



HÖGSKOLAN
I SKÖVDE

KURSPLAN

Statistisk analys och datavisualisering, Forskarnivå 6 högskolepoäng

Kurskod: IT0945F

Revisionsnummer: 3.1

Gäller från: 2024-07-01

Beslutad av: Utbildningskommittén för utbildning på forskarnivå

Beslutsdatum: 2024-03-11

1. Allmänt om kursen

Kursen ges av Högskolan i Skövde och benämns Statistisk analys och datavisualisering, Forskarnivå (Statistical Analysis and Data Visualization, Post-graduate level). Omfattningen är 6 högskolepoäng (hp).

Kursen ingår i forskarutbildningsämnet informationsteknologi.

2. Behörighetskrav

För tillträde till kursen krävs grundläggande behörighet för utbildning på forskarnivå, d.v.s. avlagd examen på avancerad nivå eller fullgjorda kursfordringar om minst 240 högskolepoäng, varav minst 60 högskolepoäng på avancerad nivå (eller motsvarande).

För att uppfylla kravet på särskild behörighet krävs att den sökande har fullgjort kursfordringar om minst 60 högskolepoäng, inklusive självständigt uppsatsarbete omfattande minst 15 högskolepoäng på avancerad nivå, inom ämnet informationsteknologi, närliggande tillämpningsområden eller andra ämnesområden som bedöms som direkt relevanta för avhandlingsarbete i informationsteknologi.

Vidare krävs godkänt betyg i Engelska 6 (eller motsvarande kunskaper). Motsvarande kunskaper visas normalt genom ett internationellt erkänt språktest, till exempel IELTS eller TOEFL.

3. Innehåll

Kursen behandlar viktiga statistiska metoder inom deskriptiv (beskrivande) och statistisk inferens, samt explorativ (undersökande) analys genom multivariat datavisualisering. Kursen inleds med en introduktion till statistiska begrepp, vanliga sannolikhetsmodeller och deras tillämpningar. Klassiska univariata och bivariata visualiseringsmetoder beskrivs, följt av multivariata visualiseringsmetoder baserade på 'machine learning', som inkluderar klustring, dimensionsreducerande projektioner och tvådimensionell avbildning.

Kursens avsnitt om statistisk inferens tar upp det frekventistiska angreppssättet, men ger även en kort översikt av Bayesianska metoder. Grunden för statistisk hypotesprövning och dess grund i den vetenskapliga metoden diskuteras. I detta avsnitt förklaras vanligt förekommande experimentell design och samplingsmetoder. Ett stort urval av hypotestester, inklusive anpassningsgrads-, centralmått- och spridningsmåttstester behandlas, för scenarier med ett, två eller flera stickprov. Statistiska regressionsmetoder beskrivs också i statistisk inferens.

Kursen avslutas med en kritisk analys av statistiska analysmetoder som använts i vetenskapliga artiklar från olika forskningsområden och diskussioner om hur felaktig experimentell design, felaktig tolkning av statistiska tester och oetisk statistisk rapportering kan påverka debatten negativt inom vetenskapssamhället och samhället i övrigt.

Kursen använder statistiska verktygslådor och bibliotek i ett eller flera vetenskapliga datorprogram och programmeringsmiljöer. I de praktiska övningarna används data från verkligheten, varav några väljs för att främja diskussioner om samhällsfrågor såsom klimatförändringar, ekonomisk ojämlikhet, ojämlikhet mellan könen och diskriminering.

4. Mål

Efter avslutad kurs ska doktoranden kunna:

- beskriva och skilja mellan olika metoder inom deskriptiv (beskrivande) och statistisk inferens,
- demonstrera användningen av multivariata visualiseringsmetoder för explorativ (undersökande) dataanalys,
- formulera statistiska hypoteser och testa dessa med hjälp av statistisk programvara,
- utvärdera och motivera användningen av specifika statistik- och visualiseringsmetoder för vissa dataanalysuppgifter,
- analysera och bedöma statistiska analysmetoder som använts i ett urval av vetenskapliga artiklar från olika forskningsområden, och
- argumentera för hur experimentell design och statistisk rapportering kan påverka debatten inom vetenskapssamhället och samhället i övrigt.

5. Examination

Kursen bedöms med betygen G (Godkänd) eller U (Underkänd).

För att få betyget Godkänd på hel kurs krävs att alla examinationsmoment är godkända.

Kursen har följande examinationsmoment:

- **Laborationer**
2 hp, betyg: G/U
- **Seminarium**
1 hp, betyg: G/U
- **Projekt rapport**
3 hp, betyg: G/U (bestämmer betyg på hel kurs)

Doktorander med varaktig funktionsnedsättning som har fått beslut om riktat pedagogiskt stöd kan erbjudas anpassad eller alternativ examination.

6. Undervisningsformer och undervisningsspråk

Undervisningen består av föreläsningar, laborationer, seminarier och projektarbeten.

Undervisningen bedrivs på engelska.

7. Kurslitteratur och övriga läromedel

Referenslitteratur

Chen, C., Härdle, W. K., Unwin, A. (2008). *Handbook of Data Visualization*. Berlin Heidelberg: Springer-Verlag. ISBN 9783662500743.

Efron, B., & Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge, UK: Cambridge University Press.

Field, A. (2017). *Discovering statistics using IBM SPSS Statistics* (5th ed.). London: SAGE. ISBN 9781526419521.

Nussbaumer Knaflic, C. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Hoboken, New Jersey: John Wiley & Sons. ISBN 978-1-119-00225-3.

Peck, R., Olsen, C. & Devore, J. (2019). *Introduction to Statistics and Data Analysis* (6th ed.). Cengage Learning. ISBN 9781337793612.

Wasserman, L. (2004). *All of Statistics*. New York, NY: Springer. ISBN 9781441923226.

Weiss, N. A. (2016). *Introductory Statistics*. Pearson. ISBN 9780321989178.

Vetenskapliga artiklar enligt lärarnas anvisningar.

8. Doktorandinflytande

Doktorandinflytande i kursen säkerställs genom kursvärdering. Doktoranderna informeras om resultatet av kursvärderingen och eventuella åtgärder som genomförts eller planeras, grundat på kursvärderingen.

9. Övrigt

På Högskolan i Skövdes webbplats finns ytterligare information om kursen samt nationella och lokala styrdokument för högskoleutbildning.