

# Longitudinální a panelová data – NMST 422

Akademický rok 2023/2024 | Letný semester

‡ <b>Prenášky/cvičenia:</b>	Doc. RNDr. Matúš Maciak, Ph.D. <a href="http://www.karlin.mff.cuni.cz/~maciak">http://www.karlin.mff.cuni.cz/~maciak</a>
☉ <b>Prednáška:</b>	Po 12:20 – 13:50   @K2
☉ <b>Cvičenie:</b>	Po 14:00 – 15:30   @K11
@ <b>Konzultačné hodiny:</b>	St: 09:00 – 10:30    @K151 (1. patro) (prípadne podľa individuálnej dohody emalom/osobne)
☰ <b>Web stránka predmetu:</b>	<a href="http://www.karlin.mff.cuni.cz/~maciak/nmst422_2324.php">http://www.karlin.mff.cuni.cz/~maciak/nmst422_2324.php</a>

## 📖 Základné informácie

Predmet *Longitudinální a panelová data (NMST 422)* predpokladá základné znalosti z oblasti teórie pravdepodobnosti, matematickej štatistiky a problematiky lineárneho regresného modelu (v rozsahu preberanom na predmete *Linear Regression – NMSA 407*). Študenti si môžu zapísať predmet NMST422 ak už absolvovali predmet NMSA407 (prerekvizita).

Výuka predmetu NMST422 je organizovaná v rozsahu 2/2 (t.j., 1-krát týždenne prednáška a 1-krát týždenne cvičenie k prednáške). Cieľom prednášky je oboznámiť študentov so základnými stochastickými modelmi a štatistickými postupmi pre zhlukové a časovo/priestorovo korelované (závislé) data.

Základnou náplňou cvičenia, ktoré naväzuje a prakticky dopĺňa prednášku, je riešenie rôznych teoretických a praktických úloh a ilustrácia jednotlivých metód, postupov a modelov na reálnych datach. Z aplikačného hľadiska bude pre účely cvičenia používaný štatistický software SAS—konkrétne jeho online cloudová služba SAS OnDemand (ďalšie podrobnosti budú upresnené na prvom cvičení).

Podmienkou pre prihlásenie sa na skúšku je získanie zápočtu. Podmienkou pre získanie zápočtu je **povinná účasť na cvičení** (dovolené sú maximálne tri neospravedlnené absencie v priebehu semestra) a **včasne odovzdanie dostatočne kvalitne vypracovaného záverečného projektu** (zadaného v priebehu semestra). Od každého študenta je vyžadovaná **individuálna príprava**, samostatné precvičovanie príkladov a úloh preberaných na cvičení a tiež aktívna účasť na celkovej výuke. Aktuálne informácie k cvičeniu sú k dispozícii na webovej stránke cvičenia.

## 📖 Orientačný sylabus cvičenia:

- Úvod a základy štatistického programu SAS;
- Lineárny regresný model a zobecný lineárny regresný model;
- Mnohorozmerná lineárna regresia;
- Longitudinálne data, časovo závislé profily;
- Lineárny model s náhodnými efektami;
- GLM model, GEE postupy;
- GLM models náhodnými efektami;
- Data a modely s chýbajúcimi pozorovaniami;
- Bayesovské prístupy;

Aktuálne informácie a konkrétny harmonogram prednášky spolu s učebnými materiálmi (PDF slidy k prednáške a podkladové markdowny k jednotlivé cvičenia) budú v priebehu semestra zverejňované na webovej stránke cvičenia.

### ☑ **Hodnotenie (podmienky udelenia zápočtu a priebeh zkoušky):**

Na cvičení je vyžadovaná **prezenčná účasť**. Pre účely získania zápočtu (pripustenia ku skúške) musí študent/študentka splniť všetky nasledujúce podmienky:

#### 1. **Zápis predmetu v systéme SIS.**

#### 2. **Účasť na cvičení**

Dovolené sú maximálne tri neospravedlnené absencie (z celkového počtu 13 cvičení) v priebehu semestra. Absenciu na cvičení je možné nahradiť individuálnou aktivitou vo forme vyriešenia dodatočných príkladov, ale vždy explicitne a vopred konzultovaných s cvičiacim.

#### 3. **Záverečný projekt**

V priebehu semestra bude zadaný samostatný (individuálny) záverečný projekt, ktorý sa bude vzťahovať ku konkrétnemu datovému súboru. Pre účely získania zápočtu musí študent/študentka (dostatočne kvalitne) vypracovať projekt podľa zadania (t.j., analyzovať data pomocou niektorej z metód diskutovaných na prednáške a na cvičení) a svoje riešenie odovzdať v rámci stanoveného deadlinu.

#### 4. **Záverečná skúška**

Záverečná skúška bude prebiehať kombinovanou formou (písomná časť + ústna skúška). Písomná časť skúšky bude obsahovať základné otázky k látke preberanej na prednáške. Ústna časť (ktorá sa bude konať len v prípade úspešnej písomnej časti, t.j. **aspoň 60 % bodov**) bude pozostávať s diskusie nad riešením písomnej časti, resp. riešením záverečného (samostatného) projektu (viď bod 3).

### 📖 **Povinná a doporučená literatúra:**

- ❑ Diggle, P., Diggle, P. J., Heagerty, P., Liang, K. Y., and Zeger, S. (2002). Analysis of longitudinal data. Oxford university press.
- ❑ Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2012). Applied longitudinal analysis. John Wiley & Sons.
- ❑ Hardin, J.W. and Hilbe, J.M. (2007). Generalized Linear Model and Extensions. StataPress.
- ❑ Pinheiro, J. and Bates, D. (2006). Mixed-effects models in S and S-PLUS. Springer science & business media.
- ❑ Kulich, M. (2022). NMST432 Advanced Regression Models: Extended Course Notes. Charles University [18.02.2022]

Ďalšie doplňujúce materiály, prípadne doplňujúca literatúra budú dodatočne uvedené v podkladových materiáloch (napr. PDF slidy k prednáške, poznámky k prednáške, alebo HTML markdowny pre SAS) pre prednášku, resp. cvičenie.

### ► **Disclaimer:**

V rámci platných Pravidiel pro organizaci studia na Matematicko-fyzikální fakultě Univerzity Karlovy (ze dne 23.června, 2023), sa vzhľadom k Čl. 8, ods.2 týchto pravidiel týmto vyhlasuje, že povaha predmetu vylučuje právo študenta na jeden riadny a dva opravné termíny pro získaní zápočtu. Získaní zápočtu sa riadi výhradne pravidlami uvedenými vyššie.

### ✉ **Kontakt:**

Webová stránka: <http://www.karlin.mff.cuni.cz/~maciak>

E-mail: [maciak@karlin.mff.cuni.cz](mailto:maciak@karlin.mff.cuni.cz)

Praha, 13. února 2024