



Statistiek III: regressieanalyse (C003560)

Cursusomvang (nominale waarden; effectieve waarden kunnen verschillen per opleiding)

Studiepunten 6.0 Studietijd 165 u Contacturen 45.0 u

Aanbodssessies en werkvormen in academiejaar 2019-2020

A (semester 1)	Engels, Nederlands	werkcollege: PC- klasoefeningen	10.0 u
		practicum	5.0 u
		hoorcollege	30.0 u

Lesgevers in academiejaar 2019-2020

Vansteelandt, Stijn	WE02	Verantwoordelijk lesgever
---------------------	------	---------------------------

Aangeboden in onderstaande opleidingen in 2019-2020

	stptn	aanbodssessie
Bachelor of Science in de wiskunde	6	A
Vorbereidingsprogramma tot Master of Science in de wiskunde	6	A

Onderwijstalen

Nederlands, Engels

Trefwoorden

Multivariate lineaire modellen, statistische data-analyse

Situering

De student een basis geven voor statistische modelbouw. Haar/hem toelaten om voor een waaier aan praktische problemen via een relatief eenvoudig (multivariaat lineair) statistisch model een oplossing te vinden. Met behulp van software het lineaire model kunnen aanpassen aan een dataset en er de gepaste conclusies bij kunnen trekken. In de context van het lineaire model de basisbegrippen van design, analyse, confounding, interactie, informatie, efficiëntie, multi-collineariteit, significantie, kracht, goodness-of-fit, invloedrijke observaties, robuustheid en modelbouw bijbrengen. De theoretische en praktische eigenschappen kunnen afleiden van de schatters en predicties die voortvloeien uit het multivariaat lineair model.

Inhoud

- De nood aan modelleren
- Het enkelvoudige en meervoudige lineaire model
- Regression to the mean
- Associatie, predictie versus causatie
- Het kleinste kwadratenprincipe
- Maximum likelihood theorie voor lineaire modellen.
- Betrouwbaarheidsintervallen voor conditionele gemiddelden
- Sandwich schatters
- Residuals en goodness-of-fit
- Grafische en formele diagnostische methoden voor de inspectie van residuen
- Goodness-of-fit testen en controle van gemaakte vooronderstellingen
- Predictie-intervallen voor nieuwe metingen
- De impact van het transformeren van variabelen
- Meervoudige correlatiecoëfficiënt
- Multicollineariteit
- Het afwegen van vertekening en efficiëntie
- Best subset selectie, ridge regression, lasso
- Counterfactuals en causale effecten
- Corrigeren voor covariaten in gerandomiseerde studies
- De impact van model misspecificatie
- Procedures voor de stapsgewijze opbouw van een regressiemodel

- Post-selection inference
- Detectie van invloedrijke metingen via leverage en de Cook's afstand
- Het klassieke en het Berkson error model voor de predictor
- Het verband tussen lineaire regressie en variantie-analyse
- Confounding en effectmodificatie
- Splines

Begincompetenties

Eindcompetenties van het vak Statistiek II.

Eindcompetenties

- 1 De student moet in staat zijn praktische problemen te herkennen die kunnen opgelost worden met behulp van een (multivariaat) lineair model.
- 2 Hij/zij moet het onderscheid kunnen maken tussen associatie, predictie en causatie en het begrip confounding in die context kunnen plaatsen.
- 3 Hij/zij moet het lineaire model op correcte wijze kunnen interpreteren, het kunnen aanpassen aan een gegeven dataset en er verantwoorde conclusies kunnen uit trekken, zowel in formele als in toegepaste zin.
- 4 Hij/zij moet de eigenschappen van de schatters kennen en kunnen afleiden, met aandacht voor het onderscheid tussen een betrouwbaarheidsinterval en een predictie-interval.
- 5 Hij/zij moet het verband kennen tussen lineaire regressie en variantie-analyse.
- 6 Hij/zij moet ook de beperkingen van het lineaire model onderkennen, in een concrete context de nodige uitbreidingen kunnen suggereren en die met de maximum likelihood methode kunnen uitwerken.
- 7 Hij/zij moet een goede balans tussen efficiëntie en vertekening kunnen nastreven bij het ontwerpen van een statistische studie.
- 8 Hij/zij moet de methode en resultaten van een lineaire regressie analyse zowel schriftelijk als mondeling correct en helder kunnen rapporteren. Hij/zij moet in deze context ook in team kunnen werken aan een oplossing voor een praktisch probleem.
- 9 Hij/zij moet regressie-analyse met hoog-dimensionale predictoren kunnen uitwerken.

Creditcontractvoorwaarde

Toelating tot dit opleidingsonderdeel via creditcontract is mogelijk mits gunstige beoordeling van de competenties

Examencontractvoorwaarde

Dit opleidingsonderdeel kan niet via examencontract gevolgd worden

Didactische werkvormen

Hoorcollege, practicum, werkcollege: PC-klasoefeningen

Toelichtingen bij de didactische werkvormen

Oefeningen: werksessies en PC-practica onder begeleiding, uitvoering in team en ook zelfstandig rapporteren van een data-analyse

Leermateriaal

Uitgebreide slides, en materiaal uit tekstboeken:
Wooldridge, J. (2012). Introductory Econometrics: A Modern Approach, 5th Edition.

Referenties

Wooldridge, J. (2012). Introductory Econometrics: A Modern Approach, 5th Edition.

Vakinhoudelijke studiebegeleiding

De studenten worden begeleid door assistenten bij praktische oefeningensessies en tijdens PC practica. Buiten de lessen kunnen zij via de elektronische leeromgeving met elkaar en met de lesgevers vragen en antwoorden uitwisselen. Een project zal hen in staat stellen data-analytische vaardigheden daadwerkelijk in de praktijk te brengen.

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Schriftelijk examen met open vragen, openboekexamen, mondeling examen

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Schriftelijk examen met open vragen, openboekexamen, mondeling examen

Evaluatievormen bij niet-periodegebonden evaluatie

Werkstuk, vaardigheidstest, verslag

Tweede examenkans in geval van niet-periodegebonden evaluatie

Examen in de tweede examenperiode is mogelijk

Toelichtingen bij de evaluatievormen

Periodegebonden evaluatie: de kennis en het probleemoplossend vermogen van de student worden schriftelijk getoetst.

Niet-periodegebonden evaluatie: een project waarin een dataset wordt geanalyseerd.

Eindscoreberekening

Niet-periodegebonden evaluatie (15%) + periodegebonden evaluatie (85%).