

Pravděpodobnost a statistika - verze 240603A

(0 – 60) bodů, požadované minimum: 30 bodů

Praktická část - (0 – 50) bodů

1. Dle historických dat očekáváme, že na jistém místě v poušti prší v průměru pět dní v roce (předpokládejme, že rok má 365 dní). Ve dnech, kdy prší, meteorologové předvíдали déšť v 90 % případů. Meteorologové však předvíдали déšť i v 10 % dnů, kdy ve skutečnosti nepršelo.
 - a) V kolika procentech dnů v roce meteorologové předvídají na daném místě v poušti déšť? **(5b)**
 - b) Marie se zítra vdává a má na daném místě v poušti naplánován venkovní obřad. Bohužel meteorologové předpovídají na zítřek déšť. Jaká je pravděpodobnost, že v den Mariiny svatby bude na daném místě v poušti pršet? **(5b)**
2. Distribuční funkce náhodné veličiny X je dána předpisem:

$$F_X(x) = \begin{cases} 0, 0 & \text{pro } x \in (-\infty, -2) \\ 0, 1 & \text{pro } x \in (-2, 2) \\ 0, 5 & \text{pro } x \in (2, 5) \\ c & \text{pro } x \in (5, \infty) \end{cases}$$

- a) Určete hodnotu konstanty c . **(1b)**
 - b) Určete předpis pravděpodobnostní funkce $P_X(x)$ náhodné veličiny X . **(2b)**
 - c) Načrtněte graf pravděpodobnostní funkce náhodné veličiny X . **(1b)**
 - d) Určete pravděpodobnost, že náhodná veličina X nabývá hodnoty z intervalu $(-\infty, 4)$. **(2b)**
 - e) Určete modus \hat{x} , střední hodnotu $E(X)$ a rozptyl $D(X)$ náhodné veličiny X . **(4b)**
3. Raketa typu Schmayd má šest motorů. Maximální dobu hoření jednoho motoru označme X . Předpokládejme, že $X \sim N(\mu, \sigma^2)$, kde $\mu = 11$ s a $\sigma = 1$ s.
 - a) Načrtněte graf hustoty pravděpodobnosti a graf distribuční funkce doby hoření jednoho motoru. **(2b)**
 - b) Určete pravděpodobnost, že motor bude hořet nejméně 10 s (pro další výpočty tuto pravděpodobnost zaokrouhlete na dvě desetinná místa) a vyznačte tuto pravděpodobnost v obou grafech načrtnutých v úloze (a). **(3b)**
 - c) Určete dobu hoření motoru, která bude překročena s pravděpodobností 30 %. **(2b)**
 - d) Aby se raketa dostala na oběžnou dráhu, potřebuje, aby alespoň 5 (z 6) jejích motorů hořelo nejméně 10 s. Určete pravděpodobnost, že se raketa dostane na oběžnou dráhu. **(3b)**

Popis datového souboru potřebného pro úlohy 4 a 5:

Byla provedena studie hodnotící výnosy čtyř odrůd pšenice (A, B, C, D). Každá odrůda byla pěstována v letech 2022 a 2023 na 100 polích. V roce 2023 byly z hlediska počasí výrazně lepší podmínky, což pravděpodobně přispělo k vyšším výnosům.

V datovém souboru jsou pro každé pole uvedeny následující údaje:

- *id* ... identifikační číslo pole,
- *odruda* ... název odrůdy, která byla na tomto poli pěstována (A/B/C/D),
- *vynos_2022* ... výnos pole v roce 2022 v tunách na hektar (t/ha),
- *vynos_2023* ... výnos pole v roce 2023 v tunách na hektar (t/ha),
- *typ_hnojiva* ... typ hnojiva použitý na poli (syntetické/přirodní).

V následujících úkolech bude zapotřebí analyzovat meziroční zvýšení výnosu, tedy

$$\text{zvyseni} = \text{vynos_2023} - \text{vynos_2022} \text{ (t/ha)}.$$

- Na základě datového souboru analyzujte meziroční **zvýšení výnosů pro pšenici odrůdy A a odrůdy B**. Nezapomeňte ověřit předpoklady pro použití metod statistické indukce.
 - Identifikujte odlehlá pozorování a z následující analýzy je vyřaďte. Uveďte *id* polí, která nebyla do další analýzy zařazena. **(1b)**
 - Určete bodové a 95% oboustranné intervalové odhady středních hodnot (popřípadě mediánů) **zvýšení výnosu** pšenic odrůd **A** a **B**. Porovnejte oba intervalové odhady a okomentujte. **(3b)**
 - Určete bodový a 95% jednostranný intervalový odhad rozdílu středních hodnot $\mu_A - \mu_B$ (popřípadě rozdílu mediánů $x_{0,5}^A - x_{0,5}^B$) **zvýšení výnosů** pšenic odrůd **A** a **B**. **(3b)**
 - Výrobce odrůdy **A** tvrdí, že jejich odrůda pšenice umí využít lepších podmínek lépe (a tedy má větší meziroční **zvýšení** výnosu) než její přímá konkurence, odrůda **B**. Zhodnoťte toto tvrzení **v návaznosti na intervalový odhad** z minulého bodu a doplňte váš závěr o příslušný čistý test významnosti na hladině významnosti 5 %. Výsledek okomentujte. **(3b)**
- Proveďte analýzu meziročního **zvýšení** výnosu u pšenice **všech odrůd**. Na hladině významnosti 5 % rozhodněte, zda se střední hodnoty nebo mediány meziročních zvýšení výnosů statisticky významně liší mezi odrůdami. Pokud ano, **identifikujte homogenní skupiny odrůd a seřadte je sestupně podle zvýšení výnosů**. Nezapomeňte na odstranění odlehlých pozorování a ověření předpokladů metod statistické indukce. **(10b)**