



Basisstatistiek voor onderzoekers



UNIVERSITEIT
GENT

POSTACADEMISCHE OPLEIDING
3 OKTOBER 2016 - 12 JUNI 2017

Een goede keuze maken uit het ruime aanbod van statistische methoden voor het analyseren van data is geen eenvoudige opdracht. Het berust immers op een grondige kennis van de voorwaarden waaronder de statistische methode werd ontwikkeld. Ook bij de interpretatie van de resultaten is inzicht in de kracht en de tekortkomingen van de gebruikte statistische techniek onontbeerlijk. Om tot een geldige statistische analyse en een verantwoord besluit te komen, dient men ook aandacht te besteden aan de opzet van de studie (Design of Experiment). Op die manier kan men met een minimum aan experimenten tot een maximale kennis over het onderzoek komen.

Deze lessenreeks gaat op deze verschillende fasen in en tijdens de begeleide oefeningen zullen de cursisten zelf gegevens kunnen analyseren en interpreteren. Er wordt gebruik gemaakt van het statistisch softwarepakket R. De voorbije jaren is R uitgegroeid van een veelgebruikte tool voor wetenschappelijk onderzoek tot een standaard in de industrie. Zeker nu Revolution Analytics door Microsoft werd opgekocht, heeft R ook de zakenwereld veroverd. R wordt veelvuldig gebruikt voor analyse van data (Microsoft, Google, Facebook, ...), marktanalyses (ANZ Bank, Bank of America, ...), visualisatie van data (New Scientist, NY Times) en zoveel meer. R is bij uitstek geschikt voor het opstellen van unieke oplossingen en automatisatie van de analyse processen in onderzoek en de zakenwereld. R leren vraagt -zeker in het begin- enige inspanning, maar daarvoor krijg je een goed gestructureerde en gedocumenteerde tool met een onwaarschijnlijk grote ondersteunende gebruikersgemeenschap.

De deelnemers zullen tijdens de cursus over de vereiste software licenties met handleiding beschikken. Iedere module wordt afgesloten met een lesavond waarin alleen oefeningen aan bod komen.

Doelpubliek

De cursus richt zich niet alleen tot onderzoekers aan Universiteiten en Hogescholen, maar ook tot mensen uit de dienstensector, de overheid en de bedrijfswereld die in ondersteunende disciplines gebruik maken van statistische methoden en modellen.

De cursus is toegankelijk voor iedereen die door zijn basisvorming of door zijn ervaring voldoende met het onderwerp vertrouwd is om met interesse en motivatie de opleiding te volgen. Door het modulaire karakter van de opleiding kan de cursist de modulekeuze aanpassen aan zijn interesseprofiel en specifieke voorkennis. Deze opleiding kan – mits deelname aan het examen – opgenomen worden in de doctoraatsopleiding van de Universiteit Gent.

Voor alle cursussen is een basiskennis R vereist met uitzondering natuurlijk van de basiscursus R zelf.

Wetenschappelijke Coördinatie

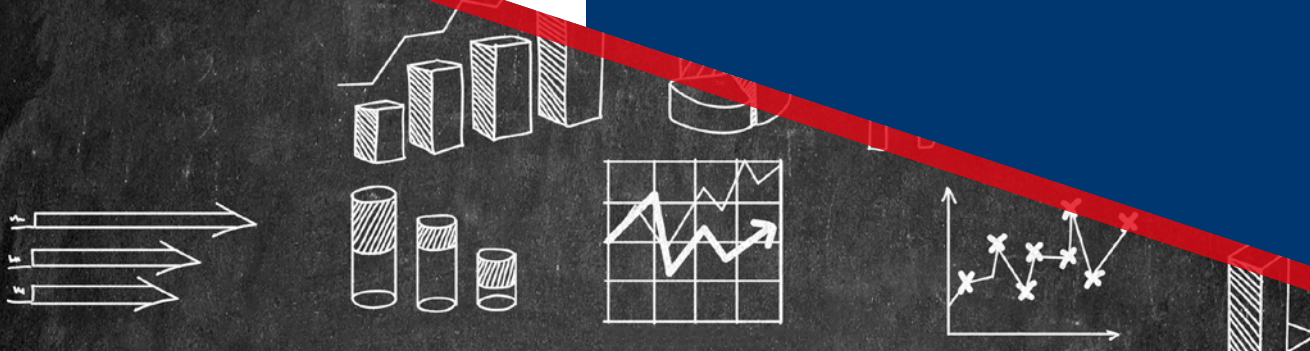
Em. Prof. Jean-Pierre Ottoy, Vakgroep Wiskundige Modelling, Statistiek en Bio-Informatica, Universiteit Gent

Lesgevers

- Prof. Jan De Neve, Vakgroep Data-Analyse, Universiteit Gent
- Dhr. Joris Meys, Vakgroep Wiskundige Modelling, Statistiek en Bio-Informatica, Universiteit Gent
- Em. Prof. Jean-Pierre Ottoy, Vakgroep Wiskundige Modelling, Statistiek en Bio-Informatica, Universiteit Gent
- Mevr. Annelies Tourny, Vakgroep Wiskundige Modelling, Statistiek en Bio-Informatica, Universiteit Gent
- Dr. Karel Vermeulen, Vakgroep Wiskundige Modelling, Statistiek en Bio-Informatica, Universiteit Gent
- Ir. Matthijs Vynck, Vakgroep Wiskundige Modelling, Statistiek en Bio-Informatica, Universiteit Gent

Getuigschrift

U ontvangt een getuigschrift, indien u deelneemt aan minstens 2 van de 4 modules en slaagt in het bijbehorende examen [per module wordt een afzonderlijk examen voorzien].



1. Basiscursus R

De populariteit van het pakket R heeft de voorbije jaren een vlucht genomen. Microsoft, Google en Facebook gebruiken R al langer in hun analyses. Maar ook bedrijven die niks met software te maken hebben, schakelen steeds vaker over naar R. Zo gebruiken bvb ANZ Bank en Bank of America de software voor marktanalyses, en de New York Times gebruikt R voor visualisatie van data.

De cursus begint bij een verkenning van R en de interface RStudio. Daarna komen de verschillende stappen in het werken met data aan bod: inlezen en opslaan van data, het uitvoeren van extra berekeningen, het omvormen van data naar de gewenste structuur voor analyse en het grafisch weergeven van data. Daarnaast ligt de focus op de tools in R die nodig zijn voor het volgen van de andere cursussen in deze module. Doorheen de cursus worden extra tips gegeven om vlotter te werken met R en RStudio.

R heeft de kracht en de mogelijkheden van een programmeertaal, maar daardoor ook een zwaardere leercurve. Daarom worden alle belangrijke concepten in R aan de hand van praktische voorbeelden en oefeningen uitgelegd.

Voorafgaande kennis van programmeren is niet vereist. De cursisten worden wel verwacht vlot met een computer te kunnen werken.

Duur: 4 avonden met 50% demonstratie en 50% oefeningen

Lesgevers: Joris Meys en Annelies Tourny

Datum: 3, 10, 17 en 24 oktober 2016

2. Basiscursus statistiek

Er wordt gestart met een herhaling van de meest voorkomende distributies (normale, t , F , ...), de principes van een statistische toets, de constructie van hypothesen, betrouwbaarheidsintervallen, steekproeven, ...

Vervolgens komen de klassieke statistische technieken aan bod: de t - en de F -toets en de variantieanalyse met 1 of met meerdere factoren.

Nadruk wordt gelegd op het probleemoplossend vermogen van de statistische technieken, de voorwaarden voor toepassing, de interpretatie en de opzet van experimenten. Reële voorbeelden worden uitgewerkt.

Duur: 6 avonden met 50% theorie en 50% oefeningen (R) en 1 extra avond met 100% oefeningen

Lesgevers: Jean-Pierre Otttoy, Annelies Tourny en Karel Vermeulen

Data: 7, 14, 21 en 28 november, 5, 12 en 19 december 2016

3. Regressieanalyse

De regressieanalyse is één van de bekendste en meest gebruikte statistische technieken, zowel in onderzoek als in het bedrijfsleven. Niettegenstaande de techniek redelijk eenvoudig is in gebruik, worden er in de praktijk toch veel problemen ondervonden. Aan welke veronderstellingen moeten de gegevens voldoen? Wat zijn de gevolgen als ze daar niet aan voldoen? Hoe worden betrouwbaarheids- en predictie-intervallen geïnterpreteerd? Hoe wordt het beste regressiemodel met meerdere predictoren geconstrueerd? Kan een discrete variabele als predictor gebruikt worden?

Deze en vele andere problemen worden in deze module uitvoerig behandeld, steeds met de nadruk op de toepassingen. Zowel de enkelvoudige lineaire regressie, de meervoudige lineaire regressie als de logistische en de niet-lineaire regressie komen aan bod.

Duur: 6 avonden met 50% theorie en 50% oefeningen (R) en 1 extra avond met 100% oefeningen

Lesgevers: Jean-Pierre Otttoy, Karel Vermeulen en Matthijs Vynck

Data: 23 en 30 januari, 6, 13 en 20 februari, 6 en 13 maart 2017

4. Proefopzet

Dagelijks worden experimenten uitgevoerd, maar vaak zijn de resultaten beperkt, is het moeilijk om een zinvol en betrouwbaar besluit te vormen en staan de resultaten niet in verhouding tot de tijd, energie en kosten die erin gestoken worden.

Een goede statistische proefopzet, ook bekend als Design Of Experiment (DOE), zorgt ervoor dat de statistische analyse correct interpreteerbaar en betrouwbaar is. Het is noodzakelijk dat de probleemstelling vooraf goed geformuleerd wordt en dat het meest geschikte statistische model gekozen wordt om zo de beoogde informatie te verkrijgen met een minimum aantal experimenten. Een goed ontworpen experiment heeft vaak een substantiële invloed op de efficiëntie van het onderzoek in termen van kosten en precisie.

De cursus is opgebouwd vertrekkende van de theorie van regressie analyse en ANOVA. Niettegenstaande het ontwerp (d.i. het design) van een experiment de statistische analyse voorafgaat, wordt een design steeds opgesteld in functie van deze toekomstige data analyse. Een goede kennis van de lineaire statistische modellen is daarom noodzakelijk.

Het doel van deze module is tweeledig. Enerzijds wordt er inzicht verworven in het opstellen van designs voor verschillende studies/ onderzoeksvragen en anderzijds wordt er dieper ingegaan op de relatie tussen het design en de statistische analyse. Dit laatste laat toe om het paradigma van statistische inferentie beter te begrijpen.

Duur: 6 avonden met 50% theorie en 50% oefeningen (R) en 1 extra avond met 100% oefeningen

Lesgever: Jan De Neve

Data: 24 april, 8, 15, 22 en 29 mei, 6 en 12 juni 2017

MEER INFO & INSCHRIJVEN

www.ivpv.ugent.be/stato

Prijs

Deelnameprijs omvat lesgeld, hand-outs, frisdranken, koffie en broodjes.
 Betaling geschiedt na ontvangst van de factuur. Alle facturen zijn betaalbaar dertig dagen na dagtekening. Alle vermelde bedragen zijn vrij van BTW.
 Voor iedere module kan er afzonderlijk ingeschreven worden.

Module 1: Basiscursus R	€ 600
Module 2: Basiscursus statistiek	€ 1.050
Module 3: Regressieanalyse	€ 1.050
Module 4: Proefopzet	€ 1.050

Korting

- ▶ Bij het volgen van 2 modules ontvangt u 5% korting.
 Bij het volgen van 3 of 4 modules ontvangt u 10% korting.
- ▶ Indien minstens één deelnemer van een bedrijf inschrijft voor de volledige opleiding, wordt voor alle bijkomende gelijktijdige inschrijvingen van hetzelfde bedrijf een korting van 20% verleend. Facturatie geschiedt dan d.m.v. een gezamenlijke factuur.
- ▶ 10% korting op de in de tabel vermelde prijzen voor leden AIG en VBIG.
- ▶ Aangepaste prijzen voor personeel van UGent en geassocieerde hogescholen.
- ▶ Kortingen zijn niet cumuleerbaar.

KMO-portefeuille

Universiteit Gent aanvaardt betalingen via de KMO-portefeuille (www.kmo-portefeuille.be; gebruik autorisatiecode DV.O103194).

Annuleren

Raadpleeg onze annulatievoorwaarden op www.ivpv.ugent.be/annulatievoorwaarden

Tijdstip en locatie

- ▶ De lessen worden gegeven van **16u30 tot 20u**, in 2 delen, gescheiden door een broodjesmaaltijd en vinden plaats aan de **Universiteit Gent, Instituut voor Permanente Vorming, IVPV leszaal, Technologiepark 904, 9052 Zwijnaarde**.
- ▶ Data onder voorbehoud van wijzigingen om onvoorziene omstandigheden.

Handboeken

De opleiding wordt ondersteund door de handboeken:

- ▶ R for Dummies van Joris Meys
- ▶ Applied Linear Statistical Models van Neter, Wasserman en Kutner (module 2 en 3)
- ▶ Design and Analysis of Experiments van Montgomery (module 4)
- ▶ Optimal design of experiments: a case study approach van Goos Peter en Bradley Jones (module 4)

Deze handboeken zijn optioneel en zijn niet inbegrepen in de deelnameprijs.

Andere IVPV opleiding
 Statistiek in het bedrijfsleven

MEER INFO & INSCHRIJVEN

www.ivpv.ugent.be/stato



ORGANISATIE

Universiteit Gent – Instituut voor Permanente Vorming (IVPV)
 Technologiepark 904, 9052 Zwijnaarde
 Tel: +32 9 264 55 82 / Fax: +32 9 264 56 05
 E-mail: ivpv@UGent.be