

Matematika A4 - Valószínűesszámitás 6. gyakorlat

Normális eloszlás, CHT

Normális eloszlás

Egy X valószínűségi változó normális eloszlású m, s^2 paraméterekkel (jel: $X \sim \mathcal{N}(m, s^2)$), 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 $\mathbb{D}^2(X) = s^2$.

Standard normális eloszlásról akkor beszélünk, ha $m = 0, s = 1$. Ekkor a sűrűség- és eloszlásfüggvénye:

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

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

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

1. feladat Legyen $X \sim \mathcal{N}(0, 1)$. Számítsuk ki az alábbi értékeket!

- (a) $\mathbb{P}(-1 < X < 1)$ (b) $\mathbb{P}(-2 < X < 2)$ (c) $\mathbb{P}(-1 < X < 2.5)$

Megoldásvázlat Azt használjuk, hogy folytonos valószínűségi változóra

$$\mathbb{P}(a < X < b) = F(b) - F(a),$$

ahol $F(x)$ az X eloszlásfüggvénye. Most $X \sim \mathcal{N}(0, 1)$, így az eloszlásfüggvénye a fenti $\Phi(x)$ függvény, aminek értékeit pozitív x esetén a táblázatból tudjuk kiolvasni, negatív x esetén pedig a $\Phi(-x) = 1 - \Phi(x)$ azonosságot használjuk.

- (a) $\mathbb{P}(-1 < X < 1) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 2 \cdot 0.8413 - 1 = 0.6826$
(b) $\mathbb{P}(-2 < X < 2) = \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 \approx 2 \cdot 0.9772 - 1 = 0.9544$
(c) $\mathbb{P}(-1 < X < 2.5) = \Phi(2.5) - \Phi(-1) = \Phi(2.5) + \Phi(1) - 1 \approx 0.9938 + 0.8413 - 1 = 0.8351$

2. feladat Legyen $X \sim \mathcal{N}(10, 36)$. Számítsuk ki az alábbi értékeket!

- (a) $\mathbb{P}(X > 5)$ (b) $\mathbb{P}(4 < X < 16)$

Megoldásvázlat Az előzőhöz hasonlóan számolható, de most először standardizálnunk kell. Legyen $Z := \frac{X-10}{6}$ az X standardizáltja. Ekkor $Z \sim \mathcal{N}(0, 1)$, így az eloszlásfüggvénye $\Phi(x)$.

- (a) $\mathbb{P}(X > 5) = \mathbb{P}\left(\frac{X-10}{6} > \frac{5-10}{6}\right) = \mathbb{P}(Z > -\frac{5}{6}) = 1 - \Phi\left(-\frac{5}{6}\right) = \Phi\left(\frac{5}{6}\right) \approx 0.7967$
(b) $\mathbb{P}(4 < X < 16) = \mathbb{P}(-1 < Z < 1) \approx 0.6826$

3. feladat Legyen $X \sim \mathcal{N}(5, \sigma^2)$. Mennyi lesz $\mathbb{D}(X)$, ha tudjuk, hogy $\mathbb{P}(X > 9) = 0.2$?

Megoldásvázlat Legyen Z az X standardizáltja! Ekkor

$$\mathbb{P}(X > 9) = \mathbb{P}\left(Z > \frac{4}{\sigma}\right) = 1 - \Phi\left(\frac{4}{\sigma}\right).$$

Tehát $\Phi\left(\frac{4}{\sigma}\right) = 0.8$, vagyis a táblázatból (amit most „visszafelé” használunk, azaz a táblázat belsejében keressük meg a 0.8-at és a táblázat széléről leolvassuk a hozzátartozó értéket): $\frac{4}{\sigma} \approx 0.84$. Vagyis $\mathbb{D}(X) = \sigma \approx \frac{4}{0.84} = 4.76$.

4. feladat Egy új adózási törvényt úgy szerkesztenek meg, hogy az kedvezzen a „középosztálybeli” családoknak, vagyis azoknak, amelyeknek a jövedelme 20 000 \$ and 30 000 \$ között van. Ha a családok jövedelme normális eloszlást követ 25 000 \$ várható értékkel és 10 000 \$ szórással, akkor a családok hány százalékának kedvez a törvény?

Megoldásvázlat Legyen X egy család jövedelme 1000 \$-ban és Z a standardizáltja! Ekkor tudjuk, hogy $X \sim \mathcal{N}(25, 10^2)$, 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 A4-et tanuló diákok végső százalékpontos eredményei normális eloszlásúnak tekinthetők 65 várható értékkel és 15 szórással.

- (a) Mi a valószínűsége, hogy egy véletlenszerű diák középest (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 középest?

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

- (a) $\mathbb{P}(60 < X < 70) = 2\Phi\left(\frac{1}{3}\right) - 1 \approx 0.2586$
- (b) $\mathbb{P}(60 < X < 70 \mid X > 50) = \frac{2\Phi(\frac{1}{3}) - 1}{\Phi(1)} \approx 0.31$

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

- (a) Mi a valószínűsége, hogy egy dohányos nikotinszintje meghaladja a 450-et?
- (b) Mi a valószínűsége, hogy egy dohányos nikotinszintje 150 és 400 közé esik?
- (c) Mi az a nikotinszint, amely fölöttivel a dohányosok 5%-a bír?
- (d) 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 $X \sim \mathcal{N}(315, 131^2)$. Így

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

7. feladat Egy részvény éves hozama normális eloszlást követ 10% várható értékkel és 20% szórással.

- (a) Mennyi a valószínűsége, hogy a részvény ára egy év múlva kisebb lesz, mint most?
- (b) Mennyi a valószínűsége, hogy az éves hozam 20% felett lesz?
- (c) Mennyi a valószínűsége, hogy a következő 10 év közül pontosan négyszer lesz az éves hozam 20% felett?

Megoldásvázlat Legyen X a részvény éves hozama és Z a standardizáltja! Ekkor $X \sim \mathcal{N}(10, 20^2)$.

- (a) $\mathbb{P}(X < 0) = \mathbb{P}\left(Z < -\frac{1}{2}\right) = 1 - \Phi\left(\frac{1}{2}\right) \approx 0.3085$
- (b) $\mathbb{P}(X > 20) = \mathbb{P}\left(Z > \frac{1}{2}\right) = 1 - \Phi\left(\frac{1}{2}\right) \approx 0.3085$

Megjegyzés: Vegyük észre, hogy a fenti két kérdés a várható értékre szimmetrikus. Így a normális eloszlás szimmetriájából is következik, hogy ezek a valószínűségek megegyeznek.

- (c) Legyen Y a következő 10 év közül azon évek száma, amelyekben a hozam 20% felett van! Ekkor $Y \sim \text{BIN}(10; 0.3085)$. Így a kért valószínűség:

$$\binom{10}{4} \cdot 0.3085^4 \cdot 0.6915^6.$$

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

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

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.

Ennek speciális esete, amikor $X_i \sim \text{BER}(p)$. Ekkor $S_n \sim \text{BIN}(n, p)$, így $m = p$, $s = \sqrt{p(1-p)}$. Ezt nevezik **de Moivre–Laplace-tételnek**.

Ha S_n egy egészértékű változó, akkor pontosabb eredményt kapunk, ha **folytonossági korrekciót** használunk, azaz $\mathbb{P}(S_n = i)$ helyett $\mathbb{P}(i - \frac{1}{2} < S_n < i + \frac{1}{2})$ -vel számolunk.

8. feladat Számos, gyermekek számára készített társasjátékban annyit kell előrelépni 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 Egy kockadobás várható értéke: $\frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = 3.5$, így

$$\begin{aligned} \mathbb{P}(90 \leq S_{30} \leq 120) &\stackrel{\text{folytonossági}}{\underset{\text{korrekció}}{=}} \mathbb{P}(89.5 < S_{30} < 120.5) = \\ &= \mathbb{P} \left(\frac{89.5 - 30 \cdot 3.5}{\sqrt{30} \cdot \sqrt{35/12}} < \frac{S_{30} - 30 \cdot 3.5}{\sqrt{30} \cdot \sqrt{35/12}} < \frac{120.5 - 30 \cdot 3.5}{\sqrt{30} \cdot \sqrt{35/12}} \right) \stackrel{\text{CHT}}{\approx} \\ &\approx 2\Phi \left(\frac{31}{\sqrt{350}} \right) - 1 \approx 0.8904 \end{aligned}$$

9. feladat Egy félvezető chipket 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

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

10. feladat Egy raktárban 101 láda van, melyek súlyai egymástól független és – kg-ban mérve – a $(20, 40)$ intervallumban egyenletes eloszlású valószínűségi változók. Egy 3 tonna teherbírású teherautó érkezik, melyre elkezdik felpakolni a ládákat. Becsüljük meg annak valószínűségét, hogy mind a 101 láda felrakható a teherautóra!

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

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

11. 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 kieggett.

- (a) Becsüljük meg annak valószínűségét, hogy 525 óra után még van működő égőnk!
- (b) Becsüljük meg, hogy még hány égőre van szükségünk ahhoz, hogy 90%-os valószínűséggel 525 óra után is legyen működő égőnk!

Megoldásvázlat

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

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

(b) Legyen n a szükséges égők száma! Ekkor az kell, hogy

$$\mathbb{P}(S_n > 525) = 0.9,$$

vagyis standardizálva

$$\mathbb{P}\left(\frac{S_n - n \cdot 5}{\sqrt{n} \cdot 5} > \frac{525 - 5n}{5\sqrt{n}}\right) = 0.9.$$

Komplementert véve:

$$\Phi\left(\frac{525 - 5n}{5\sqrt{n}}\right) = 0.1.$$

Hogy a jobb oldalon lévő értéket megtaláljuk a táblázatban, a $\Phi(-x) = 1 - \Phi(x)$ formulát kell használni, amiből azt kapjuk, hogy

$$-\frac{525 - 5n}{5\sqrt{n}} = \Phi^{-1}(0.9) = 1.28.$$

A $\sqrt{n}$ -es taggal beszorozva, négyzetre emelve, majd egy oldalra rendezve egy másodfokú egyenletet kapunk n -re: $25n^2 - 5290.96n + 275625 = 0$, aminek gyökei $n_1 = 118.96$ és $n_2 = 92.68$, de ebből n_2 a négyzetre emelésnél bejövő hamis gyök, vagyis még kb 19 égőt kell szereznünk.

12. feladat Egy gyár kétfajta érmét gyárt: egy igazságosat, és egy hamisat ami 55% eséllyel mutat fejet. Van egy ilyen érménk, de nem tudjuk igazságos-e vagy pedig hamis. Ennek eldöntésére a következő statisztikai tesztet hajtjuk végre: feldobjuk az érmét 1000-szer, ha legalább 525-ször fejet mutat, akkor hamisnak nyilvánítjuk, ha 525-nél kevesebb fej lesz a dobások között, akkor az érmét igazságosnak tekintjük. Mi a valószínűsége, hogy a tesztünk téved abban az esetben, ha az érme igazságos volt? És ha hamis volt?

Megoldásvázlat Legyen I és H a dobott fejek száma az igazságos, illetve a hamis érme esetén! Ekkor $I \sim \text{BIN}(1000, 0.5)$, $H \sim \text{BIN}(1000, 0.55)$. Így annak a valószínűsége, hogy a teszt téved:

$$\mathbb{P}(I \geq 525) = \mathbb{P}(I > 524.5) \approx 1 - \Phi\left(\frac{24.5}{\sqrt{250}}\right) \approx 0.0606,$$

$$\mathbb{P}(H < 525) = \mathbb{P}(H < 524.5) \approx \Phi\left(\frac{-25.5}{\sqrt{1000 \cdot 0.55 \cdot 0.45}}\right) \approx 0.0526.$$

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

Megoldásvázlat Legyen X a felvettek száma! Ekkor

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

14. 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:

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 2x^2 dx = \frac{2}{3}, \\ \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 2x^3 dx = \frac{1}{2}, \\ \mathbb{D}(X) &= \sqrt{\mathbb{E}(X^2) - (\mathbb{E}(X))^2} = \sqrt{\frac{1}{18}}.\end{aligned}$$

Így centrális határeloszlás tétellel:

$$\mathbb{P}(S_{50} \leq 30) \approx \Phi\left(\frac{30 - 50 \cdot \frac{2}{3}}{\sqrt{50} \cdot \sqrt{\frac{1}{18}}}\right) = \Phi(-2) = 0.0228.$$

15. 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! 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 (az egyszerűség kedvéért most folytonossági korrekció nélkül):

$$\mathbb{P}(S_{79} < 300) \approx \Phi\left(\frac{300 - 79 \cdot 3.5}{\sqrt{79} \cdot \sqrt{35/12}}\right) = 0.9394.$$