Felsőbb Matematika Villamosmérnököknek – Sztochasztika
házi feladat megoldások, 2017 ősz

Minden héten összesen 2 pontot érnek a kitűzött feladatok.

1.HF: (Beadási határidő: 2017.09.27.)

HF 1.1 Egy országban minden tízezredik ember HIV fertőzött. A fertőzöttség vizsgálatára hivatott AIDS teszt megbízhatósága 99%, abban az értelemben, hogy a fertőzötteken elvégzett tesztek 1%-a ad (hibásan) negatív eredményt, illetve az egészségeseken elvégzett tesztek 1%-a ad (hibásan) pozitív eredményt.

Mórickát véletlenszerűen kisorsolták a teljes népességből. Elvégezték rajta a tesztet, és pozitív lett. Mi a valószínűsége, hogy Móricka valóban fertőzött?

Megoldás: Jelölje A_1 azt az eseményt, hogy Móricka fertőzött, A_2 azt, hogy egészséges, B pedig azt, hogy a tesztje pozitív. A szövegbeli adatok szerint $P(A_1) = \frac{1}{10000}$, $P(A_2) = \frac{9999}{10000}$, $\mathbb{P}(B|A_1) = \frac{99}{100}$ és $\mathbb{P}(B|A_2) = \frac{1}{100}$. Mivel $\{A_1, A_2\}$ teljes eseményrendszer, Bayes tétele szerint

$$\begin{aligned}\mathbb{P}(A_1|B) &= \frac{\mathbb{P}(A_1)\mathbb{P}(B|A_1)}{\mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2)} = \\ &= \frac{\frac{1}{10000}\frac{99}{100}}{\frac{1}{10000}\frac{99}{100} + \frac{9999}{10000}\frac{1}{100}} = \frac{99}{99+9999} \approx 0.01 = 1\%.\end{aligned}$$

Ez azért ilyen kevés, mert kevés a fertőzött és sok az egészséges ember, így arányaiban nagyon sok lesz a hibás pozitív teszt: 99 helyes pozitív tesztre 9999 hibás pozitív teszt jut (bár az összes tesztnek csak egy százaléka hibás).

HF 1.2 Pistike minden nyári este tesz egy sétát, és közben az eget nézi, hullócsillagokat figyelve. Egy este átlagosan 4-et szokott látni. Ennek megfelelően, ha 4-et vagy többet lát, akkor vidáman megy haza, ha viszont kevesebbet, akkor bánatosan.

Pistike augusztus 16-án bánatosan ment haza. Ezt tudva, mennyi a valószínűsége, hogy egyetlen hullócsillagot sem látott?

(Rávezető kérdés: Legyen X a Pisike által augusztus 16-án látott hullócsillagok száma - ami persze egy valószínűségi változó. Mi X eloszlása? Pontosabban: Milyen eloszlással jó modellezni X-et?)

Megoldás: X-et Poisson eloszlással modellezzük, mert nagyon sok meteor próbálkozik azzal, hogy pont Pistike szeme láttára égjen el aznap este, és mindegyik kis valószínűséggel jár sikerrel, egymástól függetlenül. X pedig a sikeres próbákozások száma. A szöveg szerint $\mathbb{E}X = 4$, így $X \sim Poi(\lambda)$ ahol $\lambda = 4$. Vagyis minden $k = 0, 1, 2, \ldots$-re

$$p_k := \mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} = e^{-4}\frac{4^k}{k!}.$$

Így a kérdésre a válasz a feltételes valószínűség definíciója szerint

$$\begin{aligned}\mathbb{P}(X = 0|X < 4) &= \frac{\mathbb{P}(X = 0 \text{ és } X < 4)}{\mathbb{P}(X < 4)} = \frac{\mathbb{P}(X = 0)}{\mathbb{P}(X < 4)} = \frac{p_0}{p_0 + p_1 + p_2 + p_3} = \\ &= \frac{e^{-4}}{e^{-4}\left(1 + 4 + \frac{4^2}{2} + \frac{4^3}{6}\right)} = \frac{3}{71} \approx 0.042 = 4.2\%.\end{aligned}$$

2.HF: (Beadási határidő: 2017.10.11.) *(Hibás határidő javítva, bocs.)*

HF 2.1 Legyen X és Y két független valószínűségi változó, X binomiális $n_1 = 10$ és $p = \frac{2}{3}$ paraméterrel, Y pedig binomiális $n_2 = 5$ és $p = \frac{2}{3}$ paraméterrel.

a.) Mi X generátorfüggvénye?

b.) Mi Y generátorfüggvénye?

c.) Mi $Z := X + Y$ generátorfüggvénye?

d.) Mi Z eloszlása? (Értsd: adjuk meg Z lehetséges értékeit és ezek valószínűségeit, *vagy* mondjuk meg az eloszlás nevét.)

Megoldás: Az (n, p) paraméterű binomiális eloszlás generátorfüggvénye $g_{Bin(n,p)}(z) = (q + pz)^n$, ahol $q = 1 - p$. Így

a.) $g_X(z) = g_{Bin(10,\frac{2}{3})}(z) = \left(\frac{1}{3} + \frac{2}{3}z\right)^{10}$.

b.) $g_Y(z) = g_{Bin(5,\frac{2}{3})}(z) = \left(\frac{1}{3} + \frac{2}{3}z\right)^{5}$.

c.) Mivel X és Y független, $g_Z(z) = G_X(z)g_Y(z) = \left(\frac{1}{3} + \frac{2}{3}z\right)^{10}\left(\frac{1}{3} + \frac{2}{3}z\right)^{5} = \left(\frac{1}{3} + \frac{2}{3}z\right)^{15}$.

d.) g_Z éppen az $n = 15$, $p = \frac{2}{3}$ paraméterű binomiális eloszlás generátorfüggvénye. Mivel a generátorfüggvény egyértelműen meghatározza az eloszlást, $Z \sim Bin(15, \frac{2}{3})$.

HF 2.2 Egy véletlen rekurzív algoritmus futása során kezdetben elindul egyetlen folyamat, melynek sorsa kétféle lehet:

i.) Eredményesen lefut és kilép anélkül, hogy újabb folyamatot indítana. Ennek valószínűsége legyen p.

ii.) Mielőtt lefutna, véletlen számú újabb folyamatot indít, amik ugyanolyanok, mint ő maga. Ezek száma lehet 1, 2 vagy 3, azonos (vagyis $\frac{1-p}{3}$) valószínűséggel. Az eredeti folyamat csak akkor lép ki, amikor ezek mind lefutottak.

Az algoritmus futása során elinduló egyes folyamatok mind ugyanolyan eloszlással hoznak létre újabb folyamatokat, függetlenül egymástól és az előzményektől.

Kezdetben egyetlen folyamat indul el – legyen ez egymaga a „nulladik generáció”. Az „első generáció” álljon a legelső folyamat által indított folyamatokból (ha van ilyen). A „második generáció” álljon az első generáció tagjai által indított folyamatokból. És így tovább, az $n + 1$-edik generáció álljon az n-edik generáció tagjai által indított folyamatokból, $n = 0, 1, 2, \ldots$-re. Jelölje Z_n az n-edik generáció elemeinek számát. Legyen $X = Z_1$, és legyen $N = \sum_{n=0}^{\infty} Z_n$ az algoritmus futása során elinduló folyamatok össz-száma.

Válaszoljuk meg az alábbi kérdéseket

I.) ha $p = 0.6$

II.) ha $p = 0.4$.

a.) Mi X eloszlása?

b.) $\mathbb{E}X =?$

c.) Mi X generátorfüggvénye?

d.) $\mathbb{E}Z_{40} =?$

e.) Mi Z_2 generátorfüggvénye?

f.) $\mathbb{E}N =?$

g.) $\mathbb{P}(Z_4 = 0) =?$

h.) Mennyi a valószínűsége, hogy az algoritmus előbb-utóbb lefut, vagyis hogy valamelyik generáció már üres? (Avagy: annak a valószínűsége, hogy a legelső folyamat előbb-utóbb kilép.)

Megoldás:

I.) a.)

k	0	1	2	3
$p_k = \mathbb{P}(X = k)$	$\frac{6}{10}$	$\frac{4}{30}$	$\frac{4}{30}$	$\frac{4}{30}$

b.) $m := \mathbb{E}X = 0 \cdot \frac{6}{10} + 1 \cdot \frac{4}{30} + 2 \cdot \frac{4}{30} + 3 \cdot \frac{4}{30} = \frac{4}{5}$.

c.) X generátorfüggvénye $g(z) = \sum_{k=0}^{\infty} p_k z^k = \frac{6}{10}z^0 + \frac{4}{30}z^1 + \frac{4}{30}z^2 + \frac{4}{30}z^3 = \frac{18+4z+4z^2+4z^3}{30}$.

d.) $\mathbb{E}Z_{40} = m_{40} = m^{40} = \left(\frac{4}{5}\right)^{40} \approx 0.00013$.

e.) Z_2 generátorfüggvénye

$$g_2(z) = g(g(z)) = \frac{18 + 4\left(\frac{18+4z+4z^2+4z^3}{30}\right) + 4\left(\frac{18+4z+4z^2+4z^3}{30}\right)^2 + 4\left(\frac{18+4z+4z^2+4z^3}{30}\right)^3}{30}.$$

f.) Mivel $m < 1$ (a folyamat szubkritikus), $\mathbb{E}N = \frac{1}{1-m} = \frac{1}{1-\frac{4}{5}} = 5$.

g.) Az $r_n := \mathbb{P}(Z_n = 0)$ sorozatról tudjuk, hogy eleget tesz az $r_{n+1} = g(r_n)$ rekurziós szabálynak $r_0 = 0$ kezdeti értékkel. Így az r_n-ek egyesével számolhatók:

* $r_0 = 0$
* $r_1 = g(r_0) = g(0) = \frac{6}{10}$
* $r_2 = g(r_1) = g(\frac{6}{10}) = 0.7568$
* $r_3 = g(r_2) \approx 0.8351$
* $r_4 = g(r_3) \approx 0.8820$.

h.) $m < 1$, vagyis a folyamat szubkritikus (és nem elfajult, vagyis $\mathbb{P}(X = 1) \neq 1$), ezért a kihalás valószínűsége $r_\infty = 1$.

II.) a.)

k	0	1	2	3
$p_k = \mathbb{P}(X = k)$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$

b.) $m := \mathbb{E}X = 0 \cdot \frac{4}{10} + 1 \cdot \frac{2}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{2}{10} = \frac{6}{5}$.

c.) X generátorfüggvénye $g(z) = \sum_{k=0}^{\infty} p_k z^k = \frac{4}{10}z^0 + \frac{2}{10}z^1 + \frac{2}{10}z^2 + \frac{2}{10}z^3 = \frac{4+2z+2z^2+2z^3}{10}$.

d.) $\mathbb{E}Z_{40} = m_{40} = m^{40} = \left(\frac{6}{5}\right)^{40} \approx 1470$.

e.) Z_2 generátorfüggvénye

$$g_2(z) = g(g(z)) = \frac{4 + 2\left(\frac{4+2z+2z^2+2z^3}{10}\right) + 2\left(\frac{4+2z+2z^2+2z^3}{10}\right)^2 + 2\left(\frac{4+2z+2z^2+2z^3}{10}\right)^3}{10}.$$

f.) Mivel $m \geq 1$, $\mathbb{E}N = \infty$.

g.) Az $r_n := \mathbb{P}(Z_n = 0)$ sorozatról tudjuk, hogy eleget tesz az $r_{n+1} = g(r_n)$ rekurziós szabálynak $r_0 = 0$ kezdeti értékkel. Így az r_n-ek egyesével számolhatók:

* $r_0 = 0$
* $r_1 = g(r_0) = g(0) = \frac{4}{10}$
* $r_2 = g(r_1) = g(\frac{4}{10}) = 0.5248$
* $r_3 = g(r_2) \approx 0.5890$
* $r_4 = g(r_3) \approx 0.6280$.

h.) $m > 1$, vagyis a folyamat szuperkritikus, ezért számolni kell: a kihalás valószínűsége a $z = g(z)$ fixpont-egyenlet egyetlen $[0, 1)$-beli megoldása. Vagyis meg kell oldani a

$$z = \frac{4 + 2z + 2z^2 + 2z^3}{10}$$

egyenletet, ami nullára redukálva

$$z^3 + z^2 - 4z + 2 = 0.$$

Ez első ránézésre ijesztő harmadfokú, de szerencsére tudjuk, hogy $g(1) = 1$ mindig, így $z = 1$ mindig megoldás. Vagyis a baloldalból $z - 1$ kiemelhető. És valóban:

$$z^3 + z^2 - 4z + 2 = (z-1)(z^2 + 2z - 2),$$

tehát az egyenletünk

$$(z-1)(z^2 + 2z - 2) = 0.$$

Ebből $z - 1 = 0$ vagy $z^2 + 2z - 2 = 0$. Ez utóbbi már csak másodfokú, megoldásai $z = -1 \pm \sqrt{3}$.
Összefoglalva: a $z = g(z)$ fixpontegyenlet megoldásai $-1-\sqrt{3}$, $-1+\sqrt{3}$ és 1. Ezek közül pontosan egy esik a $[0, 1)$ intervallumba (ahogy annak lenni kell), és ez a kihalási valószínűség:

$$r_\infty = \mathbb{P}(\text{kihalás}) = -1 + \sqrt{3} \approx 0.7321$$

3.HF: (Beadási határidő: 2017.11.15.)

HF 3.1 Egy bolha a számegyenesen ugrál: a nullából indul, és minden másodpercben 1 egységnyit ugrik: $\frac{1}{2}$ valószínűséggel pozitív irányba (jobbra), $\frac{1}{2}$ valószínűséggel negatív irányba (balra), az előzményektől függetlenül. Legyen X_n a bolha helye n ugrás után. Vagyis ha az i-edik ugrás $\xi_i \in \{-1, 1\}$, akkor $X_n = \xi_1 + \xi_2 + \cdots + \xi_n$. (Ez az egyszerű szimmetrikus bolyongás).

a.) Közelítsük a centrális határeloszlás tétel (CHT) segítségével annak a valószínűségét, hogy $n = 2500$ ugrás után a bolha legalább $K = 200$ lépésre jobbra van a kiinduló helyétől.

b.) Legfeljebb mekkora lehet a CHT közelítés hibája a Berry-Esseen tétel szerint?

c.) Adjunk ugyenerre a valószínűségre felső becslést a Hoeffding-egyenlőtlenség segítségével!

d.) Becsüljük meg a keresett valószínűséget a Cramér tétel alkalmazásával is.

(Segítség: a p paraméterű Bernoulli eloszlás Cramér féle rátafüggvénye

$$I(x) = x \ln\left(\frac{x}{1-x}\frac{1-p}{p}\right) + \ln\left(\frac{1-x}{1-p}\right)$$

(ha $0 < x < 1$).)

Megoldás:

a.) Az $m = \mathbb{E}\xi_i = 0$, $\sigma = \sqrt{Var\xi_i} = \sqrt{1} = 1$ jelöléssel a CHT szerint

$$\begin{aligned}\mathbb{P}(X_n \geq K) = 1 - \mathbb{P}(X_n < K) &\approx 1 - \Phi\left(\frac{K - nm}{\sqrt{n}\sigma}\right) = 1 - \Phi\left(\frac{200 - 0}{\sqrt{2500} \cdot 1}\right) = \\ &= 1 - \Phi(4) \approx 0.000032\end{aligned}$$

Itt Φ a standard normális eloszlásfüggvény. A weblapon lévő táblázatban csak $\Phi(3.09) = 0.9990$-ig van benne, tehát onnan csak annyit lehet leolvasni, hogy a CHT szerint $\mathbb{P}(X_n \geq K) \approx 1 - \Phi(4) \leq 0.001$.

b.) A $C = 0.4748$, $\delta := \mathbb{E}\left(|\xi_i - m|^3\right) = \mathbb{E}\left(|\xi_i - 0|^3\right) = \mathbb{E}1^3 = 1$ jelöléssel a Berry-Esseen tétel szerint a CHT becslés hibája legfeljebb

$$\text{hiba} \leq \frac{C\delta}{\sqrt{n}\sigma^3} = \frac{0.4748 \cdot 1}{\sqrt{2500} \cdot 1^3} = \frac{0.4748}{50} \approx 0.0095.$$

Tanulság: így a Berry-Esseen tétel állítása az, hogy

$$|\mathbb{P}(X_n \geq K) - 0.000032| \leq 0.0095,$$

vagy másképpen:

$$-0.009468 = 0.000032 - 0.0095 \le \mathbb{P}(X_n \ge K) \le 0.009532 = 0.000032 + 0.0095.$$

Az alsó becslés semmitmondó, tehát biztosak csak annyiban lehetünk, hogy

$$\mathbb{P}(X_n \ge K) \le 0.009532$$

Ebben az esetben a hibakorlát nagyságrendekkel nagyobb, mint a becsült érték maga – sőt, mint látni fogjuk, a tényleges értéknél is.

c.) A ξ_i-k korlátosak: az $a_i := -1$, $b_i := 1$ jelöléssel $a_i \le \xi_i \le b_i$. A Hoeffding egyenlőtlenség alkalmazásához kelleni fog, hogy $\sum_{i=1}^n (b_i - a_i)^2 = \sum_{i=1}^n (1-(-1))^2 = \sum_{i=1}^n 4 = 4n = 10000$. Mivel $\mathbb{E}X_n = nm = 0$, a $t := K = 200$ jelöléssel A Hoeffding egyenlőtlenség szerint

$$\begin{aligned}
\mathbb{P}(X_n \ge K) &= \mathbb{P}(X_n \ge \mathbb{E}X_n + t) \le \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right) = \\
&= \exp\left(-\frac{2 \cdot 200^2}{10000}\right) = e^{-8} \approx 0.000335
\end{aligned}$$

d.) A ξ_i eloszlásának nincs megadva a rátafüggvénye, ezért úgy járunk a legjobban, ha egy kicsit átjelöljük a feladatot: Legyen $\xi_i = -1 + 2\eta_i$, avagy $\eta_i = \frac{1+\xi_i}{2}$. Így $\eta_i \in \{0, 1\}$ már Bernoulli eloszlású $p = \frac{1}{2}$ paraméterrel, és az $S_n := \eta_1 + \cdots + \eta_n$ jelöléssel $X_n = -n + 2S_n$. Vagyis a keresett valószínűség

$$\begin{aligned}
\mathbb{P}(X_n \ge K) &= \mathbb{P}(-n + 2S_n \ge K) = \mathbb{P}\left(S_n \ge \frac{n+K}{2}\right) = \\
&= \mathbb{P}\left(\frac{S_n}{n} \ge \frac{1}{2} + \frac{K}{2n}\right) = \mathbb{P}\left(\frac{S_n}{n} \ge 0.54\right).
\end{aligned}$$

Pont ilyenek becslésére való a Cramér tétel. $m = \mathbb{E}\eta_i = 0.5$, $a := 0.54$ és $b := \infty$ jelöléssel $a > m$, ezért a tétel állítása

$$\mathbb{P}\left(\frac{S_n}{n} \ge 0.54\right) = \mathbb{P}\left(\frac{S_n}{n} \in [a, b)\right) \lessapprox e^{-nI(a)},$$

ahol I a $p = \frac{1}{2}$ paraméterű Bernoulli eloszlás rátafüggvénye, a segítség szerint

$$I(a) = a \ln\left(\frac{a}{1-a}\right) + \ln\left(\frac{1-a}{\frac{1}{2}}\right) = 0.54 \ln \frac{0.54}{0.46} + \ln 0.92 \approx 0.0032034$$

Így

$$\mathbb{P}(X_n \ge K) = \mathbb{P}\left(\frac{S_n}{n} \ge 0.54\right) \lessapprox e^{-2500 \cdot 0.0032034} = e^{-8.0085} \approx 0.000333$$

(Megjegyzés: mint láttuk, ha $p = \frac{1}{2}$ és $a \approx 0.5$, akkor a Bernoulli eloszlás nagy eltérés becslésére a Hoeffding egyenlőtlenség nagyon jó (közel olyan jó, mint a Cramér tétel). Ha $p \ne \frac{1}{2}$, akkor ez nincs így.)

(Figyelmeztetés: a Cramér tétel alkalmazásánál nagyon óvatosan kell bánni a kerekítéssel. Ha pl. az $I(a) \approx 0.0032034$-et könynelműen $I(a) \approx 0.003$-ra kerekítjük, akkor $nI(a) \approx 8.0085$ helyett $nI(a) \approx 7.5$-et kapunk. Ez pedig negyon nem mindegy, mert $nI(a)$ a kitevőbe kerül: a túl bátor kerekítéssel a helyes $\mathbb{P}(X_n \ge K) \lessapprox 0.000333$ helyett $\mathbb{P}(X_n \ge K) \lessapprox 0.000553$- at kapunk, ami több, mint 60%-os hiba.)

HF 3.2 Bergengóciában minden nap $\frac{1}{3}$ valószínűséggel esik az eső, az előzményektől függetlenül. A bergengóc királyi palota kertészének akkor kell locsolnia, ha három napja nem érte víz a kertet – vagyis az előző három napon se eső nem volt, se locsolás. Hosszú távon a napok hány százalékán nem kapnak vizet a növények? *(Segítség: jelöljük X_n-nel, hogy az n-edik napon éjfélkor éppen hány napja nem érte víz a kertet.)*

Megoldás: Az X_n Markov lánc állapottere $S = \{0, 1, 2, 3\}$. A $k = 0, 1, 2$ állapotból $k + 1$-be urgunk, ha nem esik az eső – vagyis $\frac{2}{3}$ valószínűséggel. Ha viszont esik (aminek a valószínűsége $\frac{1}{3}$), akkor 0-ba. Ezzel szemben a 3 állapotból biztosan 0-ba ugrunk, mert a kertész locsol. Így az átmenetmátrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

A rendszer pontosan akkor van az $1, 2, 3$ állapotok valamelyikében, ha előző nap a növények nem kaptak vizet. Vagyis pont annyi napon nem kapnak vizet, ahányszor az $1, 2, 3$ állapotok valamelyikébe lépünk. A kérdés tehát az $f(X_n)$ időátlaga, ahol az $f : S \to \mathbb{R}$ függvény

$$f(k) = \begin{cases} 1, & \text{ha } k \in \{1, 2, 3\} \\ 0, & \text{ha } k = 0 \end{cases}.$$

Ugyanez (oszlop)vektor-jelöléssel:

$$f = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Látható, hogy a Markov lánc véges allapotterű és irreducibilis (mindenhonnan mindenhova el lehet jutni), így az ergodtétel szerint $f(X_n)$ időátlaga 1 valószínűséggel

$$\overline{f(X_n)} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(X_n) = \mathbb{E}_\pi f = \sum_{k\in S} \pi_k f(k) = \pi f,$$

ahol π az egyetlen stacionárius eloszlás (sorvektor).

A π kereséséhez a $(P^T - I)\pi^T = 0$ homogén lineáris egyenletrendszert kell megoldani, ahol I az egységmátrix. Esetünkben

$$P^T - I = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{2}{3} & -1 & 0 & 0 \\ 0 & \frac{2}{3} & -1 & 0 \\ 0 & 0 & \frac{2}{3} & -1 \end{pmatrix},$$

így az egyenletrendszer a szokásos bővített mátrix jelöléssel

$$\left(\begin{array}{cccc|c} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & -1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & -1 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & -1 & 0 \end{array}\right).$$

Ezt Gauss eliminációval megoldva

$$\left(\begin{array}{cccc|c} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & -1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & -1 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} & 1 & 0 \\ 0 & \frac{2}{3} & -1 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} & 1 & 0 \\ 0 & 0 & -\frac{2}{3} & 1 & 0 \\ 0 & 0 & \frac{2}{3} & -1 & 0 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc|c} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & -\frac{2}{3} & \frac{1}{3} & 1 & 0 \\ 0 & 0 & -\frac{2}{3} & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -1 & \frac{1}{2} & \frac{1}{2} & \frac{3}{2} & 0 \\ 0 & -1 & \frac{1}{2} & \frac{3}{2} & 0 \\ 0 & 0 & -1 & \frac{3}{2} & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -1 & \frac{1}{2} & 0 & \frac{9}{4} & 0 \\ 0 & -1 & 0 & \frac{9}{4} & 0 \\ 0 & 0 & -1 & \frac{3}{2} & 0 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc|c} -1 & 0 & 0 & \frac{27}{8} & 0 \\ 0 & -1 & 0 & \frac{9}{4} & 0 \\ 0 & 0 & -1 & \frac{3}{2} & 0 \end{array}\right).$$

Ezt kiolvasva $\pi_0 = \frac{27}{8}\pi_3$, $\pi_1 = \frac{9}{4}\pi_3$, $\pi_1 = \frac{3}{2}\pi_3$, vagyis pl. $\pi_3 = 8$ választással kapjuk, hogy egy megoldás a $\tilde{\pi} = (27, 18, 12, 8)$ vektor, és persze megoldás ennek minden számszorosa is. Minket az egyetlen normált megoldás érdekel (amire $\sum_{k \in S} \pi_k = 1$), vagyis

$$\pi = \frac{1}{65}\tilde{\pi} = \begin{pmatrix} \frac{27}{65} & \frac{18}{65} & \frac{12}{65} & \frac{8}{65} \end{pmatrix}.$$

Így a kérdésre a válasz

$$\begin{aligned} \overline{f(X_n)} &= \pi f = \begin{pmatrix} \frac{27}{65} & \frac{18}{65} & \frac{12}{65} & \frac{8}{65} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \frac{27 \cdot 0 + 18 \cdot 1 + 12 \cdot 1 + 8 \cdot 1}{65} = \\ &= \frac{38}{65} \approx 0.58 = 58\%. \end{aligned}$$

4.HF: (Beadási határidő: 2017.11.29.)

HF 4.1 Egy lépcsőházban 3 villanykörte van, és folyamatosan égnek – ha csak nincsenek éppen kiégve. Az egyes villanykörték élettartama független és exponenciális eloszlású, 1 év várható értékkel. A lépcsőházban évente átlag kétszer megjelenik a gondnok (Poisson folyamat szerint), és az összes kiégett körtét újra cseréli. Jelöljük $X(t)$-vel a t idő elteltével működő körték számát. Az időt mérjük években.

a.) Adjuk meg az $X(t)$ Markov lánc állapotterét és az átmenetrátákat (rátamátrixot). *(Vigyázat: ha éppen 2 körte működik, milyen rátával ég ki közülük **valamelyik**?)*

b.) Írjuk fel az infinitezimális generátort!

c.) Tegnap délben pont 2 körte működött. Körülbelül mennyi a valószínűsége, hogy holnap délben mind működni fog?

d.) Tegnap délben pont 2 körte működött. Körülbelül mennyi a valószínűsége, hogy pont 20 évvel később mind működni fog?

e.) A lépcsőházban akkor van zavaróan sötét, ha legfeljebb 1 körte világít. Hosszú távon az idő hány százalékában van zavaróan sötét?

Megoldás:

a.) Az állapottér $S = \{0, 1, 2, 3\}$. Ha egy körte ég, annak kiégési rátája a várható élettartam reciproka, esetünkben $1/1 = 1$, vagyis az 1 állapotból a 0 állapotba $\lambda_{10} = 1$ rátával ugrik a rendszer. Ha 2 körte ég, akkor már $1 + 1 = 2$ rátával fog kiégni valamelyik, így $\lambda_{21} = 2$. Ugyanígy $\lambda_{32} = 3$. A gondnok látogatásainak rátája 2, így a 0, 1 és 2 állapotból is $\lambda_{03} = \lambda_{13} = \lambda_{23} = 2$ rátával ugrik a rendszer

3-ba. (Ha a rendszer éppen a 3 állapotban van és jön a gondnok, akkor nem történik semmi.) Más átmenet nem lehetséges, így a ráta-mátrix

$$\underline{\underline{\lambda}} = \begin{pmatrix} * & 0 & 0 & 2 \\ 1 & * & 0 & 2 \\ 0 & 2 & * & 2 \\ 0 & 0 & 3 & * \end{pmatrix}.$$

(Ennek a főátlójában nincs semmi, mert helyben ugrás nincs.)

b.) Az infinitezimális generátort úgy kapjuk, hogy a $\underline{\underline{\lambda}}$ főátlóját kitöltjük úgy, hogy minden sorösszeg nulla legyen:

$$G = \begin{pmatrix} -2 & 0 & 0 & 2 \\ 1 & -3 & 0 & 2 \\ 0 & 2 & -4 & 2 \\ 0 & 0 & 3 & -3 \end{pmatrix}.$$

c.) Legyen tegnap délben a $t = 0$ időpont. Így a kérdés

$$\mathbb{P}(X(\Delta t) = 3 \,|\, X(0) = 2) = P_{23}(\Delta t) = ?$$

ahol $\Delta(t) = \frac{2}{365}$, mert az időt években mérjük, és $P(t)$ a t idejű átmenetmártix. Mivel Δt kicsi (a rendszerbeli tipikus várakozási időkhöz képest), $P(\Delta t) \approx P(0) + \Delta t P'(0) = I + \Delta t G$, ahol I az egységmátrix. Vagyis

$$P_{23}(\Delta t) \approx I_{23} + \frac{2}{365} G_{23} = 0 + \frac{2}{365} \cdot 2 = \frac{4}{365} \approx 0.01 = 1\%.$$

(Más szóval: Két nap alatt a $2 \to 3$ átmenet valószínűsége első közelítésben megegyezik annak valószínűségével, hogy a két nap alatt jön a gondnok, mert már ennek is elég kicsi a valószínűsége – annak meg, hogy több esemény is történik, még sokkal kisebb. A gondnok látogatásának valószínűsége pedig rövid idő alatt ráta $\cdot$ idő.

d.) A kérdés most is

$$\mathbb{P}(X(t) = 3 \,|\, X(0) = 2) = P_{23}(t) = ?$$

de ezúttal $t = 20$ hosszú idő. A Markov lánc véges állapotterű és irreducibilis (valamint folytonos idejű, így periodikus nem lehet), ezért a Markov láncok alaptétele szerint a $\mathbb{P}(X(t) = 3)$ valószínűséget a stacionárius eloszlás szerinti π_3 valószínűséggel közelíthetjük (a kiinduló állapottól függetlenül), ahol π az egyetlen stacionárius eloszlás. Ennek kereséséhez a $G^T \pi^T = 0$ lineáris egyenletrendszert kell megoldani:

$$\left(\begin{array}{cccc|c} -2 & 1 & 0 & 0 & 0 \\ 0 & -3 & 2 & 0 & 0 \\ 0 & 0 & -4 & 3 & 0 \\ 2 & 2 & 2 & -3 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -2 & 1 & 0 & 0 & 0 \\ 0 & -3 & 2 & 0 & 0 \\ 0 & 0 & -4 & 3 & 0 \\ 0 & 3 & 2 & -3 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -2 & 1 & 0 & 0 & 0 \\ 0 & -3 & 2 & 0 & 0 \\ 0 & 0 & -4 & 3 & 0 \\ 0 & 0 & 4 & -3 & 0 \end{array}\right)$$

$$\sim \left(\begin{array}{cccc|c} -2 & 1 & 0 & 0 & 0 \\ 0 & -3 & 2 & 0 & 0 \\ 0 & 0 & -4 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -1 & 1/2 & 0 & 0 & 0 \\ 0 & -1 & 2/3 & 0 & 0 \\ 0 & 0 & -1 & 3/4 & 0 \end{array}\right)$$

Ezt kiolvasva $\pi_0 = \frac{1}{2}\pi_1$, $\pi_1 = \frac{2}{3}\pi_2$, $\pi_2 = \frac{3}{4}\pi_4$, vagyis pl. $\pi_3 = 4$ választással kapjuk, hogy egy megoldás a $\tilde{\pi} = (1, 2, 3, 4)$ vektor, és persze megoldás ennek minden számszorosa is. Minket az egyetlen normált megoldás érdekel (amire $\sum_{k\in S}\pi_k = 1$), vagyis

$$\pi = \frac{1}{10}\tilde{\pi} = \begin{pmatrix}\frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10}\end{pmatrix}.$$

Így a kérdésre a válasz

$$\mathbb{P}(X(20) = 3 \,|\, X(0) = 2) \approx \pi_3 = \frac{3}{10}.$$

e.) Legyen $f : S \to \mathbb{R}$ a „zavaróan sötét van” esemény indikátora, vagyis $f(0) = f(1) = 1$, $f(2) = f(3) = 0$. Vektor-jelöléssel

$$f = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}.$$

Így a $t \in [0, T]$ időtartam alatt a 0 és 1 állapotok valamelyikében eltöltött idő $\int_0^T f(X(t))\mathrm{d}t$, a kérdés pedig, hogy az idő hány százalékát tölti a rendszer a 0 és 1 állapotok valamelyikében, az $f(X(t))$ időátlaga:

$$\overline{f(X)} = \lim_{T\to\infty}\frac{1}{T}\int_0^T f(X(t))\mathrm{d}t = ?$$

A Markov lánc véges állapotterű és irreducibilis, ezért az ergodtétel szerint egy valószínűséggel

$$\overline{f(X)} = \mathbb{E}_\pi f = \sum_{i\in S}\pi_i f(i) = \pi f,$$

ahol π az egyetlen stacionárius eloszlás. Esetünkben

$$\overline{f(X)} = \pi f = \begin{pmatrix}\frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10}\end{pmatrix}\begin{pmatrix}1\\1\\0\\0\end{pmatrix} = \frac{3}{10}.$$

HF 4.2 Egy fagyishoz a gyerekek Poisson folyamat szerint érkeznek, percenként átlagosan ketten, és beállnak a sorba. A fagyis bácsi minden gyereket független, exponenciális eloszlású idő alatt szolgál ki, 1 perc várható értékkel. Ha a sorban nem áll senki, a fagyis bácsi unatkozik. Hosszú távon az idő hány százalékában fog unatkozni, ha

a.) A sor hossza legfeljebb 5 lehet, mert ha már 5-en állnak sorban, akkor a további érkező gyerekeket az apukájuk elrángatja.

b.) A sor hossza akármennyi lehet, de ha legalább 5, akkor csak a gyerekek legelszántabb $\frac{1}{3}$-a áll be. (Minden gyerek, a többitől függetlenül, $\frac{1}{3}$ valószínűséggel legelszántabb.)

(Segítség: Legyen $X(t)$ a sor hossza t perc elteltével. Szabad kihasználni, hogy $X(t)$ születési-halálozási folyamat.)

Megoldás:

a.) Legyen a segítség szerint $X(t)$ a sor hossza t perc elteltével. Ennek állapottere $S = \{0, 1, 2, 3, 4, 5\}$. Ugrani csak szomszédos állapotokba lehet, vagyis $X(t)$ születési-halálozási folyamat. A felfelé ugrás, vagyis a gyerekek érkezésének rátája mindig 2, vagyis $\lambda_{01} = \lambda_{12} = \lambda_{23} = \lambda_{34} = \lambda_{45} = 2$. A lefelé ugrás, vagyis a kiszolgálás rátája mindig 1, vagyis $\lambda_{10} = \lambda_{21} = \lambda_{32} = \lambda_{43} = \lambda_{54} = 1$. A többi $\lambda_{ij} = 0$. A gráf-reprezentáció

$$0 \underset{1}{\overset{2}{\rightleftarrows}} 1 \underset{1}{\overset{2}{\rightleftarrows}} 2 \underset{1}{\overset{2}{\rightleftarrows}} 3 \underset{1}{\overset{2}{\rightleftarrows}} 4 \underset{1}{\overset{2}{\rightleftarrows}} 5$$

A születési-halálozási folyamat stacionárius eloszlása mindig olyan, hogy minden i-re (amire ez értelmes) $\pi_i\lambda_{i,i+1} = \pi_{i+1}\lambda_{i+1,i}$. Esetünkben $i = 0, 1, 2, 3, 4$-re $2\pi_1 = 1\pi_{i+1}$, vagyis az egyetlen stacionárius eloszlás

$$\pi = const \begin{pmatrix} 1 & 2 & 4 & 8 & 16 & 32 \end{pmatrix} = \begin{pmatrix} \frac{1}{63} & \frac{2}{63} & \frac{4}{63} & \frac{8}{63} & \frac{16}{63} & \frac{32}{63} \end{pmatrix}.$$

A fagyis bácsi akkor unatkozik, ha $X(t) = 0$, vagyis az $f : S \to \mathbb{R}$, $f = (1, 0, 0, 0, 0, 0)^T$ (oszlopvektor) időátlagát keressük. Ez az ergodtétel szerint

$$\overline{f(X)} = \pi f = \pi_0 = \frac{1}{63} \approx 0.016 = 1.6\%.$$

b.) Az állapottér ezúttal $S = \{0, 1, 2, 3, 4, 5, \dots\}$ végtelen, de $X(t)$ továbbra is születési-halálozási folyamat. A lefelé ugrás rátája továbbra is mindig 1, vagyis $\lambda_{i+1,i} = 1$ ($i = 0, 1, 2, \dots$). A felfelé ugrás rátája viszont csak akkor 2, ha legfeljebb 4 gyerek áll a sorban, vagyis $\lambda_{01} = \lambda_{12} = \lambda_{23} = \lambda_{34} = \lambda_{45} = 2$, Ha $i \geq 5$, akkor $\lambda_{i,i+1} = \frac{2}{3}$, mert csak minden harmadik érkező gyerek áll be a sorba. A gráf-reprezentáció ezúttal

$$0 \underset{1}{\overset{2}{\rightleftarrows}} 1 \underset{1}{\overset{2}{\rightleftarrows}} 2 \underset{1}{\overset{2}{\rightleftarrows}} 3 \underset{1}{\overset{2}{\rightleftarrows}} 4 \underset{1}{\overset{2}{\rightleftarrows}} 5 \underset{1}{\overset{2/3}{\rightleftarrows}} 6 \underset{1}{\overset{2/3}{\rightleftarrows}} 7 \underset{1}{\overset{2/3}{\rightleftarrows}} 8 \quad \cdots\cdots\cdots$$

Így a $\pi_i\lambda_{i,i+1} = \pi_{i+1}\lambda_{i+1,i}$ szabály minden $i = 0, 1, 2, \dots$-re érvényes. $i = 0, \dots, 4$-re azt mondja, hogy $\pi_{i+1} = 2\pi_i$, viszont $i \geq 5$-re azt, hogy $\pi_{i+1} = \frac{2}{3}\pi_i$. Vagyis az egyetlen stacionárius eloszlás, ha létezik,

$$\pi = c \begin{pmatrix} 1 & 2 & 4 & 8 & 16 & 32 & \frac{64}{3} & \frac{128}{9} & \dots & \frac{2^i}{3^{i-5}} & \dots \end{pmatrix}.$$

Esetünkben létezik, mert a vektorban szereplő számok összege véges:

$$\begin{aligned} \frac{1}{c} &:= 1 + 2 + 4 + 8 + 16 + 32 + \frac{64}{3} + \frac{128}{9} + \dots + \frac{2^i}{3^{i-5}} + \dots = \\ &= 31 + 32 \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \dots \right) = 31 + \frac{32}{1 - \frac{2}{3}} = 127 < \infty. \end{aligned}$$

Az ergodtétel állítása ebben az esetben is érvényes az $f = (1, 0, 0, 0, 0, 0, \dots)^T$ függvényre, vagyis a fagyis bácsi az idő

$$\overline{f(X)} = \pi f = \pi_0 = \frac{1}{127} \approx 0.8\%$$

-ában unatkozik.

5.HF: (Beadási határidő: 2017.12.07.)

HF 5.1 a.) Az X valószínűségi változó sűrűségfüggvénye

$$f_\theta(x) = \begin{cases} \theta x^{\theta-1}, & \text{ha } 0 < x < 1 \\ 0, & \text{ha nem} \end{cases},$$

ahol $0 < \theta < \infty$ paraméter, és nem ismerjük. Mintát vettünk X-ből, és azt kaptunk, hogy 0.997; 0.8; 0.853; 0.975; 0.986; 0.927; 0.936; 0.879; 0.767; 0.912. Adjunk maximum likelihood becslést θ-ra!

Megoldás: A minta elemszáma $N = 10$, a likelihood-függvény

$$L(\theta; \underline{x}) = \prod_{i=1}^{N} f_\theta(x_i) = \prod_{i=1}^{N} \theta x_i^{\theta-1} = \theta^N \left(\prod_{i=1}^{N} x_i \right)^{\theta-1},$$

mert minden x_i mintaelemünk $(0,1)$-be esik. *(Különben $L(\theta; \underline{x}) = 0$ lenne, ami azt jelentené, hogy a minta nem lehet a megadott eloszlásból.)* A log-likelihood függvény ennek logaritmusa:

$$l(\theta; \underline{x}) = \ln L(\theta; \underline{x}) = N \ln\theta + (\theta - 1)\ln\left(\prod_{i=1}^{N} x_i \right).$$

Ennek, mint θ függvényének a maximum-helyét keressük, amihez megnézzük, hogy hol nulla a derivált:

$$\frac{\mathrm{d}}{\mathrm{d}\theta} l(\theta; \underline{x}) = \frac{N}{\theta} + \ln\left(\prod_{i=1}^{N} x_i \right) := 0,$$

aminek a megoldása

$$\theta_{ML} = -\frac{N}{\ln\left(\prod_{i=1}^{N} x_i\right)} = -\frac{N}{\sum_{i=1}^{N} \ln x_i}.$$

Esetünkben $\sum_{i=1}^{N} \ln x_i \approx -1.052856$, amiből

$$\theta_{ML} \approx 9.498$$

Ellenőrizhető hogy ez tényleg globális maximumhely.

b.) **bónusz:** A táblázatkezelőben egy n-szer tizes táblázat minden elemébe `rand()` -ot írtam, minek hatására a táblázatkezelő kitöltötte a táblázatot független, $[0,1]$-en egyenletes eloszlású (pszeudo-)véletlen számokkal. Ezek után kikerestettem mind a 10 oszlop maximumát, és átmásoltam (kerekítve) a matekpéldába. Azt kaptam, hogy 0.997; 0.8; 0.853; 0.975; 0.986; 0.927; 0.936; 0.879; 0.767; 0.912. Vajon mennyi lehetett az n?

Megoldás: A táblázat első oszlopába n darab független, $[0,1]$-en egyenletes eloszlású valószínűségi változó került – legyen a nevük $\xi_1, \ldots, \xi_n$. Az első mintaelem, vagyis x_1 ezeknek a maximuma: $x_1 = \max\{\xi_1, \ldots, \xi_n\}$. Először ennek az eloszlását határozzuk meg. Az eloszlásfüggvény könnyű, mert x_1 akkor kisebb, mint valami x, ha minden $\xi_i < x$.

$$F_X(x) = \mathbb{P}(x_1 < x) = \mathbb{P}(\xi_1 < x, \xi_2 < x, \ldots, \xi_n < x) = \left(\mathbb{P}(\xi_1 < x)\right)^n,$$

mert a ξ_i-k függetlenek és azonos eloszlásúak. Mivel egyenletesek is $[0,1]$-en

$$\mathbb{P}(\xi_1 < x) = \begin{cases} 0, & \text{ha } x \le 0 \\ x, & \text{ha } 0 < x < 1 \\ 1, & \text{ha } x > 1. \end{cases}$$

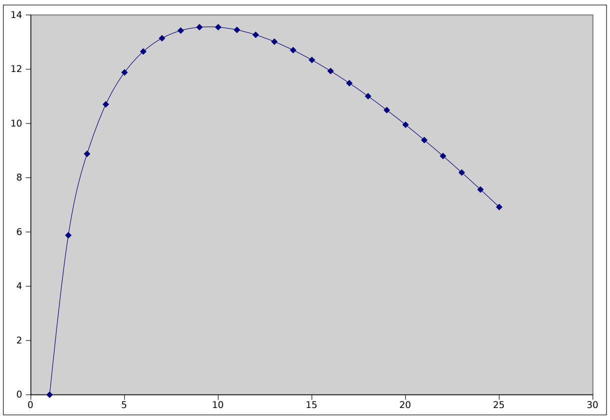
Vagyis

$$F_X(x) = \begin{cases} 0, & \text{ha } x \le 0 \\ x^n, & \text{ha } 0 < x < 1 \\ 1, & \text{ha } x > 1. \end{cases}$$

Az x_1 sűrűségfüggvénye ennek deriváltja:

$$f_X(x) = \begin{cases} nx^{n-1}, & \text{ha } 0 < x < 1 \\ 0, & \text{ha nem.} \end{cases}$$

Az $x_1, x_2, x_3, \ldots, x_{10}$ mintaelemek persze mind ugyanilyen eloszlásúak, és függetlenek. Vagyis pontosan az előző feladat szerinti eloszlásból vettünk $N = 10$ elemű mintát, csak az ismeretlen paraméter most nem $\theta \in (0, \infty)$, hanem $n \in \{1, 2, 3 \ldots\}$. Az n-et maximum likelihood becsléssel próbáljuk megtippelni. Az $n := \theta$ átjelöléssel a likelihood-függvény és a log-likelihood-függvény *ugyanaz* lesz, mint az előző feladatban, csak a maximumát most nem a valós, hanem az egész számokon keressük. Viszont a deriválásból most is látszik, hogy $l(\theta)$ növekvő a $(0; \theta_{ML}) \approx (0; 9.5)$ intervallumon és csökkenő a $(\theta_{ML}; \infty) \approx (9.5; \infty)$ intervallumon. Így az egész számokon vett maximumát csak a $\theta_{ML} \approx 9.5$ valamelyik szomszédján érheti el – vagyis $n = 9$-ben vagy $n = 10$-ben. Lásd az 1. ábrát.



1. ábra. $l(\theta) = 10 \ln \theta - (\theta - 1) \cdot 1.05285618792923$

Hogy a 9 és a 10 közül hol nagyobb az l, azt konkrét számolással dönthetjük el:

$$l(9) = N \ln 9 + (9 - 1) \ln \left(\prod_{i=1}^{N} x_i \right) \approx 13.549$$

$$l(10) = N \ln 10 + (10 - 1) \ln \left(\prod_{i=1}^{N} x_i \right) \approx 13.550$$

Éppen hogy csak, de $l(10) > l(9)$, így a tippünk $n_{ML} = 10$. *(És ez történetesen helyes.)*

HF 5.2 A közönséges szibériai turul vérnyomása (higanymilliméterben) normális eloszlású valószínűségi változó, várható értéke m, ami egyedenként változó, szórása mindig $\sigma = 20$. A madár akkor egészséges, ha $m = 200$. Egy közönséges szibériai turulnak megmértük a vérnyomását 10 különböző napon (ezek már függetlennek tekinthetők), és azt kaptuk, hogy 216; 179; 199; 199; 163; 175; 193; 190; 201; 223. Döntsünk 95%-os szinten arról a hipotézisről, hogy a turul egészséges!

Megoldás: egy $X \sim \mathcal{N}(m, \sigma^2)$ eloszlásból vettünk $n = 10$ elemű mintát, $\sigma = 20$ ismert, de m ismeretlen. A hipotetikus várható érték $\mu = 200$, így a null-hipotézis $m = \mu$. Ennek tesztelésére egymintás kétoldali u-próbát végzünk, a konfidencia-szint $1 - \varepsilon = 95\%$, vagyis $\varepsilon = 0.05$. A minta-átlag $\bar{x} = 193.8$, így a teszt-statisztika

$$u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{193.8 - 200}{20} \sqrt{10} \approx -0.98$$

Az elfogadási küszöb

$$K_\varepsilon = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right) = \Phi^{-1}(0.975) \approx 1.96$$

(itt Φ a standard normális eloszlásfüggvény).
Döntés: $|u| \leq K_\varepsilon$, ezért a null-hipotézist elfogadjuk.