

Základy matematiky a statistiky pro humanitní obory II

Vojtěch Kovář

Fakulta informatiky, Masarykova univerzita
Botanická 68a, 602 00 Brno, Czech Republic

`xkovar3@fi.muni.cz`

část 1

Obsah přednášky

- 1** Motivace (opakování)
- 2** Informace o předmětu
- 3** Shrnutí Základů matematiky I
- 4** Základy matematiky II

Rozdíl mezi SŠ a VŠ matematikou

■ Středoškolská matematika

- = počty s čísly:
- → kolik budu platit v obchodě (sčítání)
- → jaké daně budu mít (zlomky, procenta)
- → k čemu to ***** je? (matice, integrály)

■ Vysokoškolská matematika

- = umění abstrakce + přemýšlení v obecnostech
- → zásobárna abstraktních pojmů
- → přesné definice
- → spolehlivé vyvozování závěrů (důkazy)
- **společný základ** pro všechny technické obory
- **statistika** – silný nástroj společný všem oborům

Proč potřebují lingvisté matematiku?

■ Počítačová lingvistika

- potřeba spolupracovat s technicky zaměřenými lidmi
- → pochopit jejich způsob myšlení
- počítačové modely jazyka jsou založeny na matematických faktech

■ Abstraktní myšlení

- schopnost rozumově uchopit složité pojmy
- schopnost zobecňovat
- schopnost rozkládat složité problémy na jednodušší
- → nejsou tak důležité vědomosti samotné jako dovednosti, kterým se při jejich vstřebávání naučíte

■ Konkrétní praktické nástroje k analýze dat

Informace o předmětu

■ Obsah předmětu

- vybrané kapitoly z diskrétní matematiky
- zejména statistika a pravděpodobnost

■ Ukončení předmětu

- zkouška (formou dvou písemek)
- 25 bodů vnitrosemestrální písemka 9.4.
- 75 bodů závěrečná písemka

■ Úspěšné ukončení

- min. 60 % bodů z písemek

Co už umíme z minulého semestru

- Může se objevit i na zkoušce v tomto semestru
- Matematická logika
 - čtení a zápis výrokových a predikátových formulí
 - převod do „běžného“ jazyka
 - vyhodnocování pravdivosti formulí
- Teorie množin
 - základní množinové zápisy a operace
 - relace a funkce a jejich vlastnosti
 - čísla

Co čekat v tomto semestru

- Vybrané kapitoly z diskrétní matematiky
- Statistika a pravděpodobnost
 - průměr, modus, medián, korelace
 - kombinatorika, pravděpodobnostní prostor, entropie
 - vyhodnocování úspěšnosti, statistická významnost
 - precision, recall, p-value
 - cca 7 přednášek
- Základy lineární algebry
 - vektory, matice (\rightarrow word embeddings)
 - cca 1 přednáška

Co čekat v tomto semestru

- Něco málo z diferenciálního počtu
 - primitivní funkce, derivace
 - → trénování neuronových sítí
 - cca 1 přednáška
- Teorie grafů
 - graf, vrcholy, hrany, cesty v grafu
 - některé grafové algoritmy
 - cca 1 přednáška
- Něco ze zpracování přirozeného jazyka
 - korpusové statistiky, jazykové modely, word embeddings, LLM
 - cca 2 přednášky
- I další oblasti na základě vašich podnětů

Studijní materiály

- Studijní materiály a další informace
 - studijní text
 - předtočené [přednášky na YouTube](#) (7-13)
 - příklady řešené během on-line výuky na diskusních fórech: [jaro 2021](#), [jaro 2020](#)
 - slidy, diskuze, konzultace
- Vnitrosemestrální písemka
 - 9. 4. v termínu přednášky
- Odpadající přednášky
 - 26. 2.
 - 23. 4. (reading week)