**Matematicko-štatistické metódy spracovania dát 2019/2020**

**Cvičenie 6.**

Cvičenie je zamerané na náhodné premenné – určenie rozdelenia NP, strednej hodnoty, výpočet pravdepodobnosti z distribučnej funkcie (spojitá NP) resp. pravdepodobnostnej funkcie (diskrétna NP), stredná hodnota a rozptyl NP.

1. Pri trojnásobnom hode vyváženou mincou je možných 8 výsledkov, ktoré sú rovnako pravdepodobné. Napríklad HHZ znamená, že prvý výsledok je *hlava*, druhý *hlava* a tretí *znak*. Nech *X* predstavuje dosiahnutý počet *hláv* v troch hodoch.

a) Aké hodnoty môže nadobúdať náhodná premenná *X*?

b) Nájdite pravdepodobnostné rozdelenie náhodnej premennej *X*.

c) Vypočítajte $P\left(X=2\right).$

d) Vypočítajte pravdepodobnosť, že padnú nanajvýš dve *hlavy*.

2. Nech náhodná premenná *Y* predstavuje počet osôb žijúcich v náhodne vybratej obytnej jednotke. Jej pravdepodobnostné rozdelenie je uvedené v tabuľke

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| y | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| P(Y=y) | 0,265 | 0,327 | 0,161 | 0,147 | 0,065 | 0,022 | 0,013 |

a) Nájdite a interpretujte strednú hodnotu náhodnej premennej *Y.*

b) Nájdite a interpretujte smerodajnú odchýlku náhodnej premennej.

3. Obsah plochy pod hustotovou krivkou pre určitú náhodnú premennú naľavo od hodnoty 105 je 0,6227. Koľko percent všetkých pozorovaní náhodnej premennej s touto funkciou hustoty má hodnoty napravo od hodnoty 105?

4. Doplňte tabuľku pre normovanú normálne rozdelenú náhodnú premennú.

|  |  |  |
| --- | --- | --- |
| Interval | Obsah plochy | Percento celkovej plochy |
| (-∞; -3) | 0,0013 | 0,13 |
| (-3; -2) |  |  |
| (-2; -1) |  |  |
| (-1; 0) |  |  |
| (0; 1) | 0,3413 | 34,13 |
| (1; 2) |  |  |
| (2; 3) |  |  |
| (3; ∞) |   |   |

Koľko percent hodnôt náhodnej premennej $X\~N\left(μ;σ^{2}\right)$ sa nachádza v intervale $\left(μ-kσ;μ+kσ\right)$ sa pre hodnoty $k=1,2,3$.

5. Nech $Z\~N(μ=0;σ^{2}=1)$. Doplňte tabuľku kritických hodnôt $u\_{α}$ náhodnej premennej *Z*.

|  |  |  |  |
| --- | --- | --- | --- |
| $$u\_{0,10}$$ | $$u\_{0,05}$$ | $$u\_{0,02}$$ | $$u\_{0,01}$$ |
| 1,644 |  |  |  |

6. Čas bežcov v 10 km preteku New Yorkom má normálne rozdelenie so strednou hodnotou 61 minút a smerodajnou odchýlkou 9 minút. Nech *X* predstavuje dosiahnutý čas bežcov.

a) Načrtnite hustotovú funkciu náhodnej premennej *X*.

b) Vyjadrite štandardizovanú verziu (*Z*) náhodnej premennej *X*.

c) Načrtnite rozdelenie náhodnej premennej *Z*.

d) „Podiel bežcov, ktorí dosiahli čas medzi 50 min. a 70 min. sa rovná obsahu plochy pod krivkou štandardizovaného normálneho rozdelenia medzi ............. a .......... . Doplňte!

e) Koľko percent bežcov zabehlo trať pod 70 minút?

7. (Samostatne) Jedným z najväčších druhov tarantúl je *Grammostola mollicoma*. Telo tarantule sa skladá z dvoch častí. Predná časť tela je pokrytá škrupinou. V článku "*Reproductive Biology of Uruguayan Theraposids*" (Journal of Arachnology) autori Costa a Perez-Miles píšu, že dĺžka panciera dospelého samčeka *G. Mollicoma* má normálne rozdelenie so strednou hodnotou 18,14 mm a smerodajnou odchýlkou 1,76 mm.

a) Určte podiel dospelých samčekov *G.mollicoma*, ktorých dĺžka panciera je od 16 mm do 17 mm.

b) Určte podiel dospelých samčekov *G.mollicoma*, ktorých dĺžka panciera je väčšia ako 19 mm.

c) Vypočítajte a interpretujte kvartily dĺžky panciera dospelého samčeka *G.mollicoma*.

d) Určte 95-ty percentil dĺžky panciera dospelého samčeka *G.mollicoma*.

8. Karl Pearson (1857-1936), jeden z otcov modernej štatistiky skúmal oblačnosť (denná úroveň oblačnosti) počas dňa na stupnici 0-10 (0-žiadna oblačnosť, 10- celodenná oblačnosť) vo Vroclavi počas dekády 1876-1885. Rozdelenie početností je uvedené v tabuľke.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stupeň | Počet dní |   | Stupeň | Počet dní |
| 0 | 751 |  | 6 | 21 |
| 1 | 179 |  | 7 | 71 |
| 2 | 107 |  | 8 | 194 |
| 3 | 69 |  | 9 | 117 |
| 4 | 46 |  | 10 | 2089 |
| 5 | 9 |   |   |   |

Na základe histogramu pre dáta uvedené v tabuľke rozhodnite, či stupeň oblačnosti má približne normálne rozdelenie.

9. Z databázy zemetrasení s magnitúdom aspoň 5,5 v období rokov 1965-2016 je možné zistiť, že medzi nastaním dvoch za sebou nasledujúcich zemetrasení uplynie v priemere 0,81 dňa (približne 19,5 hodiny) so smerodajnou odchýlkou 0,97 dňa (približne 23,3 hodiny). Vysvetlite, prečo iba na základe hodnôt týchto dvoch parametrov je nereálne predpokladať, že čas ktorý uplynie medzi dvomi za sebou nasledujúcimi zemetraseniami má normálne rozdelenie.