

# A6M33SSL: Domácí úloha: DU6

Varianta: elhajzen

1. [stat\_test\_cars, 0.300 b.] Naše malá začínající autopůjčovna půjčuje tři různé automobily, 1, 2, a 3. Chceme zjistit, jestli se jejich spotřeba liší. Proto vždy, když zákazník automobil vrátí, zjistíme počet najetých kilometrů, množství spotřebovaného paliva a vypočteme spotřebu  $X$  automobilu. Po jisté době máme  $n_1$ ,  $n_2$  a  $n_3$  naměřených hodnot pro jednotlivé automobily. Z dat vypočtené průměrné spotřeby v jednotlivých skupinách jsou  $m_1$ ,  $m_2$  a  $m_3$ , podobně směrodatné odchylky  $s_1$ ,  $s_2$  a  $s_3$ . Pomocí analýzy rozptylu na hladině významnosti  $\alpha = a$  ověřte hypotézu, že automobily mají stejnou průměrnou spotřebu.

- Určete celkovou průměrnou spotřebu  $m$ .
- Určete odhad rozptylu  $s_B^2$  vycházející z rozdělení výběrových průměrů (tzv. vysvětlený nebo meziskupinový rozptyl).
- Určete odhad rozptylu  $s_W^2$  vycházející z rozptylů v jednotlivých skupinách (tzv. sdružený odhad rozptylu, nevysvětlený, reziduální nebo vnitroskupinový rozptyl).
- Určete hodnotu  $t$  testové statistiky.
- Kolik stupňů volnosti má čitatel ( $dof_B$ ) a jmenovatel ( $dof_W$ )?
- Pro hladinu významnosti  $\alpha = a$  zjistěte kritickou hodnotu testu  $k$ .
- Rozhodněte, zda můžete zamítnout  $H_0$  (pak do JSON souboru uveďte výsledek "z": "ano"), nebo nemůžete (JSON: "z": "ne")

**Parametry:**  $n_1 = 27$ ,  $n_2 = 25$ ,  $n_3 = 25$ ,  $m_1 = 7.9$ ,  $m_2 = 7.5$ ,  $m_3 = 8.3$ ,  $s_1 = 0.7$ ,  $s_2 = 0.8$ ,  $s_3 = 0.7$ ,  $a = 0.01$

**Požadované výsledky:**  $m$ ,  $s_B^2$ ,  $s_W^2$ ,  $t$ ,  $dof_B$ ,  $dof_W$ ,  $k$ ,  $z$

2. [stat\_test\_dice, 0.150 b.] Opakovaně jsme házeli hrací kostkou a pozorovali jsme  $n_1, \dots, n_6$  výskytů čísel 1 až 6. Pro  $\alpha = a$  ověřte hypotézu, že kostka je férová.

- Vypočtete hodnotu testové statistiky  $t$ .
- Zjistěte kritickou hodnotu testu  $k$ .
- Rozhodněte, zda můžete zamítnout  $H_0$  (pak do JSON souboru uveďte výsledek "z": "ano"), nebo nemůžete (JSON: "z": "ne").

**Parametry:**  $n_1 = 287$ ,  $n_2 = 300$ ,  $n_3 = 278$ ,  $n_4 = 305$ ,  $n_5 = 248$ ,  $n_6 = 304$ ,  $a = 0.001$

**Požadované výsledky:**  $t$ ,  $k$ ,  $z$

3. [stat\_test\_iqpartners, 0.250 b.] Studujeme inteligenci mužů a žen - manželských párů. Z populace jsme náhodně vybrali  $n$  manželských párů. Každému z manželů jsme změřili inteligenci na stupnici A - podprůměrná, B - průměrná, C - nadprůměrná. Výsledky měření jsou dány jako počty  $n(i, j)$ , kde  $i, j \in \{A, B, C\}$ , takže např.  $n(A, C)$  je počet párů, kde muž má inteligenci podprůměrnou, zatímco žena nadprůměrnou. Na hladině významnosti  $\alpha = a$  ověřte, zda inteligence jednoho z partnerů závisí na inteligenci druhého.

- Vypočtete hodnotu testové statistiky  $t$ .
- Zjistěte kritickou hodnotu testu  $k$ .

f) Rozhodněte, zda můžete zamítnout  $H_0$  (pak do JSON souboru uveďte výsledek "z\_2": "ano"), nebo nemůžete (JSON: "z\_2": "ne").

**Parametry:**  $n_{A,A} = 70$ ,  $n_{A,B} = 46$ ,  $n_{A,C} = 14$ ,  $n_{B,A} = 23$ ,  $n_{B,B} = 128$ ,  $n_{B,C} = 27$ ,  $n_{C,A} = 24$ ,  $n_{C,B} = 47$ ,  $n_{C,C} = 67$ ,  $a = 0.01$

**Požadované výsledky:**  $t$ ,  $k$ ,  $z$

4. [stat\_test\_lactacid, 0.300 b.] Měříme koncentrace kyseliny mléčné [mg/dl] v krvi matek ( $X$ ) a novorozenců ( $Y$ ). Máme k dispozici realizaci náhodného výběru z dvourozměrného rozdělení  $(X, Y)$ . Pro  $\alpha = a$  posuďte, zda jsou veličiny  $X$  a  $Y$  korelované.

a) Vypočtete výběrové směrodatné odchylky  $s_x$  a  $s_y$ .

b) Vypočtete výběrový koeficient kovariance  $c$ .

c) Vypočtete výběrový koeficient korelace  $r$ .

d) Určete hodnotu testové statistiky  $t$ .

e) Pro  $\alpha = a$  určete kritickou hodnotu testu  $k$ .

f) Rozhodněte, zda můžete zamítnout  $H_0$ , že veličiny nejsou korelované (pak do JSON souboru uveďte výsledek "z": "ano"), nebo nemůžete (JSON: "z": "ne")

**Parametry:**  $x = [41, 53, 20, 37]$ ,  $y = [48, 48, 16, 35]$ ,  $a = 0.05$

**Požadované výsledky:**  $s_x$ ,  $s_y$ ,  $c$ ,  $r$ ,  $t$ ,  $k$ ,  $z$