

BPM_STAE: Koncepty a procedury - cvičení 4, kapitola 3

1. Jak se vypočítá medián pro data s lichým počtem pozorování? Jak se vypočítá medián pro data se sudým počtem pozorování?
2. Stručně vysvětlete, co znamená odlehlá hodnota. Který z ukazatelů středu dat (průměr nebo medián) je vhodnější, pokud data obsahují odlehlé hodnoty? Uveďte příklad.
3. Na příkladu ukažte, jak mohou odlehlé hodnoty ovlivnit hodnotu průměru.
4. Který z pěti ukazatelů středu (průměr, medián, oříznutý průměr, vážený průměr, modus) lze vypočítat pouze pro kvantitativní data? Které ukazatele středu lze vypočítat pro kvantitativní i kvalitativní data? Uveďte příklady.
5. Který z ukazatelů středu (průměr, medián, oříznutý průměr, vážený průměr, modus) může mít více než jednu hodnotu pro datovou sadu? Uveďte příklad datové sady, kde daný ukazatel středu nabývá více než jedné hodnoty.
6. Je možné, aby kvantitativní datová sada neměla průměr, medián nebo modus? Uveďte příklad datové sady, pro kterou některý z těchto ukazatelů středu neexistuje.
7. Vysvětlete vztahy mezi průměrem, mediánem a modem pro symetrické a asymetrické histogramy. Uveďte grafické ilustrace.
8. Ceny automobilů mají rozdělení, které je nakloněné doprava, s odlehlými hodnotami v pravém chvostu. Který z ukazatelů středu je nejlepší pro shrnutí této datové sady? Vysvětlete.
9. Následující datová sada patří do populace. Vypočítejte průměr, medián a modus.

5 -7 2 0 -9 16 10 7

10. Rozptyl jako míra variability má nevýhodu v tom, že je ovlivněn odlehlými hodnotami. Ilustrujte na příkladu, jak odlehlé hodnoty ovlivňují rozptyl.
11. Může mít směrodatná odchylka zápornou hodnotu? Vysvětlete.
12. Kdy je hodnota směrodatné odchylky pro datovou sadu nulová? Uveďte příklad a proveďte výpočet směrodatné odchylky, která je nulová.
13. Stručně vysvětlete rozdíl mezi populačním parametrem a statistikou získanou ze vzorku. Uveďte příklad každého.
14. Následující datová sada patří do populace. Vypočítejte rozpětí, rozptyl a směrodatnou odchylku.

5 -7 2 0 -9 16 10 7

15. Stručně vysvětlete Chebyshevovu větu a její aplikace.
16. Stručně vysvětlete empirické pravidlo. Na jaký typ rozdělení se aplikuje?
17. Výběr 2000 pozorování má průměr 74 a směrodatnou odchylku 12. pomocí Chebyshevovy věty určete minimální procento pozorování, které spadá do intervalů $\bar{x} \pm 2s$, $\bar{x} \pm 2.5s$, $\bar{x} \pm 3s$. Poznámka: interval $\bar{x} \pm 2s$ zahrnuje interval od $\bar{x} + 2s$ do $\bar{x} - 2s$ a tak dále.
18. Velká populace má průměr 230 a směrodatnou odchylku 41. Pomocí Chebyshevovy věty určete minimální procento pozorování, které spadá do intervalů $\mu \pm 2\sigma$, $\mu \pm 2.5\sigma$ a $\mu \pm 3\sigma$.
19. Velká populace má rozdělení ve tvaru zvonu s průměrem 310 a směrodatnou odchylkou 37. Pomocí empirického pravidla určete přibližné procento pozorování, které spadá do intervalů $\mu \pm 1\sigma$, $\mu \pm 2\sigma$ a $\mu \pm 3\sigma$.
20. Výběr 3000 pozorování má rozdělení ve tvaru zvonu s průměrem 82 a směrodatnou odchylkou 16. Pomocí empirického pravidla určete přibližné procento pozorování, které spadá do intervalů $\bar{x} \pm 1s$, $\bar{x} \pm 2s$, $\bar{x} \pm 3s$.

21. Stručně popište, jak se vypočítávají tři kvartily pro datovou sadu. Ilustrujte na dvou příkladech: jeden s lichým počtem pozorování, druhý se sudým počtem pozorování.
22. Vysvětlete, jak se vypočítá mezikvartilové rozpětí. Uveďte jeden příklad.
23. Stručně popište, jak se vypočítají percentily pro datovou sadu.
24. Vysvětlete koncept percentilového pořadí pro pozorování v datové sadě.
25. Stručně vysvětlete, jaké shrnující ukazatele se používají k vytvoření krabicového grafu (Boxplot).
26. Připravte krabicový graf (Boxplot) pro následující data. Obsahuje tato datová sada nějaké odlehlé hodnoty?

36	43	28	52	41	59	47	61	24	55	63	73
32	25	35	49	31	22	61	42	58	65	98	34