־$a_n \to a$ אז:

$$P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq a_n\right) \xrightarrow[n\to\infty]{} \Phi(a)$$

אם ברשותנו טבלת התפלגות נורמלית בעלת ערכים חיוביים בלבד, נוכל להשתמש בנוסחה $\Phi(-a) = 1 - \Phi(a)$. עבור אי שוויון בערך מוחלט מתקיים:

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}\right| \leq a\right) \xrightarrow[n\to\infty]{} \Phi(a) - \Phi(-a) = 2\Phi(a) - 1$$

7.5 הילוך מקרי

תהי $\{X_n\}$ סדרת משתנים מקריים בלתי־תלויים כך ש:

$$P(X_i = 1) = p, \qquad P(X_i = -1) = 1 - p$$

אז לכל $t \in \mathbb{R}$ מתקיים:

$$P\left(\sum_{i=1}^{n} X_i > t\right) \xrightarrow[n\to\infty]{} \begin{cases} 0 & p < \frac{1}{2} \\ \frac{1}{2} & p = \frac{1}{2} \\ 1 & p > \frac{1}{2} \end{cases}$$

7.6 שדה הזנב וחוק 0־1 של קולמוגורוב

תהי $\{X_n\}$ סדרת משתנים מקריים בלתי־תלויים אך לא בהכרח שווי־התפלגות. יהי $\mathcal{F}_n = \sigma(X_n, X_{n+1}, \ldots)$ ה־σ־שדה הנוצר ע"י X_i עבור $i \geq n$. אז **שדה הזנב** הוא:

$$\mathcal{F}_\infty = \bigcap_{n=1}^{\infty} \mathcal{F}_n$$

באופן אינטואטיבי, מאורע הוא בשדה הזנב אם התרחשותו אינה מושפעת משינוי מספר סופי של משתנים בסדרה. להלן מספר דוגמאות למאורעות כאלה:

$$\{\limsup X_n < c\}, \qquad \sum_{n=1}^{\infty} X_n < \infty, \qquad \exists \lim_{n\to\infty} X_n$$

לפי **חוק 0־1 של קולמוגורוב**, הסתברותו של מאורע בשדה הזנב חייבת להיות 0 או 1.

7.7 חוק התכנסות הסכום

תהי $\{X_n\}$ סדרת משתנים מקריים בעלי תוחלת 0 ושונות סופית. אז מתקיים:

$$\sum_{k=1}^{\infty} \mathrm{Var}(X_k) < \infty \Longrightarrow P\left(\sum_{k=1}^{\infty} X_k < \infty\right) = 1$$

8 התפלגויות בדידות

8.1 התפלגות אחידה: $X \sim \mathrm{Uni}(a, \ldots, b)$

X הוא מספר שלם בין a ל־b הנבחר באקראי. יש בסה"כ $n = b-a+1$ אפשרויות וכל אחת בעלת הסתברות שווה, לכן ההתפלגות, התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$P(X = k) = \frac{1}{b-a+1} = \frac{1}{n}, \qquad \mathrm{E}(X) = \frac{a+b}{2}$$

$$\mathrm{Var}(X) = \frac{(b-a+1)^2 - 1}{12} = \frac{n^2 - 1}{12}$$

$$M_X(t) = \frac{\mathrm{e}^{at} - \mathrm{e}^{(b+1)t}}{n(1 - \mathrm{e}^t)}$$

8.2 התפלגות בינומית: $X \sim \mathrm{Bin}(n, p)$

X הוא מספר ההצלחות ב־n ניסויים בלתי־תלויים אשר סיכוי ההצלחה בכל אחד מהם הוא p. נבחר k מתוך n הניסויים. ניסויים אלה הצליחו, בהסתברות p, ו־$n-k$ הניסויים הנותרים נכשלו, בהסתברות $1-p$. לכן ההתפלגות היא:

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

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = np, \quad \mathrm{Var}(X) = np(1-p), \quad M_X(t) = \left(1 - p + p\,\mathrm{e}^t\right)^n$$

והשכיח הוא $\lfloor (n+1)p \rfloor$ או $\lfloor (n+1)p \rfloor - 1$. אם $X \sim \mathrm{Bin}(n,p)$ ו־$Y \sim \mathrm{Bin}(m,p)$ הם בלתי־תלויים, אז מתקיים $X + Y \sim \mathrm{Bin}(n+m, p)$.

8.3 התפלגות בינומית שלילית וגיאומטרית

התפלגות בינומית שלילית: $X \sim \mathrm{NegBin}(n, p)$
X הוא מספר הניסויים עד להצלחה ה־n (כולל), כשהניסויים בלתי־תלויים והסיכוי להצלחה בכל אחד מהם הוא p. נבחר $n-1$ מתוך $k-1$ הניסויים הראשונים. ניסויים אלה הצליחו, בהסתברות p^{n-1}, וה־$k-n$ הנותרים נכשלו, בהסתברות $1-p$. לאחר מכן עלינו להכפיל את התוצאה ב־p משום שהניסוי האחרון הצליח. לכן ההתפלגות היא:

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

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = \frac{n}{p}, \quad \mathrm{Var}(X) = n\frac{1-p}{p^2}, \quad M_X(t) = \left(\frac{p\,\mathrm{e}^t}{1 - (1-p)\,\mathrm{e}^t}\right)^n$$

אם $X \sim \mathrm{NegBin}(n,p)$ ו־$Y \sim \mathrm{NegBin}(m,p)$ הם בלתי־תלויים, אז מתקיים $X + Y \sim \mathrm{NegBin}(n+m, p)$.

התפלגות גיאומטרית: $X \sim \mathrm{Geo}(p) \equiv \mathrm{NegBin}(1, p)$
זהו מקרה פרטי של התפלגות בינומית שלילית כאשר $n = 1$. ההתפלגות היא:

$$P(X = k) = p(1-p)^{k-1}, \qquad P(X \leq k) = 1 - (1-p)^k$$

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = \frac{1}{p}, \quad \mathrm{Var}(X) = \frac{1-p}{p^2}, \quad M_X(t) = \frac{p\,\mathrm{e}^t}{1 - (1-p)\,\mathrm{e}^t}$$

השכיח הוא 1, כי ההסתברות יורדת ככל ש־k עולה. אם $X_i \sim \mathrm{Geo}(p)$ עבור $1 \leq i \leq n$ הם בלתי־תלויים, אז מתקיים:

$$\sum_{i=1}^{n} X_i \sim \mathrm{NegBin}(n, p)$$

ההתפלגות הגיאומטרית היא ההתפלגות הבדידה **היחידה** שהיא "חסרת זיכרון", כלומר, בכל שלב בניסוי יש אותו סיכוי להצלחה ללא תלות בכמות או בתוצאות הניסויים הקודמים (למשל, למטבע שמוטל אין "זיכרון" של ההטלות הקודמות). ניתן לתאר תכונה זו באמצעות הנוסחה:

$$P(X > m + n | X > m) = P(X > n)$$

כלומר, לכל $n, m \in \mathbb{N}$, אם ידוע לנו שכבר נערכו m ניסויים כושלים, ההסתברות ש־n ניסויים נוספים ייכשלו שווה להסתברות ש־n ניסויים ייכשלו אם נתחיל את התהליך מהתחלה. מכאן, אם מצאנו התפלגות המקיימת:

$$P(X = k + 1) = cP(X = k)$$

כאשר c קבוע כלשהו, התפלגות זו בהכרח תהיה גיאומטרית (הקבוע c הוא פשוט ההסתברות לכישלון בניסוי).

8.4 התפלגות היפרגיאומטרית: $X \sim \mathrm{HG}(N, D, n)$

X הוא מספר הפריטים "המיוחדים" שנמצאו בדגימה (ללא החזרה) בגודל n מתוך אוכלוסייה בגודל N הכוללת D פרטים מיוחדים. נבחר k פריטים מיוחדים מתוך D המיוחדים שבאוכלוסייה, ונבחר את $n-k$ הפריטים הלא־מיוחדים שנשארו מתוך $N-D$ הלא־מיוחדים באוכלוסייה. נחלק את מספר האפשרויות בגודל מרחב המדגם, שהוא n פריטים כלשהם שנדגמו מאוכלוסייה בגודל N. לכן ההתפלגות היא:

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

התוחלת והשונות הן:

$$\mathrm{E}(X) = n\frac{D}{N}, \qquad \mathrm{Var}(X) = n\frac{D}{N}\left(1 - \frac{D}{N}\right)\frac{N-n}{N-1}$$

והשכיח הוא $\left\lfloor \frac{(n+1)(D+1)}{N+2} \right\rfloor$.

8.5 התפלגות פואסונית: $X \sim \mathrm{Pois}(\lambda)$

X הוא מספר ההצלחות בתהליך בינומי כאשר $p \to 0$, $n \to \infty$, אך המכפלה $\lambda = np$ קבועה. לכן ההתפלגות משמשת לקירוב של תהליכים בהם יש הסתברות נמוכה להצלחה ומספר רב של ניסיונות, לדוגמה מספר טעויות ההדפסה בספר או מספר האנשים שמגיעים לגיל 100 באוכלוסייה. ההתפלגות, התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$P(X = k) = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}, \quad \mathrm{E}(X) = \mathrm{Var}(X) = \lambda, \quad M_X(t) = \mathrm{e}^{\lambda\left(\mathrm{e}^t - 1\right)}$$

השכיח הוא $\lfloor \lambda \rfloor$, או $\lambda - 1$ אם $\lambda \in \mathbb{Z}$. אם $X \sim \mathrm{Pois}(\lambda)$ ו־$Y \sim \mathrm{Pois}(\mu)$ הם בלתי־תלויים, אז מתקיים $X + Y \sim \mathrm{Pois}(\lambda + \mu)$.

9 התפלגויות רציפות

9.1 התפלגות אחידה: $X \sim \mathrm{Uni}(a, b),\ a, b \in \mathbb{R},\ a < b$

ההתפלגות האחידה מתארת מספר ממשי בין a ל־b אשר נבחר באקראי. פונקציית הצפיפות היא:

$$f_X(x) = \frac{1}{b-a}, \qquad x \in [a,b]$$

פונקציית ההתפלגות המצטברת ופונקציית האחוזון הן:

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}, \qquad F_X^{-1}(p) = a + p(b-a)$$

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = \frac{a+b}{2}, \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}, \qquad M_X(t) = \frac{\mathrm{e}^{bt} - \mathrm{e}^{at}}{t(b-a)}$$

ניתן לעבור להתפלגות הסטנדרטית כך:

$$\frac{X-a}{b-a} \sim \mathrm{Uni}(0,1)$$

כמו כן, אם $X \sim \mathrm{Uni}(0,1)$ אז $-\ln X \sim \exp(1)$.

9.2 התפלגות מעריכית: $X \sim \exp(\lambda),\ \lambda > 0$

ההתפלגות המעריכית מודדת את הזמן שעבר בין שני מאורעות בתהליך פואסון, כלומר תהליך בו מאורעות מתרחשים באופן רציף ובלתי־תלוי בקצב ממוצע קבוע. פונקציית הצפיפות היא:

$$f_X(x) = \lambda\,\mathrm{e}^{-\lambda x}, \qquad x \geq 0$$

פונקציית ההתפלגות המצטברת ופונקציית האחוזון הן:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - \mathrm{e}^{-\lambda x} & x \geq 0 \end{cases}, \qquad F_X^{-1}(p) = -\frac{1}{\lambda}\ln(1-p)$$

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = \frac{1}{\lambda}, \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}, \qquad M_X(t) = \frac{\lambda}{\lambda - t}$$

ניתן לעבור להתפלגות הסטנדרטית כך: $\lambda X \sim \exp(1)$. ההתפלגות המעריכית היא ההתפלגות הרציפה היחידה שהיא חסרת זיכרון:

$$P(X > x + y | X > y) = P(X > x)$$

9.3 התפלגות גמא: $X \sim \Gamma(k, \lambda),\ k, \lambda > 0$

התפלגות גמא מייצגת את זמן ההמתנה עד שמאורע מסוים יתרחש. אם הפרמטר k הוא שלם, אז X הוא סכום של k משתנים מקריים מעריכיים בעלי פרמטר λ. בפרט מתקיים $\Gamma(1, \lambda) = \exp(\lambda)$. הפרמטר λ משמש לקביעת קנה המידה: אם $X \sim \Gamma(k, \lambda)$ אז $\alpha X \sim \Gamma(k, \lambda/\alpha)$.
פונקציית הצפיפות היא:

$$f_X(x) = \frac{1}{\Gamma(k)} x^{k-1} \lambda^k\,\mathrm{e}^{-\lambda x}, \qquad x \geq 0$$

כאשר Γ היא פונקציית גמא:

$$\Gamma(k) = (k-1)! = \int_0^\infty t^{k-1}\,\mathrm{e}^{-t}\,\mathrm{d}t, \qquad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

פונקציית ההתפלגות המצטברת (עבור $k \in \mathbb{N}$) היא:

$$F_X(x) = 1 - \sum_{n=0}^{k-1} \frac{(\lambda x)^n}{n!}\,\mathrm{e}^{-\lambda x}$$

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = \frac{k}{\lambda}, \qquad \mathrm{Var}(X) = \frac{k}{\lambda^2}, \qquad M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^k$$

עבור סכום של משתנים מקריים בלתי־תלויים $X_i \sim \Gamma(k_i, \lambda)$ מתקיים:

$$\sum X_i \sim \Gamma\left(\sum k_i, \lambda\right)$$

בנוסף מתקיים:

$$X \sim \mathcal{N}(\mu, \sigma^2) \Longrightarrow \frac{(X-\mu)^2}{2\sigma^2} \sim \Gamma\left(\frac{1}{2}, 1\right)$$

$$X \sim \mathcal{N}(0,1) \Longrightarrow X^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$$

9.4 התפלגות נורמלית: $X \sim \mathcal{N}(\mu, \sigma^2),\ \mu \in \mathbb{R},\ \sigma^2 > 0$

ההתפלגות הנורמלית מתארת בקירוב משתנים מקריים שנוטים להתקבץ סביב ערך יחיד μ. פונקציות הצפיפות, ההתפלגות המצטברת והאחוזון שלה הן:

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

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right)$$

$$F_X^{-1}(p) = \mu + \sigma\sqrt{2}\,\mathrm{erf}^{-1}(2p-1)$$

התוחלת, השונות והפונקציה יוצרת המומנטים הן:

$$\mathrm{E}(X) = \mu, \qquad \mathrm{Var}(X) = \sigma^2, \qquad M_X(t) = \mathrm{e}^{\mu t + \sigma^2 t^2/2}$$

ניתן לעבור להתפלגות הסטנדרטית כך: $(X-\mu)/\sigma \sim \mathcal{N}(0,1)$. פונקציות הצפיפות, ההתפלגות המצטברת והאחוזון שלה הן:

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\,\mathrm{d}t = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right)$$

$$\Phi^{-1}(p) = \sqrt{2}\,\mathrm{erf}^{-1}(2p-1)$$

בנוסף, אם $X \sim \mathcal{N}(\mu, \sigma^2)$ אז:

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

ואם $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ ו־$X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ הם בלתי־תלויים אז:

$$aX_1 + bX_2 \sim \mathcal{N}\left(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right)$$

להפך, אם X_1, X_2 הם בלתי־תלויים וסכומם מתפלג נורמלית, אז בהכרח גם X_1, X_2 מתפלגים נורמלית.