

Basiscursus Statistiek

Ingangsniveau:

Er wordt geen kennis van het vak statistiek noch van wiskunde verondersteld.

Eindtermen:

Na het doorlopen van de cursus is de cursist in staat om:

- 1) de basisbegrippen en -technieken van zowel beschrijvende als inductieve statistiek te omschrijven
- 2) statistische software-output correct te interpreteren
- 3) in de context van een specifieke onderzoeksvraag en bijbehorende data set op correcte wijze de geëigende statistische methoden en technieken toe te passen en de resulterende statistische uitkomsten correct te interpreteren

Opzet:

8 bijeenkomsten van 90 minuten met tussentijdse bestudering van theorie en het maken van opdrachten via een elektronische leeromgeving om de theorie toe te passen op concrete onderzoeksvragen en bijbehorende data sets.

Lesmateriaal:

De keuze van een studieboek hangt af van de doelgroep. Verder wordt er lesmateriaal uitgereikt.

Inhoud cursus:

Onderwerpen:

Inleiding:

waar gaat het vak statistiek over? In wat voor situaties is ze relevant? Om wat voor soort redering en uitspraken gaat het bij statistiek? Reikwijdte, beperkingen en korte geschiedenis van het vak

Beschrijvende statistiek:

Centrum- en spreidingsmaten

Meetniveaus

(Standaard)normale verdeling, z-score, standaardiseren

samenhang (*association*) en causaliteit

Inductieve statistiek:

beknopte inleiding tot de kansrekening, wet van de grote aantallen

data en kansverdeling(en): de basiselementen van een statistisch model

soorten kansverdelingen

Centrale Limiet Stelling

De twee hoofdvormen: schatten en toetsen

Basispatroon van een (frequentistische) schatting en toetsing

Parametrisch en non-parametrisch toetsen: diverse voorbeelden

bias and variability

observational en experimental designs

Reflecties op het vak statistiek

misverstanden rond de mogelijkheden en de beperkingen van het vak statistiek

kritiek op de significantietoets en enkele veel voorkomende misverstanden en misinterpretaties

beginselen van Bayesiaanse statistiek

Statistische software: R (*freeware en open source*), JASP (*freeware*), SPSS

Begrippenlijst:

Beschrijvende statistiek:

variabele

categorical and quantitative variables

gemiddelde, mediaan, modus, bereik (*range*), variantie, standaarddeviatie, interkwartielafstand

staaf- en cirkeldiagram, histogram, boxplot, *stem-and-leaf plot*

scheefheid, kurtosis, uitbijters (*outliers*)

covariantie, correlatie, puntenwolk (*scatter plot*), *least-squares regression line* en betekenis van *r-square*, *lurking variable*,

two-way tables: joint distribution – marginal distribution – conditional distribution

Inductieve statistiek:

kans, kansvariabele (*random variable*), discrete en continue kansvariabelen

kansverdeling, kansdichtheid (*density*) en cumulatieve kansverdeling

voorwaardelijke kans, produkt- en somregel voor kansen, begrip complementaire kans

populatie, steekproef, parameter en steekproefgrootheid (*statistic*)

statistisch model, modelspecificatie en modelselectie

soorten steekproeven

Cronbach's Alpha, betrouwbaarheid en validiteit

sampling variability, steekproevenverdeling (*sampling distribution*)

de drieslag: populatieverdeling, steekproefverdeling, steekproevenverdeling

betrouwbaarheidsinterval (*confidence interval*), foutmarge (*margin of error*), *confidence level*

onderzoeken van de aannames die aan een significantietoets ten grondslag liggen, het begrip robuustheid van een toets

toetsingsgrootheid (*test statistic*), significantietoets, statistische significantie, eenzijdig en tweezijdig toetsen, nulhypothese,

alternatieve hypothese, p-waarde, randomisatie (*randomness*), *independent trials*,

factoren die de steekproefomvang bepalen

statistical inference for sample counts, for sample proportions, for a sample mean, for a difference between sample proportions, for a difference between sample means, for regression coefficients

binomiale verdeling, Poissonverdeling, t-verdeling, F-verdeling, Chi-kwadraatverdeling

t-toets, variantieanalyse (ANOVA), enkelvoudige en meervoudige lineaire regressie, begrip *Generalized Linear Regression*

(GLM), factoranalyse, chi-kwadraattoets

interactie-effect (of gemodereerd effect), *lurking variable*, *confounding variable*, *covariate*, multicollineariteit

overdispersion, overfitted model

prior en posterior kansverdeling

voordelen van Bayesiaanse data-analyses