# התפלגות נורמלית ומשפט הגבול המרכזי

יותם חרובי

סמסטר א' תשע"ח

1. ממבחן של ד"ר שלומי רובינשטיין ינואר 2007
יהי $X \sim N(0.5, 4)$ משתנה מקרי נורמלי. חשבו את $P(X^2 < 1)$
**פתרון:**

$$P(X^2 < 1) = P(-1 < X < 1) = P\left(\frac{-1-0.5}{\sqrt{4}} < \frac{X-0.5}{\sqrt{4}} < \frac{1-0.5}{\sqrt{4}}\right) = P\left(-\frac{3}{4} < \frac{X-0.5}{\sqrt{4}} < \frac{1}{4}\right)$$

המעבר האחרון הוא תקנון המשתנה הנורמלי. אנו מבצעים תקנון משום שטבלת ההתפלגות הנורמלית היא עבור משתנה מתוקנן בלבד. כעת, אם $Z$הוא מ"מ נורמלי סטנדרטי, כלומר $Z \sim N(0,1)$, ניתן לכתוב:

$$P(X^2 < 1) = P\left(-\frac{3}{4} < Z < \frac{1}{4}\right) = \Phi\left(\frac{1}{4}\right) - \Phi\left(-\frac{3}{4}\right)$$

כאשר $\Phi(t) = F_Z(t) = P(Z \le t)$היא פונקצית ההתפלגות המצטברת של משתנה נורמלי סטנדרטי (וערכיה הם אלה שמופיעים בטבלת ההתפלגות הנורמלית).
בטבלה מופיעות הסתברויות $\Phi(t)$ רק עבור $t \ge 0$. אם אנו נדרשים לדעת את ערכה של $\Phi(t)$ עבור $t < 0$ עלינו להשתמש בתכונת הסימטריה של $\Phi(t)$ סביב $t = 0$. כלומר:

$$\Phi\left(-\frac{3}{4}\right) = 1 - \Phi\left(\frac{3}{4}\right)$$

נציב זאת ונקבל:

$$P(X^2 < 1) = \Phi\left(\frac{1}{4}\right) - \left(1 - \Phi\left(\frac{3}{4}\right)\right) = 0.5987 - (1 - 0.7734) = 0.3721$$

2. יהיו $\{T_i\}_{i=1}^n$ ־ סדרה של מ"מ ב"ת המייצגים זמני ההמתנה לאוטובוס בתחנה מסויימת ביום ה $i$. ידוע ש $T_i \sim Exp(\lambda = 1)\ \forall i$

(א) חשבו את ההסתברות שביום הראשון זמן ההמתנה לאוטובוס יעלה על 2 דקות.

(ב) חשבו את ההסתברות שזמן ההמתנה הממוצע ב 64 הימים הראשונים יעלה על שתי דקות. באיזה משפט השתמשתם? אילו לא היה ניתן להשתמש במשפט, מה הייתם צריכים לעשות כדי לענות על שאלה זו

**פתרון:**

(א) חישוב ישיר:

$$P\left(T_1>2\right)=1-P\left(T_1\leq 2\right)=1-F_{T_1}(2)=1-\left(1-e^{-1\cdot 2}\right)=0.135$$

(ב) אנו מתעניינים ב $P\left(\bar{T}_{64}>2\right)$ אבל לא ידועה לנו **ההתפלגות המדוייקת** של המשתנה המקרי $\bar{T}_{64}$. נעזר במשפט הגבול המרכזי המבטיח לנו התכנסות בהתפלגות של סידרת הממוצעים של תצפיות ב"ת. במילים אחרות, משפט הגבול המרכזי נותן לנו את **ההתפלגות המקורבת** של $\bar{T}_{64}$. לפי המשפט, בגלל ש $\{T_i\}_1^{64}$ סדרה של מ"מ ב"ת ושווי התפלגות עם תוחלת $\mathbb{E}T_i=1$ ושונות $Var(T_i)=1$ אז:

$$\bar{T}_{64}\xrightarrow{D}N(1,\,\frac{1}{64})$$

לכן,

$$P\left(\bar{T}_{64}>2\right)=P\left(\frac{\bar{T}_{64}-1}{\sqrt{1/64}}>\frac{2-1}{\sqrt{1/64}}\right)\approx P\left(Z>\frac{1}{1/8}\right)=1-P\left(Z\leq\frac{1}{1/8}\right)=1-\Phi(8)\approx 1-1=0$$

ההסתברות שלאחר 64 ימים זמן ההמתנה הממוצע יהיה גדול משתי דקות, כלומר יחרוג מתוחלת זמן ההמתנה (השווה לדקה) ביותר מדקה, היא קרובה מאוד לאפס. זאת אף על פי שביום בודד יש למאורע כזה הסתברות של בערך $13.5\%$ (סעיף א'). אנו מקבלים כאן המחשה לכך שסידרת הממוצעים הולכת ומתקרבת לתוחלת האמיתית (יש הסתברות הולכת וקטנה שהיא תחרוג ממנה ביותר מדקה).

3. ממבחן של ד"ר שלומי רובינשטיין, מרץ 2007

שיכור מבצע הילוך מקרי על הציר הממשי. בזמן $t=0$ הוא נמצא במיקום $X=0$. בכל יחידת זמן שלמה הוא צועד יחידת מרחק אחת שלמה ימינה או שמאלה בסיכוי שווה ובאופן בלתי תלוי בצעדיו האחרים. חשבו בקירוב את ההסתברות שלאחר 900 צעדים ($t=900$) הוא ימצא ימינה לנקודה $X=44$.

**פתרון:**

נגדיר משתנים מקריים $\{X_i\}_{i=1}^{900}$באופן הבא: $X_i=1$ אם בזמן $t=i$ השיכור צועד ימינה, אחרת $X_i=-1$.

יהי $S$ מיקומו של השיכור לאחר 900 צעדים. אז מתקיים: $S=\sum_{i=1}^{900}X_i$.

אנו מעוניינים בהסתברות הבאה: $P(S>44)$. נחשב אותה בקירוב בעזרת משפט הגבול המרכזי ($CLT$):

הסדרה $\{X_i\}_{i=1}^{900}$ היא סדרת מ"מ ב"ת ושווי התפלגות בעלי תוחלת ושונות סופיים לכן ניתן להשתמש ב $CLT$.

נחשב את התוחלת והשונות של $S$:

$$\mathbb{E}X_i=0.5\cdot 1+0.5\cdot(-1)=0\ \ \forall i$$

$$\mathbb{E}S=\mathbb{E}\left(\sum_{i=1}^{900}X_i\right)=\sum_{i=1}^{900}\mathbb{E}X_i=900\cdot 0=0$$

$$Var(X_i)=\mathbb{E}X_i^2-\left(\mathbb{E}X_i\right)^2=1^2\cdot 0.5+(-1)^2\cdot 0.5-0^2=1\ \ \forall i$$

$$Var(S)=Var\left(\sum_{i=1}^{900}X_i\right)=\sum_{i=1}^{900}Var(X_i)=900\cdot 1=900$$

כאשר המעבר השני נובע מאי תלות.
כעת, לפי $CLT$, המשתנה המתוקנן $\frac{S-0}{\sqrt{900}}$ מתפלג בקירוב נורמלית סטנדרטית. ולכן:

$$P(S>44)=P\left(\frac{S-0}{\sqrt{900}}>\frac{44-0}{\sqrt{900}}\right)\approx 1-\Phi\left(\frac{44-0}{\sqrt{900}}\right)=1-\Phi(1.467)\approx 0.07$$

**הערה:** ניתן להציג את משפט הגבול המרכזי באופן הבא: תהי $\{X_n\}_{n=1}^{\infty}$ סדרת מ"מ שווי התפלגות ובלתי תלויים (לא בהכרח נורמליים!) בעלי תוחלת $\mu$ סופית ושונות $\sigma^2$ סופית. נגדיר את סדרת הסכומים החלקיים $\{S_n\}_{n=1}^{\infty}$ באופן הבא:

$$S_n=\sum_{i=1}^{n}X_i$$

ואת סדרת הממוצעים $\{\bar{X}_n\}_{n=1}^{\infty}$באופן הבא:

$$\bar{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{1}{n}S_n$$

אז לפי חוק הגבול המרכזי:

$$\frac{S_n-\mathbb{E}S_n}{\sqrt{Var(S_n)}}\xrightarrow{D}N(0,\,1) \tag{1}$$

זה הניסוח שהשתמשנו בו בפתרון הזה. מדובר בניסוח נכון, אבל מקובל יותר לנסח את המשפט באופן הבא: במקום (1) לכתוב:

$$\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}\xrightarrow{D}N(0,\,1) \tag{2}$$

שני הניסוחים שקולים. נראה זאת:

$$\mathbb{E}S_n=\mathbb{E}\sum_{i=1}^{n}X_i=\sum_{i=1}^{n}\mathbb{E}X_i=n\mu$$

$$Var(S_n)=Var(\sum_{i=1}^{n}X_i)=\sum_{i=1}^{n}Var(X_i)=n\sigma^2$$

כאשר המעבר האחרון נובע מאי תלות. מכאן ש

$$\frac{S_n-\mathbb{E}S_n}{\sqrt{Var(S_n)}}=\frac{n\cdot\bar{X}_n-n\mu}{\sqrt{n\sigma^2}}=\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}=\frac{\bar{X}_n-\mathbb{E}\bar{X}_n}{\sqrt{Var(\bar{X}_n)}}$$

כלומר, המ"מ המתקבל מתקנון הממוצע הוא בדיוק המ"מ המתקבל מתקנון הסכום ושניהם מתפלגים נורמלית סטנדרטית לפי משפט הגבול המרכזי.

שימו לב שהסימולציה שהצגתי בכתה, וגם שאלה 2 בקובץ זה, מדגישות ניסוח אפשרי שלישי (שקול כמובן). הסימולציה מראה את התפלגות הממוצע **הלא מתוקנן.** ככל ש $n$ גדל, התפלגות הממוצע "הולכת ומתקרבת" ל $N(\mu,\sigma^2/n)$ כלומר:

$$\bar{X}_n\xrightarrow{D}N(\mu,\,\frac{\sigma^2}{n})$$