

Statistiek II: project (C003559)

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

Studiepunten 6.0 **Studietijd 165 u**

Aanbodsessies en werkvormen in academiejaar 2024-2025

A (semester 2)	Nederlands	Gent	werkcollege peer teaching hoorcollege groepswerk
----------------	------------	------	---

Lesgevers in academiejaar 2024-2025

Luyts, Martial	WE02	Verantwoordelijk lesgever
----------------	------	---------------------------

Aangeboden in onderstaande opleidingen in 2024-2025

	stptn	aanbodssessie
Bachelor of Science in de wiskunde	6	A

Onderwijstalen

Nederlands

Trefwoorden

Data-analyse, Kansrekening, wiskundige statistiek

Situering

De stap zetten van theoretische, axiomatisch opgebouwde kennis naar relevante, correcte en efficiënte data-analyse via het softwarepakket R, en rapportering. Als wiskundige leren participeren in teamverband en in een multi-disciplinaire context zoals men die in de maatschappij en heel wat onderzoekssettings ontmoet. Assumpties en conclusies correct en begrijpelijk leren communiceren aan zowel de meer als de minder gespecialiseerde partners. Kennis maken met het beroepenveld.

Inhoud

Gedurende de cursus, worden drie onderling verbonden projecten voorgesteld. Het derde project over een real-life probleem is het examenproject en tools om het te ontwikkelen worden bestudeerd in de twee eerste projecten. Als inleiding leren de studenten betrouwbaarheidsintervallen en eenvoudige toetsen van hypothesen onderzoeken via data-analyse in het softwarepakket R. De one-sample t-test, gepaarde en ongepaarde t-testen, de Pearson chi-kwadraat test en de Fisher exact test kunnen worden betrokken.

Volgens de inhoud van het derde project, maken de studenten kennis met:

- Simulatie-experimenten voor de evaluatie van de eigenschappen van schatters,
- Principes en technieken voor beschrijvende data-analyse. Principaalcomponenten analyse, Cluster-analyse, self-organizing maps,
- Beslissingsbomen, bootstrap aggregatie, random forests,
- Gebruik van statistiek in de media, industrie en academische wereld,
- Wetenschappelijke integriteit en data confidentialiteit,
- Mondeling en schriftelijk wetenschappelijk rapporteren (in LaTeX):
 - De logische structuur van de tekst: globaal en in detail, inhoudelijk en formeel
 - Regels voor auteurschap en volgorde van auteurs.
 - Bronvermelding (acknowledgments, copyright en plagiaat).
 - Referentie-stijlen (BibTeX).
 - Structuur van een slide, verhaalopbouw en timing.

Tijdens de daaropvolgende lessen brengen de studenten in teamverband een stuk theorie aan dat noodzakelijk is voor het uitwerken van het examenproject. Deze lessen worden door de studenten onder begeleiding en in team uitgewerkt en voor de klas gebracht na feedback door

de titularis. De inhoud van deze lessen is afhankelijk van het voorliggende project. Richtingaangevend zijn basistechnieken voor data-analyse, zoals principes en technieken voor beschrijvende data-analyse, principaalcomponentenanalyse, cluster-analyse, one-way anova, het multipliciteitsprobleem, niet-parametrische testen, studie-design, bootstrap en jackknife. Deze lessen worden verrijkt met een beknopt verslag waarbij een data-analyse, gebruik makend van deze technieken in het softwarepakket R, wordt geïllustreerd. Tijdens de werkcolleges worden deze technieken onder begeleiding ge-oefend met behulp van de software R.

In de loop van het semester wordt het tweede project sequentieel door verschillende teams uitgewerkt. Ter voorbereiding op het examenproject, heeft dit project als doel om de studenten te laten kennismaken en participeren in het onderzoeksproces:

- waarbij een toegepaste probleemstelling uit het empirische onderzoek wordt vertaald naar een wiskundig/statistisch probleem;
- waarbij diverse oplossingsstrategieën worden voorgesteld en hun kwaliteit wordt vergeleken via analytisch werk en simulatie-experimenten;
- waarbij de strategie die optimaal werd bevonden, gebruikt wordt om de beschikbare gegevens te analyseren en aldus een antwoord op de gestelde onderzoeksvraag te bieden;
- waarbij de resultaten schriftelijk en mondeling worden gerapporteerd aan de medestudenten.

Wekelijks wordt een discussie gevoerd rond de voorliggende probleemstelling, waarna 1 team aan de slag gaat om een stuk van de voorliggende vraagstelling op te lossen. Vervolgens worden de bevindingen mondeling gerapporteerd, besproken in de volledige studentengroep, waarna een volgend team aan de slag gaat om een verder stuk van de probleemstelling uit te werken. Richtingaangevend is de analyse van classificatie en regressie problemen, waar de studenten diverse mogelijke statistische analyses vergelijken en evalueren. Gaandeweg wordt de studenten duidelijk gemaakt dat diverse analyses vaak mogelijk zijn, maar de kwaliteit vaak verschilt. Het belang van statistisch design en protocol van een studie wordt in dat licht besproken.

In het derde project, ontwikkelen de studenten interessante/originele oplossingen voor het voorgestelde real-life probleem; ze schrijven een volledig rapport en verdedigen dit tijdens het mondelinge examen. Wiskundig gezien hangt het derde project alleen af van de voorkennis die in de eerste twee projecten wordt ontwikkeld. De studenten moeten echter ook de kenmerken in acht nemen van het domein waarop het derde project wordt toegepast.

Begincompetenties

Eindcompetenties van de vakken Statistiek I en Analyse I.

Eindcompetenties

- 1 De student moet de relevantie van eenvoudige statistische methoden bij praktische problemen kunnen herkennen en er de oplossing voor kunnen uitwerken.
- 2 Zij/hij moet het resultaat van deze werkzaamheden op een heldere en correcte manier zowel mondeling als schriftelijk kunnen rapporteren.

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

Groepswerk, Werkcollege, Hoorcollege, Peer teaching

Toelichtingen bij de didactische werkvormen

Oefeningen, inclusief PC-practica: onder begeleiding; (digitaal) bibliotheekbezoek; projectwerk in team met rapportering en feedback; uitwisseling via de elektronische leeromgeving.

Studiemateriaal

Type: Handboek

Naam: Onderdelen uit tekstboeken'

Richtprijs: Gratis of betaald door opleiding

Optioneel: nee

Bijkomende info: Er wordt gesteund op onderdelen uit tekstboeken die online raadpleegbaar zijn.

Type: Slides

Naam: Slides'

Richtprijs: Gratis of betaald door opleiding

Optioneel: nee

Bijkomende info: Slides zijn beschikbaar op Ufora.

Referenties

T. Hastie, R. Tibshirani en J. Friedman. 'The elements of Statistical Learning' (2de editie); Springer, 2009.

G. James, D. Witten, T. Hastie en R. Tibshirani. 'An Introduction to Statistical Learning with Applications in R'; Springer, 2013.

W. Mendenhall. 'Mathematical Statistics with Applications' (6de editie); Duxbury Press, 2001.

D. Nolan en T. Speed. 'Stat Labs: Mathematical Statistics through applications'; Springer, 2000.

Vakinhoudelijke studiebegeleiding

Hoorcolleges en PC Practica, zelfstandig werk met rapportering en feedback, uitwisselingen via de elektronische leeromgeving.

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Schriftelijke evaluatie open boek

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Schriftelijke evaluatie open boek

Evaluatievormen bij niet-periodegebonden evaluatie

Werkstuk

Tweede examenkans in geval van niet-periodegebonden evaluatie

Examen in de tweede examenperiode is enkel mogelijk in gewijzigde vorm

Toelichtingen bij de evaluatievormen

De kennis en het probleemoplossend vermogen van de studenten wordt getoetst aan de hand van een projectwerk met schriftelijke en mondelinge rapportering.

Eindscoreberekening

Permanente evaluatie (mondeling en schriftelijk rapporteren):

- 1 groepswork (uitwerken van een hoorcollege onder begeleiding) met mondelinge rapportering (15%)
- 1 groepswork (in de loop van het semester uitgewerkt onder begeleiding) met schriftelijke en mondelinge rapportering (15%)
- 1 groepswork (individueel uitgewerkt als examenproject) met schriftelijke rapportering en individuele mondelinge verdediging (20%)

Periodegebonden evaluatie: open boek examen (50%)