

## Analýza dat pro Neurovědy – písemná část kolokvia

Písemnou část kolokvia je nutno odevzdat v elektronické podobě nejpozději 2 dny před ústní částí kolokvia. Z písemné části je nutno dosáhnout alespoň 60% bodů (tedy 18 z celkových 30 bodů).

**Příklad 1** (8 bodů). K následujícím zadáním analytických úloh doplňte vhodný statistický test včetně toho, zda je o test oboustranný či jednostranný:

- Chceme srovnat vliv třech typů cvičení na nárůst svalové hmoty. (Předpoklady pro použití parametrického testu nejsou splněny a nepomůže ani logaritmická transformace dat).
- Chceme zjistit, zda je hladina celkového cholesterolu vyšší u mužů než u žen. (Předpoklady pro použití parametrického testu jsou splněny).
- Chceme zjistit, zda vedla dieta ke snížení systolického tlaku u skupiny pacientů. Máme k dispozici hodnoty systolického tlaku před zahájením diety a po skončení diety. (Předpoklady pro použití parametrického testu jsou splněny).
- Chceme zjistit, zda ženy s chronickou myeloidní leukémií trpí anémií, tedy zda mají hladinu hemoglobinu v krvi nižší než 120 g/l. (Předpoklady pro použití parametrického testu nejsou splněny a nepomůže ani logaritmická transformace dat).
- U skupiny pacientů jsme změřili hmotnost před operací a čtrnáct dní po operaci. Chceme zjistit, zda vedla operace ke změně váhy pacientů. (Předpoklady pro použití parametrického testu nejsou splněny a nepomůže ani logaritmická transformace dat).
- Chceme zjistit, zda se liší počet bílých krvinek v 1 ml krve u pacientů s chronickou myeloidní leukémií a chronickou leukocytární leukémií. (Předpoklady pro použití parametrického testu nejsou splněny a nepomůže ani logaritmická transformace dat).
- Chceme ověřit, zda kuřáci v našem souboru netrpí hypertenzí, tzn., zda diastolický tlak u kuřáků v našem souboru je menší než 90 mmHg. (Předpoklady pro použití parametrického testu jsou splněny).
- Chceme srovnat výšku v jednom roce u donošených dětí, u předčasně narozených dětí, které se narodily po 30. týdnu těhotenství, a u předčasně narozených dětí, které se narodily před 30. týdnem těhotenství. (Předpoklady pro použití parametrického testu jsou splněny).

**Příklad 2** (3 body). Pilotní fáze screeningového programu karcinomu prsu se zúčastnilo 2000 žen. Výsledky jsou uvedené v tabulce. **Spočtete senzitivitu, specificitu, prediktivní hodnoty a věrohodnostní poměry a výsledky okomentujte.**

Tabulka 1. Výsledky pilotní fáze screeningového programu karcinomu prsu.

| Skutečnost – přítomnost karcinomu prsu | Výsledek mamografického vyšetření |                    | Celkem |
|--|-----------------------------------|--------------------|--------|
|  | Pozitivní výsledek                | Negativní výsledek |        |
| Ano                                    | 13                                | 2                  | 15     |
| Ne                                     | 237                               | 1748               | 1985   |
| Celkem                                 | 250                               | 1750               | 2000   |

Senzitivita =  
Specificita =  
Pozitivní prediktivní hodnota =  
Negativní prediktivní hodnota =  
Věrohodnostní poměr pozitivního testu =  
Věrohodnostní poměr negativního testu =

**Příklad 3** (4 body). **Datový soubor Data\_kardio\_1.xls zvalidujte** (tzn. datový soubor připravte pro analýzu) **a sumarizujte všechny parametry (proměnné) pomocí popisné statistiky**.  
Použijte soubor Data\_kardio\_1.xls.

**Příklad 4** (3 body). **Zhodnoťte vztah systolického tlaku (sys\_tlak) a věku** graficky i statistickou metodou (včetně p-hodnoty). Svůj výběr statistické metody okomentujte.  
Použijte soubor Data\_kardio\_2.xlsx.

**Příklad 5** (4 body). **Ověřte, zda existuje vztah mezi glykemií** (parametr glykemie\_norma, kde 0 znamená „pod normou“, 1 znamená „v normě“ a 2 znamená „nad normou“) **a pohlavím**. Dále kategorie u glykémie vhodně slučte tak, aby vznikly dvě kategorie, a ověřte vztah výsledné proměnné a pohlaví.  
Použijte soubor Data\_kardio\_2.xlsx.

**Příklad 6** (4 body). **Zjistěte, zda se liší diastolický tlak u třech věkových kategorií lidí (<30 let, 30-50 let, >=50 let)**. Pokud budete chtít použít parametrický test, nezapomeňte ověřit všechny předpoklady. Proveďte i popisnou sumarizaci diastolického tlaku podle kategorií věku. Zhodnoťte, zda má smysl počítat intervaly spolehlivosti pro průměr. Vykreslete krabicové grafy.  
Použijte soubor Data\_kardio\_3.xlsx.

**Příklad 7** (4 body). **Zjistěte, zda se liší BMI podle pohlaví**. Pokud nejsou splněny předpoklady parametrického testu, zkuste BMI logaritmovat. Zhodnoťte, zda má smysl počítat intervaly spolehlivosti pro průměr. Vykreslete krabicové grafy.  
Použijte soubor Data\_kardio\_3.xlsx.