Diskrétne náhodné vektory

Metódy riešenia úloh
z pravdepodobnosti a štatistiky
Cvičenie 3

Náhodné vektory

- Náhodný vektor $(X_1, X_2, \ldots, X_n)$ – každá jeho zložka je náhodná premenná
- Diskrétny náhodný vektor – nadobúda konečne alebo spočítateľne veľa hodnôt

Príklad 1: Hlasovanie komisie

- V komisii sú 3 osoby – A, B, C, pričom
  - A sa správne rozhoduje s pravd. 0,9
  - B sa správne rozhoduje s pravd. 0,8
  - C sa správne rozhoduje s prav. 0,75
- A, B, C hlasujú nezávisle

Príklad 1: Hlasovanie komisie

- Výsledok hlasovania komisie – náhodný vektor (X_1, X_2, X_3)
- Ak sa i-ty člen komisie rozhodol správne, $X_i = 1$, inak $X_i = 0$.
  - <u>Otázka (a):</u> Aké hodnoty môže nadobúdať tento náhodný vektor a s akými pravdepodobnosťami sa nadobúdajú?

Príklad 1: Hlasovanie komisie

- Predpokladajme, že komisia sa rozhodne tak, ako hlasovala väčšina jej členov.
- Otázka (b): Vypočítajte pravdepodobnosť, že komisia sa rozhodne správne.

Príklad 1: Hlasovanie komisie

- Otázka (c): Člen komisie C si všimol, že jeho hlasovanie sa často ukáže byť nesprávne. Uvažujme tieto dve možnosti:
  - Sedí vedľa člena B, a tak pri hlasovaní odpíše jeho rozhodnutie
    - Rezignuje a namiesto rozmýšľania nad problémom si len hodí mincou.

Pri ktorej z nich je pravd. správneho rozhodnutia komisie väčšia?

Príklad 2: Hlasovanie II.

- Uvažujme pôvodné pravd. a doplňme k pôvodnému vektoru *(X_1, X_2, X_3)* ešte:
  - *Y*, ktorá vyjadruje počet členov, ktorí hlasovali správne
  - Z, ktorá vyjadruje rozhodnutie komisie (0 = nesprávne, 1 = správne)

- Máme teda náhodný vektor *(X_1, X_2, X_3, Y, Z)*

Príklad 2: Hlasovanie II.

- Úlohy:
  a) Nájdite *marginálne rozdelenie Z*
  b) Nájdite podmienené rozdelenie *Z*, za podmienky, že $X_1=0$ – teda rozdelenie výsledku hlasovania, za podmienky, že člen A (najšikovnejší) sa rozhodol nesprávne
  c) Nájdite podmienené rozdelenie *Z*, za podmienky, že $X_1+X_2>0$ – teda rozdelenie výsledku hlasovania, za podmienky, že člen A alebo B sa rozhodol správne

Nezávislosť

- Pripomeňme si definíciu nezávislosti diskrétnych náhodných premenných X, Y:

$$\boxed{P[X = x,\ Y = y] = P[X = x]\,P[Y = y]}$$

pre všetky možné hodnoty x, y

- Užitočná stratégia, ako dokázať, že X, Y nie sú nezávislé: Nájsť hodnoty x, y tak, že pravdep. $P[X = x]$ a $P[Y = y]$ sú nenulové, ale $P[X = x,\ Y = y] = 0$

Príklad 3: Nezávislosť I.

- Pomocou uvedenej stratégie dokážte, že tieto zložky týchto náhodných vektorov nie sú nezávislé:
  a) (Y, Z) z príkladu 2
  b) (X_1, Y) z príkladu 2

- Niekedy sa nedá použiť, treba dokazovať inak - DÚ:
  c) Dokážte, že (X_1, Z) z príkladu 2 nie sú nezávislé.

Príklad 4: Nezávislosť II.

- Dokážte, že zložky nasledujúcich náhodných vektorov nie sú po dvoch nezávislé – teda žiadna dvojica zložiek nie je nezávislá:
  a) Máme 3 červené ruže, 3 ružové a 3 biele. Náhodne z nich vyberieme 5 a spravíme kyticu. *(Č, R, B)* je náhodný vektor vyjadrujúci počet ruží jednotlivých farieb v kytici.

Príklad 4: Nezávislosť II.

b) Hodíme naraz šiestimi kockami. Náhodný vektor *(X_1, X_2, X_3, X_4, X_5, X_6)* označuje počet jednotiek, dvojek, ..., šestiek, ktoré na nich padnú.

c) Hodíme naraz desiatimi kockami. Náhodný vektor *(P, N)* označuje počet párnych a počet nepárnych čísel, ktoré na nich padnú.

d) Hodíme dvakrát kockou. *(S, R)* označuje súčet a rozdiel (prvé číslo mínus druhé) čísel, ktoré padnú.

Príklad 5: Scrabble I.

- V hre Scrabble hráči zostavujú z písmeniek slová.
- Počty písmeniek v slovenskej hre (podľa bodových hodnôt) sú:

- 2 žolíky (za 0 bodov)
- *1 bod*: **A** ×9, **O** ×9, **E** ×8, **I** ×5, **N** ×5, **T** ×4, **R** ×4, **S** ×4, **V** ×4
- *2 body*: **M** ×4, **K** ×3, **L** ×3, **D** ×3, **P** ×3
- *3 body*: **J** ×2, **U** ×2
- *4 body*: **B** ×2, **H** ×1, **Y** ×1, **Z** ×1, **Á** ×1, **C** ×1
- *5 bodov*: **Č** ×1, **Ž** ×1, **Š** ×1, **Í** ×1, **Ý** ×1
- *7 bodov*: **Ľ** ×1, **Ť** ×1, **É** ×1, **Ú** ×1
- *8 bodov*: **Ď** ×1, **F** ×1, **G** ×1, **Ň** ×1, **Ô** ×1
- *10 bodov*: **Ĺ** ×1, **Ŕ** ×1, **X** ×1, **Ä** ×1, **Ó** ×1

Príklad 5: Scrabble I.

- Na začiatku si náhodne vytiahneme 7 písmen.
- Aká je pravdepodobnosť, že to budú:
  - 2 samohlásky a 5 spoluhások
  - 2 samohlásky, 4 samohlásky a žolík
    - obidva žolíky a zvyšné budú spoluhlásky

Príklad 6: Scrabble II.

- V jednej obmene hry sa vytiahne 7 písmen a každý z hráčov sa snaží vytvoriť slovo pomocou týchto písmen.
- Pomienkou je, že medzi týmito 7 písmenami musia byť aspoň dve samohlásky a aspoň dve spoluhlásky, pričom žolík môže podľa potreby predstavovať samohlásku alebo spoluhlásku.

Príklad 6: Scrabble II.

- V prvom kole sa písmená ťahajú z celej sady. Aká je pravdepodobnosť, že to bude treba zopakovať, lebo nebude splnená podmienka na minimálny počet samohlások a spoluhlások?

Society of Actuaries

- SOA – Society of Actuaries, www.soa.org
- Exam P – Probability, www.soa.org/education/exam-req/edu-exam-p-detail.aspx

The syllabus for Exam P develops the candidate's knowledge of the fundamental probability tools for quantitatively assessing risk. The application of these tools to problems encountered in actuarial science is emphasized. A thorough command of the supporting calculus is assumed. Additionally, a very basic knowledge of insurance and risk management is assumed.

Počas semestra vyriešime aj niekoľko príkladov z tejto skúšky

Príklad 7: SOA I.

- Združené rozdelenie počtu tornád v P a Q
- Vypočítajte varianciu počtu tornád v Q za podmienky, že v P neboli žiadne tornáda

		Q			
		0	1	2	3
P	0	0.12	0.06	0.05	0.02
	1	0.13	0.15	0.12	0.03
	2	0.05	0.15	0.10	0.02

Možnosti:

0,51

0,84

0,88

0,99

1,76

Príklad 8: SOA II.

- Poistenie vodičov, ktorí práve dostali vodičský preukaz – nemáme ich charakteristiky, len historické údaje o takýchto vodičoch, z ktorých sa odhadli pravdepodobnosti:
    - » 0,5 nízke riziko
    - » 0,3 stredné riziko
    - » 0,2 vysoké riziko

Príklad 8: SOA II.

- Poistenie uzavreli štyria takíto vodiči
- Prepokladáme, že ich zaradenie podľa rizika je nezávislé a riadi sa uvedeným rozdelením.

Aká je pravdep., že v do kategórie vysokého rizika spadá aspoň o 2 viac vodičov ako do kategórie nízkeho rizika?

(A) 0.006 (B) 0.012 (C) 0.018 (D) 0.049 (E) 0.073

Multinomické rozdelenie

- Máme *n* guľôčok, umiestňujeme ich nezávisle do *k* priehradiek
- Pravdepodobnosť, že guľôčku umiestnime do *i*-tej priehradky je p_i
  - Pravdapodobnosť, že v jednotlivých priehradkách bude $x_1, ..., x_k$ guľôčok, je:

$$\frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

Multinomické rozdelenie

- Príklad:

  Uvažujme pravdepodobnosti z príkladu 6. Máme 15 vodičov. Aká je pravdepodobnosť, že 9 z nich patrí do kategórie nízkeho rizika, 3 do kategórie stredného rizika a 3 do kategórie vysokého rizika?

Príklad 9: Hádzanie kockami

- Hodíme šiestimi kockami.
  - Aká je pravdepodobnosť, že padnú dve dvojky, dve trojky, päťka a šestka?
  - Aká je pravdepodobnosť, že padnú tri jednotky a tri šestky?

Viacrozmerné hypergeometrické rozdelenie

- V urne je spolu *m* guľôčok, z nich m_1 má farbu 1, m_2 farbu 2, ..., m_k farbu k.
- Bez vracania vyberieme *n* guľôčok.

Pravdep., že z nich j_1 má farbu 1, ..., j_k má farbu k, je:

$$\frac{\binom{m_1}{j_1}\binom{m_2}{j_2}\cdots\binom{m_k}{j_k}}{\binom{m}{n}}$$

Príklad 10: Kytica

- Máme 3 červené ruže, 3 ružové a 3 biele. Náhodne z nich vyberieme 5 a spravíme kyticu.
- Aká je pravdepodobnosť, že v kytici bude jedna červená, dve ružové a dve biele ruže?

Príklad 11: Výber s vracaním a bez vracania

- V urne je spolu m guľôčok, z nich m_1 má farbu 1, m_2 farbu 2, ..., m_k farbu k.
- Vyberieme n z nich – vieme, že ak je to výber:
  - s vracaním → multinomické
  - bez vracania → viacrozmerné hypergeometrické

Príklad 11: Výber s vracaním a bez vracania

- <u>Intuícia:</u>

  Pri veľkom počte guľôčok by v tom nemal byť veľký rozdiel, či guľôčky vrátime alebo nie.

Príklad 11: Výber s vracaním a bez vracania

- <u>Presne:</u>
  - Fixujme *n* (veľkosť výberu)
  - Nechajme *m* (celkový počet guľôčok) rásť do nekonečna tak, že pritom $m_i/m \rightarrow p_i$.
  - Potom pravdepodobnosti viacrozmerného hypergeometrického rozdelenia konvergujú k pravd. multinomického rozdelenia. Dokážte toto tvrdenie.

Príklad 11: Výber s vracaním a bez vracania

- <u>Numerická ukážka - DÚ:</u>
  - Ilustrujte predchádzajúce tvrdenie numericky.
  - Pre zvolené „dosť veľké“ *m* (treba vyskúšať, aké musí byť) vypočítajte pravd. niekoľkých zvolených udalostí presne pomocou viacrozmerného hypergeometrického rozdelenia a porovnajte s multinomických, kde $p_i = m_i/m$. Mali by byť podobné.

Príklad 12: SOA III.

- Ešte jeden príklad zo skúšky SOA:
  - V predajni áut je počet predaných áut za deň 0, 1 alebo 2.
  - Ak predajú auto, ponúknu kupujúcemu predĺženú záruku
  - Označme X počet predaných áut a Y počet predaných predĺžených záruk.

Príklad 12: SOA III.

– Združené rozdelenie vektora *(X, Y)* je:

$$P(X=x, Y=y) = \begin{cases} 1/6 & (x,y)=(0,0), \\ 1/12 & (x,y)=(1,0), \\ 1/6 & (x,y)=(1,1), \\ 1/12 & (x,y)=(2,0), \\ 1/3 & (x,y)=(2,1), \\ 1/6 & (x,y)=(2,2). \end{cases}$$

– Vypočítajte disperziu náhodnej premennej *X*, teda počtu predaných áut.