

Katedra aplikované matematiky si Vás v rámci [Statistické laboratoře](#) dovoluje pozvat na sérii workshopů pod názvem

Zimní škola statistické analýzy dat se softwarem R

Lektorky:

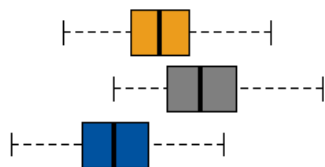
Ing. Martina Litschmannová, Ph.D.

Mgr. Adéla Vrtková

**Workshopy budou probíhat
od 22. ledna 2018 do 26. ledna 2018,
vždy od 8:30h do 11:30h, učebna EB 413 (FEI)**

Termín	Téma
22. ledna	Máme data – a co dál? (popisná statistika)
23. ledna	Neurčitost v datech – bodové a intervalové odhady
24. ledna	Testování statistických hypotéz (princip, vybrané jednovýběrové a dvouvýběrové testy)
25. ledna	Vícevýběrové testy par. hypotéz – jednoduchá analýza rozptylu, Úvod do analýzy kategoriálních dat – dvourozměrné kontingenční tabulky
26. ledna	Úvod do korelační a regresní analýzy

Kapacita zimní školy je omezená. Počítáme s 25 místy pro zájemce z řad zaměstnanců a studentů VŠB – TU Ostrava. Přihlásit se můžete na <http://k470.vsb.cz/statlab/>.



Anotace jednotlivých workshopů:

1. Máme data – a co dál? (popisná statistika)

Statistika nuda je... Nebo není? Seznámíme se se základními pojmy z popisné statistiky a následně se zaměříme na vizualizaci a charakteristiky kategoriálních (nominálních a ordinálních) dat. Seznámíme se se softwarem R a ukážeme si, jak v Rku načíst datový soubor, jak data připravit pro následnou analýzu a jak analyzovat jednorozměrná kategoriální data.

Dále budeme pokračovat s popisnou statistikou numerických (poměrových) dat. V teoretické části lekce se zaměříme zejména na správnou interpretaci měř polohy a heterogenity dat a způsoby vizualizace. Následně vše procvičíme na praktických příkladech v Rku.

2. Neurčitost v datech – bodové a intervalové odhady

Pomocí popisné statistiky zjišťujeme informace o našich datech (výběrových souborech). Statistika nám však umožňuje více. Na základě výběrového souboru dokážeme odhadnout (s jistou spolehlivostí), jak se chová příslušná populace. Ukázat jak na to je cílem tohoto workshopu.

3. Testování statistických hypotéz (princip, vybrané jednovýběrové testy, vybrané dvouvýběrové testy)

Pomocí intervalových odhadů odhadujeme na základě náhodných výběrů populační charakteristiky. Jiným způsobem, jak zjistit na základě náhodného výběru „něco“ o populaci, je testování hypotéz. Vysvětlíme si princip testování hypotéz, seznámíme se s pojmem p-hodnota a s typy chyb, které se při testování hypotéz mohou vyskytnout. Na vybraných jednovýběrových a dvouvýběrových testech si ukážeme, jak se testování provádí v praxi (samozřejmě s využitím Rka).

4. Vícevýběrové testy par. hypotéz – jednoduchá analýza rozptylu

Workshop je pokračováním v tématu Testování hypotéz. Zaměříme se na testy shody úrovně (stř. hodnot, mediánů) a ověření homogenity vícevýběrových dat, tj. ANOVu, Kruskalův – Wallisův test, Bartlettův a Leveneův test.

Druhá část workshopu bude věnována základní analýze závislosti dvou kategoriálních proměnných. Ukážeme si, jak s pomocí Rka provést základní explorační analýzu, seznámíme se s vhodnými měrami asociace, ukážeme si, jak sledovanou závislost vizualizovat a jak testovat její statistickou významnost. Pokud nám to čas dovolí, ukážeme si i speciální přístupy k analýze dat zapsaných v asociačních tabulkách.

5. Úvod do korelační a regresní analýzy

Závěrečný workshop je věnován základní analýze závislosti numerických proměnných. Ukážeme si, jak měřit sílu závislosti mezi numerickými proměnnými a podrobněji se pak budeme zabývat základy lineární regrese.