

Zadání Cvičení #5

Popis dat: Pracovní data jsou uložena v souboru **data.csv**, který je k dispozici ke stažení na Moodle stránce tohoto předmětu, ve složce příslušného cvičení. Pro načtení dat do Matlabu využijte funkce `readtable`.

```
T = readtable('data.csv', 'ReadVariableNames', true);
```

Data jsou tentokrát získána od třech skupin:

- Pacienti trpící [Huntingtonovou nemocí](#) (datový label HD)
- Pacienti trpící Parkinsonovou nemocí (datový label PD)
- Kontrolní skupina zdravých lidí (datový label HC)

Pacienti trpící PN byli klinicky vyšetřeni pomocí dotazníku *UPDRS*, který znáte už z předchozích cvičení.

Pacienti trpící HN byli klinicky vyšetřeni za pomoci dotazníku *UHDRS* ([Unified Huntington's Disease Rating Scale](#)), což je principiálně stejná škála jako *UPDRS*, jen pro hodnocení HN.

V tabulce máte ke všem subjektům k dispozici ID, label, hodnoty z dotazníků (jen pro ty skupiny, kterých se dotazníky týkají, ostatní mají hodnotu *NaN*) a hodnoty následujících vypočtených parametrů:

- **RI** – *Rhythm acceleration* ([Rusz et al. 2015](#), viz Reference)
 - Míra zrychlování řeči (držení rytmu) během řečového rytmického testu, detailní popis výpočtu viz článek, jednotkou jsou procenta.
 - Vyšší hodnoty zde ukazují na menší schopnost držet se stálého tempa při DDK testu, a tedy odpovídají i většímu postižení.
- **DVA** – *Degree of vocal arrest* ([Rusz et al. 2014](#), viz Reference)
 - Podíl ticha na celkovém řečovém záznamu při fonaci (vyluzování zvuku písmene) v procentech.
 - Vyšší hodnoty zde vypovídají o větším množství hlasových „záseků“ při fonaci, a tedy i většímu motorickému postižení.
- **MPT** – *Maximum phonation time* ([Rusz et al. 2014](#), viz Reference)
 - Maximální délka fonace (vyluzování zvuku) písmene „a“ v sekundách.
 - Vyšší hodnoty zde odpovídají menšímu motorickému postižení.
- **stdF0** – *Standard deviation of fundamental frequency* (Hodnoceno pomocí programu PRAAT, [Boersma and Weenink 2017](#), viz Reference)
 - Míra toho jako moc se mění intonace řeči během monologu vyšetřovaného, směrodatná odchylka fundamentální frekvence F0 s jednotkou *st* (*semitones*, viz reference).
 - Vyšší hodnoty zde odpovídají větší kontrole nad motorickými schopnostmi, a tedy i menšímu postižení.

Důležité: Pokud v průběhu analýzy zjistíte, že vyšetřování nějakého parametru nedává dále smysl, rozhodně ho nevyšetřujte dál a argumentujte proč, nebo to nejlépe konzultujte během cvičení.

Zbytečně si nepřidělávejte práci. 😊

Zadání úlohy	body
<p>Provedte vizualizaci dat. Vizualizaci proveďte s důrazem na čitelnost a přehlednost. Cílem je vidět co nejlépe případné rozdíly mezi skupinami. Doporučujeme využít zobrazení pomocí <code>boxplot</code> či <code>boxchart</code> a rozvržení grafů pomocí funkce <code>tiledlayout</code>.</p> <p>Vizualizaci proveďte pro všechny 4 měřené akustické parametry: RI, DVA, MPT a stdF0.</p> <p>Pohledem vyhodnoťte jestli jednotlivé skupiny vykazují normální rozdělení (respektive spíše pokud nějaká skupina vykazuje významné odchylky od normálního rozdělení) a zdali je mezi skupinami pro každý z parametrů vidět nějaký efekt (silný/slabý, či z hlediska pacienta pozitivní/negativní, ...).</p>	1
<p>Zjistěte, zda data individuálních skupin splňují podmínky pro korektní použití ANOVA (<i>Analysis Of Variance</i>), tedy:</p> <ol style="list-style-type: none"> 1. Normalitu vstupních dat <ul style="list-style-type: none"> ○ Vyberte vhodný test, viz předešlé cvičení a přednáška. ○ Není třeba vypisovat hypotézy, pouze si dejte pozor – u jednotlivých testů se liší (viz např. rozdíl mezi <code>kstest</code> a <code>swtest</code>). ○ Výsledky testu korektně reportujte. 2. Homoskedasticitu – homogenitu rozptylů <ul style="list-style-type: none"> ○ BUĎ: Korektním způsobem argumentujte, proč lze požadavek na homogenitu rozptylů považovat za splněný i bez provádění testů. <i>Toto odůvodnění zapište slovy.</i> ○ NEBO: Otestujte rovnost rozptylů pomocí vhodné analýzy – Bartlettova nebo Levenova testu. Výsledek testu korektně reportujte a interpretujte. Náповěda: funkce <code>vartestn</code>. 3. Nezávislost vstupních dat <ul style="list-style-type: none"> ○ <u>Splněno</u> již díky správně navrženému experimentu, kdy nejsou mezi vzorky žádní duplicitní pacienti. <p>Vyšetření normality bude teoreticky vyžadovat $3 \cdot 4 = 12$ testů, vyšetření homogenity rozptylů (pokud ho budete vyšetřovat statistickými testy) pak 4 testy. Doporučujeme tedy zpracovat data pomocí cyklu, který bude automaticky vypisovat reporty.</p> <p>Vhodnost použití parametrických testů (v další části) rozhodujete na základě toho, kolik skupin z celku se vymyká normalitě, jak moc jsou „nenormální“ (vzdálené od normality, viz výsledky testů), popř. jak moc je narušena homoskedasticita. Rozhodnout se můžete jakkoliv, důležité je pak následné výsledky správně interpretovat.</p>	1.5

<p>Chtěli bychom dokázat, že lze rozlišit zdravé lidi (HC), pacienty s PD a pacienty s HD jen na základě nahrávek řečových testů, a případně i potvrdit, že změny v akustických parametrech úzce souvisí s celkovým klinickým hodnocením stavu pacienta.</p> <p>Použijte informace, které jste získali v předchozích částech, pro výběr vhodných testů a následnou interpretaci. Odpovězte na následující otázky:</p> <ol style="list-style-type: none"> 1. <i>Lze mezi sebou statisticky významně rozlišit skupiny HC, PD a HD pomocí parametrů RI, MPT, DVA nebo stdF0? V této části hledáte jen odpověď ano či ne pomocí vhodného testu.</i> 2. <i>Pokud jste zjistili, že pomocí nějakého parametru lze od sebe skupiny odlišit, jaké dvojice skupin vykazují mezi sebou statisticky významné rozdíly?</i> 3. <i>Jaký je vztah mezi zrychlováním řeči (parametr RI) a klinickým hodnocením lékařem (hodnoty z dotazníků UPDRS a UHDRS)? Odpovídá horší schopnost udržet tempo řeči vyššímu motorickému postižení, hodnocenému lékařem?</i> <ul style="list-style-type: none"> • Výběr statistických testů odůvodněte. • Pro opakované testy stačí zapsat testové hypotézy jen jednou. • Reportujte výsledky testů korektně (využijte přednášku). • Výsledky testů interpretujte s ohledem na předešlá rozhodnutí o normalitě, homoskedasticitě, výběru testu a výběru post-hoc metody. 	1.5
<p>Nepovinný bonus:</p> <p>Manuálně implementujte jednosměrnou (<i>one-way</i>) ANOVA analýzu, a tu následně otestujte na vámi vygenerovaných datech (dvě přibližně normální rozdělení s vybraným efektem skupin).</p>	0.5

Reference

Boersma, P. and Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.30, retrieved 22 July 2017 from <http://www.praat.org/>

Rusz, J., Saft, C., Schlegel, U., Hoffman, R., and Skodda, S. (2014). Phonatory dysfunction as a preclinical symptom of Huntington disease. PLoS one, 9, e113412, <https://doi.org/10.1371/journal.pone.0113412>

Rusz, J., Hlavnička, J., Čmejla, R., and Růžička, E. (2015). Automatic evaluation of speech rhythm instability and acceleration in dysarthrias associated with basal ganglia dysfunction. Frontiers in bioengineering and biotechnology, 3: 104, <https://doi.org/10.3389/fbioe.2015.00104>