Matematika A3 építőmérnököknek 12. gyakorlat
Centrális határeloszlás tétel

Legyenek $X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók, melyekre $\mathbb{E}(X_i) = m$, $\mathbb{D}(X_i) = s$ minden $i = 1, 2, \dots$ esetén! Legyen $S_n = X_1 + X_2 + \dots + X_n$. Ekkor minden $a \in \mathbb{R}$ konstansra

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\cdot m}{s\cdot\sqrt{n}} \le a\right) = \Phi(a).$$

Megjegyzés: A tétel mondanivalója, hogy S_n standardizáltjának, azaz $\frac{S_n - n\cdot m}{s\cdot\sqrt{n}}$-nek az eloszlása nagy n esetén jól közelíthető $\mathcal{N}(0,1)$ standard normális eloszlással. Azaz S_n eloszlása nagy n esetén közelítőleg $\mathcal{N}(n\cdot m, s^2\cdot n)$. Így S_n-re vonatkozó valószínűségi kérdéseket meg tudunk oldani a normális eloszlásnál tanultak segítségével.

A centrális határeloszlás tételnek egy nevezetes speciális esete, amikor X_i olyan változó, ami p valószínűséggel 1, $1-p$ valószínűséggel 0 (ezt hívják p paraméterű Bernoulli eloszlásnak). Ekkor $S_n \sim \mathrm{BIN}(n,p)$, így $m = p$, $s = \sqrt{p(1-p)}$. Ezt nevezik **de Moivre–Laplace-tételnek**.

1. feladat Mennyi a valószínűsége, hogy 6000 kockadobás során előforduló hatosok száma 970 és 1050 közé fog esni?

Megoldásvázlat Legyen S_{6000} a hatosok száma 6000 dobás után! Ekkor $S_{6000} \sim \mathrm{BIN}\left(6000, \frac{1}{6}\right)$ és a kérdezett valószínűség:

$$\mathbb{P}(970 < S_{6000} < 1050) = \mathbb{P}\left(\frac{970 - 1000\cdot\frac{1}{6}}{\sqrt{\frac{1}{6}\cdot\frac{5}{6}}\cdot\sqrt{6000}} < \frac{S_{6000} - 1000\cdot\frac{1}{6}}{\sqrt{\frac{1}{6}\cdot\frac{5}{6}}\cdot\sqrt{6000}} < \frac{1050 - 1000\cdot\frac{1}{6}}{\sqrt{\frac{1}{6}\cdot\frac{5}{6}}\cdot\sqrt{6000}}\right) =$$

$$\mathbb{P}\left(-1,04 < \frac{S_{6000} - 1000\cdot\frac{1}{6}}{\sqrt{\frac{1}{6}\cdot\frac{5}{6}}\cdot\sqrt{6000}} < 1,73\right).$$

Itt a de Moivre–Laplace-tétel alapján az egyenlőtlenségek közepén szereplő $\frac{S_{6000}-1000\cdot\frac{1}{6}}{\sqrt{\frac{1}{6}\cdot\frac{5}{6}}\cdot\sqrt{6000}}$ változó jól közelíthető egy $\mathcal{N}(0,1)$ eloszlású változóval. Így a fenti valószínűséget is tudjuk jó közelítéssel számolni:

$$\mathbb{P}(970 < S_{6000} < 1050) \approx \Phi(1,73) - \Phi(-1,04) = 0,9582 - (1 - 0,8508) = 0,809.$$

Megjegyzés:

- Mivel S_{6000} eloszlása binomiális, így a pontos valószínűséget is fel tudnánk írni a súlyfüggvénye segítségével. De ekkor a kérdezett valószínűségre egy körülbelül 80 tagú összegünk lenne, ami nehezen kiértékelhető.
- Mivel egy egészértékű változóról szólt a feladat, az elején úgy is felírhattuk volna a kérdezett valószínűséget, hogy $\mathbb{P}(971 \le S_{6000} \le 1049)$, viszont ekkor egy kicsit más eredményt kaptunk volna. Mindkettő felírásból kapott eredmény jól közelíti a kérdezett valószínűséget és mindkettő jó megoldás. Ezeknél pontosabb eredményt kapunk, ha 970, 5 és 1049, 5 határokkal számolunk. Ezt hívják folytonossági korrekciónak.

2. feladat Egy gyár adott típusú termékei egymástól függetlenül elfogadható minőségűek 0,95 valószínűséggel. Becsüljük meg annak valószínűségét, hogy a következő 150 termékből legfeljebb 10 nem lesz elfogadható!

Megoldásvázlat Legyen S_{150} a 150 termékből a nem elfogadhatók száma! Ekkor tudjuk, hogy $S_{150} \sim \mathrm{BIN}(150; 0,05)$. Így a de Moivre–Laplace-tétel alapján

$$\mathbb{P}(S_{150} \le 10) = \mathbb{P}\left(\frac{S_{150} - 150\cdot 0,05}{\sqrt{150\cdot 0,95\cdot 0,05}} \le \frac{10 - 150\cdot 0,05}{\sqrt{150\cdot 0,95\cdot 0,05}}\right) \approx \Phi(0,94) = 0,8264.$$

3. feladat Kétféle érménk van. Az egyik igazságos, ami 50%-os eséllyel mutat fejet is és írást is, a másik viszont cinkelt, és 55%-os eséllyel mutat fejet. Megtaláljuk az egyik érménket, de nem tudjuk, hogy igazságos-e vagy sem. Ennek eldöntésére az alábbi tesztet hajtjuk végre: feldobjuk az érmét 1000-szer, majd ha legalább 525 fejet mutat, akkor azt mondjuk, hogy ez a cinkelt érme. Ellenkező esetben igazságosnak nyilvánítjuk az érmét.

(a) Mennyi a valószínűsége, hogy a tesztünk téved, feltéve, hogy igazságos volt az érme?

(b) Mennyi a valószínűség, ha hamis?

Megoldásvázlat Legyen I és H a dobott fejek száma az igazságos, illetve a hamis érme esetén! Ekkor $I \sim \mathrm{BIN}(1000; 0,5)$, $H \sim \mathrm{BIN}(1000; 0,55)$. Így a de Moivre–Laplace-tétel alapján annak a valószínűsége, hogy a teszt téved:

(a) $\mathbb{P}(I \geq 525) \approx 1 - \Phi\left(\frac{25}{\sqrt{250}}\right) \approx 0,0571$,

(b) $\mathbb{P}(H < 525) \approx \Phi\left(\frac{-25}{\sqrt{1000 \cdot 0,55 \cdot 0,45}}\right) \approx 0,0559$.

4. feladat Határozzuk meg azt a legkisebb k egész számot, amelyre igaz, hogy annak a valószínűsége, hogy 400 érmedobás során a fejek száma 195 és k közé esik, az legalább 0,5.

Megoldásvázlat Legyen S_{400} a fejek száma 400 dobás után! Ekkor $S_{400} \sim \mathrm{BIN}(400; 0,5)$. Olyan k értéket keresünk, amire

$$\mathbb{P}(195 < S_{400} < k) > 0,5.$$

A bal oldalon lévő valószínűséget tudjuk közelíteni a de Moivre–Laplace-tétel segítségével:

$$\mathbb{P}(195 < S_{400} < k) = \mathbb{P}\left(\frac{195-200}{10} < \frac{S_{400}-200}{10} < \frac{k-200}{10}\right) \approx$$
$$\Phi\left(\frac{k-200}{10}\right) - \Phi\left(-\frac{1}{2}\right) = \Phi\left(\frac{k-200}{10}\right) - 0,3085.$$

Ebből $\Phi\left(\frac{k-200}{10}\right) > 0,8085$, így $\frac{k-200}{10} > 0,88$, azaz $k \geq 209$.

5. feladat Hányszor kell érmével dobnunk ahhoz, hogy 0,95-nél nagyobb valószínűséggel a fej eredmények száma a dobások számának 47%-a és 53%-a közé essen?

Megoldásvázlat Legyen S_n a fejek száma n dobás után! Ekkor $S_n \sim \mathrm{BIN}(n; 0,5)$. Olyan n-et keresünk, amire

$$\mathbb{P}(0,47n < S_n < 0,53n) > 0,95.$$

A bal oldalon lévő valószínűséget tudjuk közelíteni a de Moire–Laplace-tétel segítségével:

$$\mathbb{P}(0,47n < S_n < 0,53n) = \mathbb{P}\left(\frac{-0,03n}{0,5\sqrt{n}} < \frac{S_n - 0,5n}{0,5\sqrt{n}} < \frac{0,03n}{0,5\sqrt{n}}\right) \approx 2\Phi\left(0,06\sqrt{n}\right) - 1.$$

Így olyan n kell, amire $\Phi\left(0,06\sqrt{n}\right) > 0,975$, azaz $0,06\sqrt{n} > 1,96$, így $n > 1067$.

6. feladat Mennyi a valószínűsége annak, hogy 50 db független azonos eloszlású valószínűségi változó összege a [0, 30] intervallumba esik, ha egy ilyen valószínűségi változó eloszlása a [0, 1] intervallumon

(a) egyenletes,

(b) $f(x) = 2x$ sűrűségfüggvény szerint alakul?

Megoldásvázlat Legyen X az adott részfeladatban szereplő egy darab változó és S_{50} az 50 db összege! Mivel X nemnegatív, így S_{50} is nemnegatív, így annak a valószínűségét kell kiszámolni, hogy $S_{50} \leq 30$.

(a) Mivel X eloszlása egyenletes, így $\mathbb{E}(X) = 0,5$, $\mathbb{D}^2(X) = \frac{1}{12}$. Így centrális határeloszlás tétellel:

$$\mathbb{P}(S_{50} \leq 30) \approx \Phi\left(\frac{30-25}{\sqrt{50}\cdot\sqrt{\frac{1}{12}}}\right) = \Phi(2,45) = 0,9929.$$

(b) X várható értéke és szórása definíció alapján számolható (l. 10. gyakorlat 3. feladat): $\mathbb{E}(X) = \frac{2}{3}$, $\mathbb{D}(X) = \sqrt{\frac{1}{18}}$. Így centrális határeloszlás tétellel:

$$\mathbb{P}(S_{50} \leq 30) \approx \Phi\left(\frac{30-50\cdot\frac{2}{3}}{\sqrt{50}\cdot\sqrt{\frac{1}{18}}}\right) = \Phi(-2) = 0,0228.$$

7. feladat Becsüljük meg annak valószínűségét, hogy 10 000 kockadobás összege 34 800 és 35 200 közé esik!

Megoldásvázlat Egy kockadobás várható értéke: $\frac{1}{6}\cdot(1+2+3+4+5+6) = 3,5$, szórása $\sqrt{35/12}$ (definíció alapján számolható). Legyen S_n az első n kockadobás összege! Ekkor standardizálás után tudjuk használni a centrális határeloszlás tételt:

$$\begin{aligned}&\mathbb{P}(34800 < S_{10000} < 35200) = \\ &\quad \mathbb{P}\left(\frac{34800-10000\cdot3,5}{\sqrt{35/12}\cdot\sqrt{10000}} < \frac{S_{10000}-10000\cdot3,5}{\sqrt{35/12}\cdot\sqrt{10000}} < \frac{35200-10000\cdot3,5}{\sqrt{35/12}\cdot\sqrt{10000}}\right) \overset{CHT}{\approx} \\ &\quad 2\cdot\Phi(1,17) - 1 = 0,758.\end{aligned}$$

8. feladat Egy kockát folyamatosan feldobunk addig, amíg a dobások összege meg nem haladja a 300-at. Becsüljük meg annak a valószínűségét, hogy legalább 80 dobásra van ehhez szükségünk!

Megoldásvázlat Legyen S_n a dobások összege n dobás után! Az előző feladatban láttuk, hogy egy kockadobás várható értéke: $3,5$, szórása $\sqrt{35/12}$. Pontosan akkor van szükség legalább 80 dobásra, ha 79 dobás után az összeg még kevesebb, mint 300. Ennek a valószínűségét tudjuk becsülni a centrális határeloszlás tétellel:

$$\mathbb{P}(S_{79} < 300) \approx \Phi\left(\frac{300-79\cdot3,5}{\sqrt{79}\cdot\sqrt{35/12}}\right) = 0,9394.$$

9. feladat Adott 100 égő, melyek élettartama egymástól független exponenciális eloszlású valószínűségi változó, 5 óra várható értékkel.

(a) Tegyük fel, hogy az égőket egymás után használjuk, azonnal kicserélve azokat, amint kiégnek! Becsüljük meg annak valószínűségét, hogy 525 óra után még van működő égőnk!

(b) Tegyük fel, hogy az égők kicserélésének ideje egyenletes eloszlást követ 0 és 0,5 óra között, továbbá az egyszerűbb számítás érdekében tegyük fel, hogy az utolsó kiégett égőt is "kicseréljük"! Mennyi a valószínűsége, hogy így még 550 óra múlva is lesz működő égőnk?

Megoldásvázlat

(a) Legyen X egy égő élettartama és S_n n égő összélettartama! Ekkor $X \sim \mathrm{EXP}\left(\frac{1}{5}\right)$, így

$$\mathbb{P}(S_{100} > 525) \approx 1 - \Phi(0.5) \approx 0.3085.$$

(b) Legyen Y egy égő élettartama a kicseréléssel együtt! Ekkor ez egy független exponenciális és egyenletes eloszlású változó összege, így a várható értéke és szórásnégyzete is ezek várható értékeinek, illetve szórásnégyzeteinek az összege:

$$\mathbb{E}(Y) = 5 + 0,25 = 5,25, \quad \mathbb{D}^2(Y) = 25 + \frac{1}{48} \approx 25,02.$$

Legyen T_{100} a 100 égő összélettartama a kicserélési időkkel együtt! Ekkor a centrális határeloszlás tétel alapján:

$$\mathbb{P}(T_{100} > 550) \approx 1 - \Phi\left(\frac{550 - 525}{10 \cdot \sqrt{25,02}}\right) = 0,3085.$$

10. feladat Egy raktárban 101 láda van, melyek súlyai egymástól függetlenül 20 és 40 kg közti egyenletes eloszlás szerint alakulnak. Egy 3t teherbírású teherautó érkezik, melyre elkezdik felpakolni a ládákat. Becsüljük meg annak a valószínűségét, hogy mind a 101 láda felrakható a teherautóra!

Megoldásvázlat Legyen M egy láda tömege és S_n n láda össztömege! Ekkor $M \sim \mathrm{UNI}(20, 40)$, így $\mathbb{E}(M) = 30$, $\mathbb{D}(M) = \frac{20}{\sqrt{12}}$. Ebből

$$\mathbb{P}(S_{101} < 3000) \approx \Phi\left(\frac{-3\sqrt{3}}{\sqrt{101}}\right) \approx 0.3015.$$

11. feladat Egy bizonyos népesség testsúlyának átlaga 70 kg, szórása 10 kg. Egy hajó teherbírása 7250 kg. Mennyi a valószínűsége jó közelítéssel, hogy 100 emberrel biztonságosan el tud indulni a hajó?

Megoldásvázlat Legyen S_{100} a 100 ember össztömege! A centrális határeloszlás tétellel:

$$\mathbb{P}(S_{100} < 7250) \approx \Phi\left(\frac{7250 - 7000}{100}\right) = 0,9938.$$

12. feladat Egy téglagyárban a téglák 6,3%-a selejtes. Ha a napi termelés mennyisége 2.000 db, mennyi a közelítő valószínűsége annak, hogy köztük 135-nél kevesebb lesz a selejt?

Megoldásvázlat Jelölje S_n az n termék közül a selejtesek számát! Ekkor

$$\mathbb{P}(S_{2000} < 135) \approx \Phi\left(\frac{135 - 2000 \cdot 0.063}{\sqrt{2000 \cdot 0.063 \cdot 0.937}}\right) \approx 0.7967.$$