Felsőbb Matematika Villamosmérnököknek - Sztochasztika 2. ZH

2015. december 1. 18:00, **A csoport**
Munkaidő: 50 perc. Minden feladat 15 pontot ér.

1. Egy háromállapotú, diszkrét idejű Markov lánc gráf-reprezentációját mutatja az ábra.

$$\begin{tikzcd} 1 \arrow[r, shift left, "1"] & 2 \arrow[l, shift left, "\frac{1}{2}"] \arrow[r, shift left, "\frac{1}{2}"] & 3 \arrow[l, shift left, "1"] \end{tikzcd}$$

   a.) Írjuk fel a Markov átmenetmátrixot!

   b.) A Markov lánc kezdetben az 1-es állapotból indul. Körülbelül mennyi a valószínűsége, hogy 100 lépés után éppen a 2-es állapotban van?

   Megoldás:

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

   b.) Mivel a Markov lánc periodikus $d = 2$ periódussal, az 1-es állapotból indulva páros sok lépésben csak az 1-es vagy 3-as állapotba juthat, a 2-esbe nem. Vagyis a válasz 0.

2. Egy folytonos idejű Markov lánc a $\{0, 1\}$ állapotokban lehet. A 0 állapotban átlagosan 1 percet tölt, mielőtt 1-be ugrana, az 1 állapotban pedig átlagosan 2 percet, mielőtt 0-ba ugrana. Írjuk fel a Markov lánc infinitezimális generátorát! (Az időt mérjük percben.)

   Megoldás: Mivel a Markov lánc a 0 állapotban átlagosan 1 percet tölt, a 0-ból való elugrás rátája 1. Mivel az 1 állapotban pedig átlagosan 2 percet tölt, az 1-ből való elugrás rátája $\frac{1}{2}$. Vagyis a tartózkodási idő paraméter vektor $(\lambda_0\, , \lambda_1) = (1\, , \frac{1}{2})$ és mivel urgani mindig csak a másik állapotba lehet, az éleken való ugrási ráták ugyanezek: $\lambda_{01} = \lambda_0 = 1$, $\lambda_{10} = \lambda_1 = \frac{1}{2}$. Vagyis az infinitezimális generátor

$$G = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

3. Egy fali ingaóra leolvasása bizonytalan: a leolvasott érték normális eloszlású valószínűségi változó, aminek várható értéke a ténylegesen mutatott idő, a szórása pedig ismeretlen. Egymástól függetlenül 10-szer is megmértük, hogy mennyi idő alatt ér körbe az óra. (Referenciául egy nagyon pontos órát használtunk, de a leolvasás a fentiek szerint bizonytalan.) Az elméleti 12 órához képest a következő késéseket mértük (másodpercben): -84, 18, -8, 1, 51, 36, -55, -26, -16, 33. Döntsünk 95%-os szinten arról a hipotézisről, hogy az óra pontos. *(Segítség: a fenti számok összege -50, négyzetösszege 16388.)*

   Megoldás: Egy (m, σ^2) paraméterű normális eloszlából vettünk $n = 10$ elemű mintát, ahol m és σ is ismeretlen. A nullhipotézis $m = \mu := 0$ egyenlőség, ezért kétoldali egymintás t-próbát végzünk, $\varepsilon = 1 - 95\% = 0.05$. Ehhez kell a korrigált tapasztalati szórás:

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{-50}{10} = -5 \text{ az átlag,} \\
s_n^2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = \frac{16388}{10} - (-5)^2 = 1638.8 - 25 = 1613.8 \text{ a tapasztalati szórásnégyzet,} \\
s_n^{*2} &= \frac{n}{n-1} s_n^2 = \frac{10}{9} \cdot 1613.8 \approx 1793.11 \text{ a korrigált tapasztalati szórásnégyzet,} \\
s_n^* &= \sqrt{s_n^{*2}} \approx \sqrt{1793.11} \approx 42.35 \text{ a korrigált tapasztalati szórás.}
\end{aligned}$$

   Így a teszt-statisztika

$$t = \frac{\bar{x} - \mu}{s_n^*}\sqrt{n} \approx \frac{-5 - 0}{42.35}\sqrt{10} \approx -0.373.$$

Az elfogadási küszöb a t eloszlás táblázatából a $df = n - 1 = 9$ szabadsági fokhoz és $\varepsilon = 0.05$ „két-farok” súlyhoz (avagy $\frac{\varepsilon}{2} = 0.025$ „egy-farok” súlyhoz) tartozó kvantilis, vagyis

$$K := t_{\frac{\varepsilon}{2}} = 2.262.$$

Döntés: $|t| \leq K$, azért a null hipotézist *elfogadjuk*.

Felsőbb Matematika Villamosmérnököknek - Sztochasztika 2. ZH

2015. december 1. 19:00, **B csoport**
Munkaidő: 50 perc. Minden feladat 15 pontot ér.

1. Egy nagy országban a sok szavazó 2 pártra oszlik: 70%-uk a „Mindenkit Utálunk” párt (MU) híve, 30%-uk pedig a „Becsüljetek Minket” párt (BM) támogatója. Egy közvéleménykutató intézet 500 szavazót kérdez meg a pártszimpátiájáról. Adjunk nagy eltérés becslést annak valószínűségére, hogy a kutatás a BM pártot mutatja erősebbnek.

Segítség: A p paraméterű Bernoulli eloszlás Cramér féle rátafüggvénye

$$I(x) = x \ln \frac{(1-p)x}{p(1-x)} - \ln \frac{1-p}{1-x}.$$

A λ paraméterű Poisson eloszlás Cramér féle rátafüggvénye

$$I(x) = x \ln \frac{x}{\lambda} - x + \lambda.$$

Megoldás: Legyen $n = 500$ és $i = 1, 2, \ldots, n$-re az X_i valószínűségi változó

$$X_i = \begin{cases} 1, & \text{ha az } i\text{-edik megkézdezett BM-támogató} \\ 0, & \text{ha nem.} \end{cases}$$

Az X_i-k függetlenek és $p = 0.3$ paraméterű Bernoulli eloszlásúak. Legyen $S_n = X_1 + \cdots + X_n$ a megkérdezett BM-támogatók száma. A feladat a $\mathbb{P}(S_n > 250) = \mathbb{P}\left(\frac{S_n}{n} > \frac{1}{2}\right)$ valószínűség becslése.

1. megoldás: A Cramér tételt alkalmazzuk a $p = 0.3$ paraméterű Bernoulli eloszlásra hogy megbecsüljük a $\mathbb{P}\left(\frac{S_n}{n} \in (a, b)\right)$ valószínűséget, ahol $a = \frac{1}{2}$ és $b = \infty$. A várható érték $m = p = 0.3$, ezért $m < a$, így

$$\mathbb{P}\left(\frac{S_n}{n} > \frac{1}{2}\right) = \mathbb{P}\left(\frac{S_n}{n} \in (a, b)\right) \lessapprox e^{-nI(a)} = e^{-nI\left(\frac{1}{2}\right)}.$$

A segítségben megadott rátafüggvénybe behelyettesítve $x = \frac{1}{2}$-et

$$e^{-nI\left(\frac{1}{2}\right)} = \exp\left[-n\left(\frac{1}{2}\ln\frac{(1-p)\frac{1}{2}}{p(1-\frac{1}{2})} - \ln\frac{1-p}{1-\frac{1}{2}}\right)\right] = \cdots = \sqrt{4pq}^{\,n}, \tag{1}$$

ahol $q = 1 - p$. Konkrétan $n = 500$, $p = 0.3$-ra

$$\mathbb{P}\left(\frac{S_n}{n} > \frac{1}{2}\right) \lessapprox \sqrt{4 \cdot 0.3 \cdot 0.7}^{\,500} \approx 1.17 \cdot 10^{-19}.$$

(Megjegyzés: persze nem szükséges az (1)-beli szép leegyszerűsített képletet kiszámolni vagy előadásról tudni – elég numerikusan behelyettesíteni $n = 500$, $x = 0.5$, $p = 0.3$-at.)

2. megoldás: Mivel az X_i-k korlátosak $0 = a_i \leq X_i \leq b_i = 1$ korlátokkal, használhatjuk a Hoeffding egyenlőtlenséget is. $\sum_{i=1}^{n}(b_i - a_i)^2 = 500 \cdot (1-0)^2 = 500$, valamint $\mathbf{E}S_n = np = 500 \cdot 0.3 = 150$, vagyis $t = 100$ választással

$$\mathbb{P}(S_n > 250) = \mathbb{P}(S_n > \mathbf{E}S_n + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right) = \exp\left(-\frac{2 \cdot 100^2}{500}\right) = e^{-40} \approx 4.25 \cdot 10^{-18}.$$

(Megjegyzés: látható, hogy a Hoeffding egyenlőtlenség durvább felső becslést ad, mint a Cramér tétel.)

2. Pistike szobájában két villanykörte van. Mindig mindkettőt égeti – már ha nincsenek kiégve. A körték egymástól függetlenül exponenciális eloszlású véletlen idő alatt égnek ki, 1 év várható értékkel. Ha csak az egyik van kiégve, Pistike nem törődik vele, de ha a második is kiég, akkor mindkettőt azonnal újra cseréli. Jelölje $X(t)$ a Pistike szobájában t időpontban világító villanykörték számát. Írjuk fel az $X(t)$ Markov lánc

   a.) állapotterét,

   b.) infinitezimális generátorát.

   Megoldás:

   a.) Az állapottér $S = \{1, 2\}$, mert ha a rendszer az 1 állapotban van és az egyetlen körte kiég, a rendszer rögtön a 2 állapotba kerül (mert Pistike azonnal mindkét égőt kicseréli).

   b.) Az időt mérjük években. Az 1 állapotból való elugrás rátája $\lambda_1 = \lambda_{12} = 1$, mert az egyetlen körte átlagosan $\frac{1}{\lambda_1} = 1$ év alatt ég ki. Ha viszont két égő működik, akkor ezek *egyike* kétszer ekkora rátával ég ki: $\lambda_2 = \lambda_{21} = 2$. Ezért az infinitezimális generátor

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

3. Ha Móricka kitölt egy IQ-tesztet, az elért eredmény normális eloszlású valószínűségi változó, aminek várható értéke a Móricka (általunk nem ismert) intelligencia-hányadosa, szórása pedig 5. Ha többet is kitölt, az eredmények azonos eloszlásúak és egymástól függetlenek (Móricka nem fejlődik). Móricka rögtön 10-et is kitöltött, és a következő eredményeket kapta: 86, 103, 99, 100, 109, 106, 91, 96, 97, 105. Döntsünk 95%-os szinten arról a hipotézisről, hogy Móricka intelligencia-hányadosa legalább 100. *(Segítség: a fenti számok összege* 992*, négyzetösszege* 98854*.)*

   Megoldás: Egy (m, σ^2) paraméterű normális eloszlából vettünk $n = 10$ elemű mintát, ahol m ismeretlen de $\sigma = 5$ ismert. A nullhipotézis $m \geq \mu := 100$ egyenlőtlenség, ezért egyoldali egymintás u-próbát végzünk, $\varepsilon = 1 - 95\% = 0.05$. Az átlag

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{992}{10} = 99.2$$

   kisebb, mint μ, ezért sajnos végig kell számolni a tesztet annak eldöntésére, hogy ez az eltérés szignifikáns-e. A teszt-statisztika

$$u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{99.2 - 100}{5} \sqrt{10} \approx -0.506,$$

   az elfogadási küszöb

$$K := u_\varepsilon = \Phi^{-1}(1 - \varepsilon) = \Phi^{-1}(0.95) \approx 1.645.$$

   Döntés: $u > -K$, ezért a nullhipotézist *elfogadjuk*.