Matematika A3 építőmérnököknek 12. gyakorlat
Nagy Számok Gyenge Törvénye, normális eloszlás, centrális határeloszlás-tétel

Nagy Számok Gyenge Törvénye Legyenek $X_1, X_2, X_3, \dots$ független, azonos eloszlású valószínűségi változók, melyek várható értéke $\mathbb{E}(X_i) = m$, szórásnégyzete $\mathbb{D}^2(X_i) = s^2$. Jelölje S_n az első n valószínűségi változó összegét, azaz $S_n = X_1 + X_2 + \dots + X_n$. Ekkor tetszőlegesen kicsi $\varepsilon > 0$ esetén

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| > \varepsilon\right) \xrightarrow{n\to\infty} 0.$$

1. feladat Legyen $g : [0,1] \to [0,1]$ egy integrálható függvény és válaszuk véletlenül (egyenletesen) $P_1, P_2, \dots$ pontokat a $[0,1] \times [0,1]$ négyzeten. Legyen X_i az a valószínűségi változó, ami 1, ha P_i pont a $g(x)$ függvény görbéje alatt van és 0, ha felette. Hova tart $\frac{\sum_{i=1}^{n} X_i}{n}$?

Megoldásvázlat A nagy számok gyenge törvénye alapján az érték $\mathbb{E}(X_i)$-hez tart. $X_i \sim \mathrm{BER}(p)$, ahol p annak a valószínűsége, hogy egy pont a görbe alá esik, ami épp $\int_0^1 g(x)\,\mathrm{d}x$. Így a határérték $\int_0^1 g(x)\,\mathrm{d}x$.

2. feladat Legyen $g : [0,1] \to \mathbb{R}$ egy korlátos, integrálható függvény és $X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók $\mathrm{UNI}[0,1]$ eloszlással. Legyen $Y_i = g(X_i)$ minden $i = 1, 2, \dots$ esetén. Hova konvergál Y_i valószínűségi változók átlaga?

Megoldásvázlat Nagy számok gyenge törvénye alapján az átlag az Y_i változók várható értékéhez tart. Mivel X_i sűrűségfüggvénye: $f_X(x) = 1$, ha $x \in [0,1]$ és $f_X(x) = 0$ egyébként, így

$$\mathbb{E}(Y_i) = \mathbb{E}(g(X_i)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\,\mathrm{d}x = \int_0^1 g(x)\,\mathrm{d}x.$$

Tehát az átlag most is a függvény integráljához tart.

Normális eloszlás X valószínűségi változó normális eloszlású m, s^2 paraméterekkel (jel: $X \sim N(m, s^2)$), ha sűrűségfüggvénye:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot s} e^{-\frac{(x-m)^2}{2s^2}}.$$

Ekkor várható értéke $\mathbb{E}(X) = m$, szórásnégyzete $\mathbb{D}^2(X) = s^2$.

Sztenderd normális eloszlásról akkor beszélünk, ha $m = 0$, $s = 1$. Ekkor a sűrűségfüggvény: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, eloszlásfüggvénye pedig

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

Erre nincs elemi képlet, viszont értékei táblázatból kiolvashatók pozitív x-ekre. Negatív x esetén a $\Phi(-x) = 1 - \Phi(x)$ azonosságot használhatjuk.

Ha $X \sim N(m, s^2)$, akkor az ő sztenderdizáltja $Z := \frac{X-m}{s} \sim N(0,1)$.

3. feladat Egy újszülött testsúlya normális eloszlású, 3500 g várható értékkel és 500 g szórással. Mi a valószínűsége, hogy a következő újszülött súlya meghaladja a 3200 g-ot? Mi a válasz az előző kérdésre, ha tudjuk, hogy az újszülött nem igen kis súlyú (egy újszülött igen kis súlyú, ha a súlya kisebb, mint 1500 g)?

Megoldásvázlat Legyen a következő újszülött testsúlya X és Z a sztenderdizáltja. Ekkor $X \sim N(3500, 500^2)$, így

$$\mathbb{P}(X > 3200) = \mathbb{P}\left(Z > -\frac{3}{5}\right) = \Phi\left(\frac{3}{5}\right) \approx 0.7257,$$

$$\mathbb{P}(X > 3200 \,|\, X > 1500) = \frac{\mathbb{P}(X > 3200)}{\mathbb{P}(X > 1500)} = \frac{\Phi\left(\frac{3}{5}\right)}{\Phi(4)} \approx 0.7257.$$

4. feladat Egy új adózási törvényt úgy szerkesztenek meg, hogy az kedvezzen a "középosztálybeli" családoknak, vagyis azoknak, amelyeknek a jövedelme \$20000 and \$30000 között van. Ha a családok jövedelme normális eloszlást követ \$25000 várható értékkel és $\$10000^2$ varianciával, akkor a családok hány százalékának kedvez a törvény?

Megoldásvázlat Legyen X egy család jövedelme \$1000-ban és Z a sztenderdizáltja. Ekkor $X \sim N(25, 10^2)$, így

$$\mathbb{P}(20 < X < 30) = \mathbb{P}(-0.5 < Z < 0.5) = 2\Phi\left(\frac{1}{2}\right) - 1 \approx 0.383.$$

5. feladat Az A3-at tanuló diákok végső százalékpontos eredményei normális eloszlásúnak tekinthetők 65 várható értékkel és 15 szórással.

(a) Mi a valószínűsége, hogy egy véletlenszerű diák közepest (3-ast) kap (azaz, a pontja 60 és 70 pont közé esik)?

(b) Az átment diákok (azok, akik legalább 50 pontot értek el) hány százaléka kap közepest?

Megoldásvázlat Legyen X egy véletlenszerűen választott diák százalékpontos eredménye. Ekkor $X \sim N(65, 15^2)$.

(a) $\mathbb{P}(60 < X < 70) = 2\Phi\left(\frac{2}{3}\right) - 1 \approx 0.2586$

(b) $\mathbb{P}(60 < X < 70 \,|\, X > 50) = \frac{2\Phi\left(\frac{2}{3}\right)-1}{\Phi(1)} \approx 0.31$

6. feladat Dohányosok nikotinszintje normális eloszlású valószínűségi változóval modellezhető, melynek várható értéke 315, szórásnégyzete 17 161.

- Mi a valószínűsége, hogy egy dohányos nikotinszintje meghaladja a 450-et?
- Mi a valószínűsége, hogy egy dohányos nikotinszintje 150 és 400 közé esik?
- Mi az a nikotinszint, amely fölöttivel a dohányosok 5%-a bír?
- Mi a valószínűsége, hogy egy dohányos nikotinszintje 400-nál magasabb, feltéve, hogy az az átlagnál magasabb?

Megoldásvázlat Legyen X egy véletlenszerűen választott dohányos nikotinszintje. Ekkor a feladat alapján $X \sim N(315, 131^2)$. Így

- $\mathbb{P}(X > 450) = 1 - \Phi\left(\frac{135}{131}\right) \approx 0.1515$
- $\mathbb{P}(150 < X < 400) = \Phi\left(\frac{85}{131}\right) + \Phi\left(\frac{165}{131}\right) - 1 \approx 0.6384$
- Kell a, amire $\mathbb{P}(X > a) = 0.05$, azaz $\Phi\left(\frac{a-315}{131}\right) = 0.95$, amiből $a \approx 531.2$.
- $\mathbb{P}(X > 400 \,|\, X > 315) = \frac{1-\Phi\left(\frac{85}{131}\right)}{1/2} \approx 0.5156$

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

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

Ennek speciális esete, amikor $X_i \sim \mathrm{BER}(p)$. Ekkor $S_n \sim \mathrm{BIN}(n, p)$ és

$$\lim_n \mathbb{P}\left(\frac{S_n - n \cdot p}{\sqrt{np(1-p)}} \le a\right) = \Phi(a).$$

Ezt nevezik **de Moivre–Laplace-tétel**nek.

Folytonossági korrekció: Fenti tételekben, ha S_n egy egészértékű változó, akkor pontosabb eredményt kapunk, ha

$$\mathbb{P}(S_n = i) \quad \text{helyett} \quad \mathbb{P}\left(i - \frac{1}{2} < S_n < i + \frac{1}{2}\right)\text{-et}$$

használunk. Ezt nevezzük folytonossági korrekciónak.

7. feladat Számos, gyermekek számára készített társasjátékban annyit kell előrelépnünk a bábuval, mint ahányast a kockával dobtunk. Legyen n dobás után a pozíciónk S_n. Mi a közelítő valószínűsége, hogy 30 dobás után a pozíciónk 90 és 120 közé esik (a határokat is beleértve)? Egy kockadobás szórása $\sqrt{35/12}$.

Megoldásvázlat

$$\mathbb{P}(90 \le S_{30} \le 120) \overset{\text{folytonossági korrekció}}{=} \mathbb{P}(89.5 < S_{30} < 120.5) \overset{CHT}{\approx} 2\Phi\left(\frac{31}{\sqrt{350}}\right) - 1 \approx 0.8904$$

8. feladat Egy félvezető chipeket gyártó cég termékeinek 6.3%-a selejt. Ha a napi termelés mennyisége 2000 darab, mennyi a közelítő valószínűsége annak, hogy köztük 135-nél kevesebb selejt lesz?

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

$$\mathbb{P}(S_{2000} < 135) = \mathbb{P}(S_{2000} < 134.5) \approx \Phi\left(\frac{134.5 - 2000 \cdot 0.063}{\sqrt{2000 \cdot 0.063 \cdot 0.937}}\right) \approx 0.7823.$$

9. feladat Egy kollégiumban az üres férőhelyek száma 150. Egy ide jelentkező hallgató 0.6 valószínűséggel felel meg a felvételi követelményeknek. Ha 240-en jelenkeznek, akkor mi a közelítő valószínűsége, hogy a felvettek közül senkit sem kell átirányítani másik kollégiumba?

Megoldásvázlat Legyen X a felvettek száma. Ekkor

$$\mathbb{P}(X \le 150) = \mathbb{P}(X < 150.5) \approx \Phi\left(\frac{150.5 - 240 \cdot 0.6}{\sqrt{240 \cdot 0.6 \cdot 0.4}}\right) \approx 0.8051.$$

10. feladat Egy 10000-es kisvárosban polgármestert választanak; két jelölt van, A és B. A szavazók nagyon bizonytalanok; egy adott szavazó A-ra 51% valószínűséggel szavaz majd (és így B-re 49%-os valószínűséggel). Mi a közelítő valószínűsége, hogy A nyeri a választást?

Megoldásvázlat Legyen A az A-ra leadott szavazatok száma. Ekkor $A \sim \mathrm{BIN}(10000, 0.51)$, így annak a valószínűsége, hogy A nyer:

$$\mathbb{P}(A > 5000) = \mathbb{P}(A > 5000.5) \approx \Phi\left(\frac{0.995}{\sqrt{0.51 \cdot 0.49}}\right) \approx 0.9767.$$

11. feladat Egy raktárban 101 láda van, melyek súlyai egymástól független és – kg-ban véve – a $(20, 40)$ számközben egyenletes eloszlású valószínűségi változók. Egy 3t teherbírású teherautó érkezik, melyre elkezdik felpakolni a ládákat. Becsülje meg annak 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

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

12. feladat Adott 100 égőnk, melyek élettartama egymástól független exponenciális eloszlású, 5 óra várható értékkel. Tegyük fel, hogy az égőket egymás után használjuk, azonnal kicserélve azt, amelyik kiégett. Becsüljük meg annak valószínűségét, hogy 525 óra után még van működő égőnk!

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

$$\mathbb{P}(S_{100} > 525) \approx 1 - \Phi(0.5) \approx 0.6915.$$