

Požadavky k SZZ – specializace Statistika a analýza dat

Státní závěrečná zkouška sestává z obhajoby diplomové práce a z ústní zkoušky.

Charakteristika závěrečné práce a její obhajoba

Zpracováním diplomové práce student prokazuje orientaci v problematice dané tématem práce a schopnost odborné práce pod vedením vedoucího. U obhajoby diplomové práce se hodnotí porozumění tématu a úroveň prezentace.

Charakteristika ústní zkoušky

Účelem zkoušky je prověřit, že absolvent je schopen vést debatu na odborné úrovni. Cílem ústní zkoušky není opakovat zkoušky z jednotlivých předmětů a zkoušet detailní znalost teorie a důkazů. Smyslem je prokázat všeobecný přehled o základních pojmech a výsledcích z jednotlivých oborů a širších souvislostech mezi nimi a o jejich možných aplikacích.

Technická realizace

U ústní zkoušky student obdrží tři otázky, jednu z okruhu A společných oblastí znalostí programu Aplikovaná matematika a dvě ze znalostí své specializace, které jsou uvedeny v okruhu B.

Vymezení rozsahu otázek k ústní zkoušce

A. Společný okruh – Základy matematiky

1. Základy časových řad

vlastnosti a charakteristiky náhodných posloupností a časových řad, odhady charakteristik stacionárních časových řad a modelování deterministických složek (regrese, vyhlazování a dekompozice)

2. ARMA modely

vlastnosti ARMA modelů, korelační struktura ARMA procesů, predikce a odhad parametrů v ARMA modelech, rozšíření pro sezonní řady a nestacionární řady s jednotkovými kořeny (SARIMA modely)

3. Stochastická analýza

Wienerův proces a jeho vlastnosti, stochastický integrál, Itôovo lemma, řešení stochastických diferenciálních rovnic, martingaly, Girsanovova věta

4. Stochastické modely

modelování pomocí stochastických diferenciálních rovnic, Wienerův proces s driftem, geometrický Brownův pohyb, Ornsteinův-Uhlenbeckův proces, difuze

5. Maticové numerické metody

blokové operace s maticemi, rozklady matic a jejich použití, výpočet vlastních hodnot a vlastních vektorů; metoda nejmenších čtverců – klasický přístup a přístup pomocí pseudoinverze

6. Optimalizační numerické metody

Newtonova-Raphsonova metoda, Fisherova skóringová metoda, Nelderova-Meadova metoda, metoda bisekce, metoda zlatého řezu, Brentova-Dekkerova metoda; metoda nejmenších čtverců – obyčejná, pomocí pseudoinverze, nelineární

B. Okruh specializace Statistika a analýza dat

1. Parametrická statistická inference 1

funkce věrohodnosti, jejich logaritmy a typy (relativní, profilová a odhadnutá funkce věrohodnosti); různé aproximace funkcí věrohodnosti, delta metoda; bodové a intervalové odhady parametrům

2. Parametrická statistická inference 2

testování statistických hypotéz Waldovým principem, věrohodnostním poměrem a skóre principem pro diskrétní (kategoriální) a spojitá data; vybrané parametrické statistické testy pro parametry, jejich funkce, vektory a podmnožiny parametrů

3. Neparametrická statistická inference 1

pořadové statistiky, hustota a distribuční funkce pořadové statistiky, minima a maxima; asymptotické rozdělení pořadové statistiky, rozptyl pořadové statistiky, střední hodnota a rozptyl mediánu, pás spolehlivosti pro distribuční funkci

4. Neparametrická statistická inference 2

hustota, distribuční funkce, funkce přežití, riziko a kumulativní riziko, střední hodnota času přežití a střední hodnota zůstatkového života; neparametrické odhady; neparametrické testy pro (ne)cenzorovaná data – dva a vícevýběrové testy funkcí přežití a rizika; zobecněný Spearmanův a Kendallův koeficient korelace

5. Jádrové vyhlazování

obecný princip jádrových odhad, jádrové odhady hustoty a distribuční funkce, kanonická jádra a teorie optimálních jader, jádra vyšších řádů; jádrové odhady regresní funkce, jádrové odhady dvourozměrných hustot, kritéria pro posouzení kvality odhadu

6. **Splajnové vyhlazování**

interpolace a vyhlazování křivek a ploch pomocí jednorozměrných a mnohorozměrných splajnů; aditivní modely; zobecnění interpolace a vyhlazování na náhodný výběr křivek a ploch – registrace, resampling; analýza funkcionálních dat

7. **Regresní modely 1**

lineární regresní modely (LRM) s homogenními a nehomogenními rozptyly, LRM s fixními efekty a korelovanými chybami, LRM se smíšenými a náhodnými efekty

8. **Regresní modely 2**

zobecněný lineární model (exponenciální rodiny rozdělení, linková funkce, kanonická linková funkce), důležité zobecněné lineární modely (logistická regrese, poissonovská regrese, multinomická regrese, logaritmicko-lineární modely, kontingenční tabulky)

9. **Regresní modely 3**

rozdělení pravděpodobnosti ze třídy zobecněného gama a související rozdělení; parametrické regresní modely v analýze přežití a v analýze historie událostí, Coxův regresní model

10. **Regresní modely 4**

koncepty optimality v navrhování experimentů, hřebenová regrese, LASSO, regrese s hlavními komponentami, nelineární regresní model

11. **Mnohorozměrné statistické metody**

testování mnohorozměrných hypotéz o vektorech středních hodnot a kovariančních maticích, MANOVA, profilová analýza, mnohorozměrné lineární regresní modely, PCA, PLS, faktorová analýza, analýza kanonických korelací, diskriminační analýza

12. **Stochastické modely markovského typu**

homogenní markovský řetězec se spojitým časem; proces vzniku a zániku a jeho speciální případy; teorie hromadné obsluhy – struktura systému hromadné obsluhy, Kendallova klasifikace, odvození charakteristik jednolinkového stabilizovaného systému