

Datum:	Příjmení, jméno:	Body celkem:
--------	------------------	--------------

Analýza experimentálních dat, cvičení 3

Soubor dat: data.csv

Popis dat: Skupina zdravých mluvčích (HC) a skupina pacientů s Parkinsonovou nemocí (PD) byly podrobeny diadochokinetickému testu (DDK) a kvalita jejich artikulace byla vyhodnocena příznakem *voice onset time* (VOT) měřeného automatickou metodou (Novotný et al. 2015).

Zadání úlohy	body
<p>Data si různými způsoby vizualizujte.</p> <p><i>Je viditelný efekt skupin?</i></p> <p>Spočítejte <i>Cohenovo d</i> mezi skupinami HC a PD.</p> <p><i>Jaký efekt je mezi skupinami?</i></p>	0.5
<p>Odhadněte parametry normálního rozdělení VOT pro skupinu PD a HC, určete jeho kumulativní distribuční funkci v rozsahu Vašich dat a odpovězte na otázku:</p> <p><i>Patří subjekt X do skupiny HC nebo PD nebo do žádné z nich?</i></p>	0.75
<p>Implementujte bootstrapping pro testování rozdělení 2 skupin.</p> <p>Navrhněte hypotézu (H0), reportujte p-hodnotu a odpovězte následující otázku:</p> <p><i>Liší se řečový příznak VOT významně mezi skupinami HC a PD?</i></p> <p>Proveďte stejnou analýzu pomocí dvouvýběrového t-testu, reportujte p-hodnotu a porovnejte výsledek s bootstrappingem.</p>	0.75
<p>Nepovinný bonus:</p> <p>Kam byste subjekt X zařadili s ohledem na prevalenci PD 0.3% v populaci?</p> <p>Nápověda (Bayesova věta):</p> $P(A x) = \frac{P(x A) \cdot P(A)}{P(x A) \cdot P(A) + P(x B) \cdot P(B)} \quad (1)$ <p>kde $P(A x)$ je pravděpodobnost nemoci s ohledem na výskyt, $P(x A)$ je pravděpodobnost nemoci určená testem, $P(A)$ je pravděpodobnost výskytu nemoci, $P(x B)$ je pravděpodobnost zdravého stavu určená testem a $P(B)$ je pravděpodobnost zdravého stavu.</p>	1

Reference

Novotný, M., Ruzs, J., Čmejla, R., and Růžička, E. (2014). Automatic evaluation of articulatory disorders in Parkinson's disease. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 22, 1366-1378.