

Statistiek in de Natuurkundebachelor

Samenvatting

Statistiekonderwijs is belangrijk voor een bachelorstudent Natuurkunde, omdat statistische analyses regelmatig worden gebruikt in fysisch onderzoek voor het kwantitatief evalueren van de betrouwbaarheid en validiteit van resultaten. Hoewel in de natuurkundige praktijk meestal tamelijk informeel wordt gesproken over de statistiek, is het van belang dat de student een solide wiskundige basis krijgt voordat hij op meer intuïtief niveau met de stof leert omgaan. Alleen door de statistiek wiskundig te beschrijven, wordt men immers gedwongen om de aannames van de gebruikte methodes expliciet te maken, wat noodzakelijk is om de validiteit van de analyse te beoordelen.

Ik heb onderzocht hoe het statistiekonderwijs het best kan worden ingericht, om aan te sluiten op de doelen. In dit verslag geef ik hierover een reeks aanbevelingen. Ik adviseer om het onderwijs te verdelen over de vakken Statistiek voor Natuurkunde, waarin de wiskundige basis wordt gelegd zonder de concrete toepassing uit het oog te verliezen, en het Natuurkundepracticum, waarin de statistiek in concrete situaties moet worden toegepast.

1 Inleiding en Motivatie

Enkele jaren geleden is de bacheloropleiding Natuur- en sterrenkunde aan de VU volledig geherstructureerd. Bij deze herstructurering heeft onder andere het statistiekonderwijs in de bacheloropleiding een prominentere plaats gekregen dan voor de herziening het geval was. Dit onderwijs is vooral geconcentreerd in het nieuwe vak Statistiek voor Natuurkunde, en daarnaast onder andere in het natuurkundepracticum. Het vak Statistiek voor Natuurkunde is in januari 2010 voor het eerst gegeven, daarna heeft er een evaluatie plaatsgevonden over de aansluiting van de cursus op de wensen vanuit de opleiding. Uit deze evaluatie bleek dat er aanleiding was om deze aansluiting nog eens onder de loep te nemen.

Vanuit dat beginpunt heb ik een onderzoek gedaan naar wat de beste manier is om het statistiekonderwijs in de bacheloropleiding Natuurkunde vorm te geven. Hierbij heb ik, aan de hand van theorieën over didactische principes [Kallenberg, 2006], overleg gevoerd met de docenten van Statistiek voor Natuurkunde en van het natuurkundepracticum, en op basis daarvan en van mijn eigen ervaringen bij het leren van de statistiek* een advies geformuleerd hoe het statistiekonderwijs eruit zou moeten zien. Ik heb een aantal gemotiveerde aanbevelingen opgenomen over hoe het statistiekonderwijs in de opleiding geïmplementeerd zou kunnen worden. Als aanvulling op dit onderzoek heb ik een aantal opgaven geformuleerd, die aan de studenten zouden kunnen worden opgegeven om het statistiekonderwijs te ondersteunen; behalve dat deze opgaven een directe implementatie geven van het door mij geadviseerde onderwijs, kunnen ze ook gezien worden als een concrete illustratie van de strekking van mijn advies. (De opgaven zijn opgenomen in de bijlage bij dit verslag.)

In dit verslag zet ik de resultaten van dit onderzoek uiteen. Ik geef gemotiveerde aanbevelingen over leerdoelen, leerstof, vorm van het onderwijs en vorm en inhoud van de toetsing. Ik heb ervoor gekozen om niet bij alle gemaakte keuzes de motivatie expliciet weer te geven, omdat dat weinig zou toevoegen en omdat dit verslag dan drie keer zo dik zou zijn geworden. Waar ik heb ingeschat dat de motivatie van een keuze logisch volgt uit de algemene motivering van mijn keuzes of de keuze op basis van ‘common sense’ te verklaren is, heb ik de uitleg achterwege gelaten. Bij de uiteenzetting van

* Naast Natuurkunde heb ik ook de bacheloropleiding Econometrie gevolgd, waarin veel aandacht aan de statistiek werd besteed.

mijn advies heb ik hier en daar een verwijzing opgenomen naar de opgaven, als illustratie hoe ik dit zou implementeren.

De structuur van dit verslag is als volgt. Ik begin (paragraaf 2) met het inventariseren van de eisen die aan een bachelorstudent worden gesteld: wat moet hij uiteindelijk kunnen op het gebied van de statistiek, tijdens de masteropleiding en als hij in een onderzoeksgroep aan het werk gaat. Naar aanleiding daarvan formuleer ik het hoofddoel van het statistiekonderwijs. Daarna (paragraaf 3) volgt een beschouwing hoe dit doel bereikt kan worden. Ik ga in op de vraag hoe het statistiekonderwijs in de opleiding een plaats kan krijgen en hoe wiskundig het onderwijs zou moeten zijn.

In paragraaf 4 werk ik het ontwerp van het onderwijs uit, apart voor de vakken Statistiek voor Natuurkunde en het natuurkundepracticum. Voor Statistiek voor Natuurkunde, waarin de wiskundige basis van de statistiek wordt gelegd, doe ik dit uitgebreid, met een beschouwing over voorkennis, leerdoelen, stofomschrijving en de vorm van het onderwijs. Rond het natuurkundepracticum, waar het onderwijs zich alleen nog richt op de toepassing, is mijn bespreking korter. Tot slot, in paragraaf 5, beschrijf ik hoe de adviezen die ik in dit verslag neerleg zijn gebruikt in de cursus Statistiek voor Natuurkunde die in januari 2011 is gegeven, en maak ik een vergelijking met de cursus in 2010.

2 Eisen aan een bachelorstudent

De bacheloropleiding traint de student om een goed fysicus te worden. Als we willen bepalen hoe het statistiekonderwijs in de bacheloropleiding eruit moet zien, moeten we ons dus eerst afvragen hoe een goed fysicus de statistiek gebruikt en op welk niveau hij met de materie omgaat. In dit hoofdstuk zal ik op de twee aspecten van die laatste vraag een antwoord formuleren. Daarna behandel ik een globaal antwoord op de eerste vraag, zoals dat vanuit de opleiding zelf gegeven is. Het blijkt dat de visie van de opleiding op het statistiekonderwijs goed aansluit bij mijn idee over het gebruik van de statistiek door een fysicus.

De rol van de statistiek binnen de fysica

In de fysica speelt de statistiek een ondersteunende rol in de analyse van experimentele gegevens. Meestal wordt een model met parameters gefit aan de waarnemingen of worden de parameters van een kansmodel onderzocht. De belangrijkste toepassingen zijn dan het schatten van de parameters, het bepalen van de onzekerheid in die schattingen en het vergelijken van de resultaten van de parameterschattingen met elkaar of met een literatuurwaarde.

De toepassing van statistiek in de data-analyse voegt aan de waarnemingen een kwantitatieve onderbouwing toe van de conclusies, op het gebied van validiteit –kan de theorie op de waarnemingen worden toegepast– en betrouwbaarheid –hoe zeker kan men zijn van de conclusies. Dit is een aantrekkelijke situatie omdat de statistische analyse kwantitatief en objectief controleerbaar is. Het is dus wenselijk dat iemand die wordt opgeleid tot fysicus leert de statistiek voor deze doelen te gebruiken.

Het is van essentieel belang om te bedenken wat de beperkingen zijn van de gebruikte statistische methoden. De toepasbaarheid van een statistische analyse scharniert immers op de geldigheid van de aannames die worden gedaan in de vertaling van de fysische situatie naar de statistiek – ook al worden die aannames niet altijd expliciet gemaakt, om de statistische methoden te kunnen gebruiken, moeten ze wel degelijk worden gedaan. De aannames over de waarnemingen hebben een sterke invloed op de validiteit van de analyse en om die validiteit te beoordelen moet een fysicus de rol van de aannames dus goed begrijpen. Dit is, naast het kunnen toepassen van de statistiek, een belangrijke eis die aan een fysicus in opleiding gesteld zou moeten worden.

De manier waarop statistiek wordt gebruikt

Fysici houden van intuïtief denken. Intuïtie kan gezien worden als een hoog niveau van begrip, en een intuïtief begrip van een onderwerp is dus ook in het onderwijs nastrevenswaardig. Dit geldt ook voor het statistiekonderwijs. Met de concepten uit de statistiek die een fysicus nodig heeft moet de student op intuïtief niveau leren omgaan. Intuïtie alleen is echter niet voldoende: er zijn situaties waarin met een intuïtieve benadering een probleem niet kan worden opgelost. Wanneer de intuïtie tekort schiet, moet men terug kunnen vallen op de wiskunde en het probleem in detail uitwerken. Het is dus ook belangrijk dat het onderwijs zorgt voor beheersing van de wiskundige details.

Opgemerkt moet worden dat het ook in didactische zin zinvol is om de wiskundige details in het onderwijs aan bod te laten komen. Zou het onderwijs immers alleen op de ontwikkeling van een intuïtief begrip van dit wiskundige onderwerp zijn gericht, dan zou de student nooit leren volledig op zijn intuïtie te vertrouwen omdat hij weet dat hij de achterliggende theorie niet werkelijk begrijpt.

Eisen geformuleerd door de opleiding

Vanuit de opleiding natuur- en sterrenkunde zelf zijn de eisen geformuleerd die aan een bachelorstudent worden gesteld. Deze eisen zijn onder andere geformuleerd door de commissie die enkele jaren geleden het curriculum van de opleiding heeft gewijzigd. In het rapport van de curriculumcommissie [Wuite, 2008] is de positie van data- en foutenanalyse hierbij expliciet vermeld:

“De basiskennis en vaardigheden voor de vernieuwde opleiding omvatten:

Kennis van basisbegrippen in de mechanica, quantummechanica, relativiteitstheorie, electrodynamica, thermodynamica, statistische fysica, biofysica en medische fysica;

daarbij behorende basiskennis van toepassingen op de structuur en dynamica in de natuur;

beheersing van het mathematisch-fysische vocabulair;

*meten en theorievorming middels gestructureerde experimenten, **inclusief adequate data- en fouten-analyse**;*

adequate mondelinge en schriftelijke vaardigheden voor communicatie met vakgenoten en niet-vakgenoten op bachelor niveau.”

Verderop in het rapport blijkt dat de commissie het doel van ‘adequate data- en fouten-analyse’ onder andere wil implementeren in een vak Kansrekening, Statistiek en Data-analyse, ter waarde van 3 studiepunten. De kennis die een student bij dit vak moet verwerven, wordt beschreven als volgt: *“Kansverdelingen, theorie parameterschatting, schatters (maximum likelihood, kleinste kwadraten), covariantiematrix, betrouwbaarheidsintervallen, correlatie, toetsen, modelsimulatie.”* De bijbehorende competenties die volgens het rapport moeten worden verworven zijn: *Toepassing in eenvoudige data analyse (rechte lijn, exponential decay). Interpretatie van residuen en beoordeling kwaliteit model. Gebruik van Origin.* Uit het rapport blijkt dus dat de opleiding zich inderdaad richt op de toepassing van de statistiek in experimenteel onderzoek en dat tegelijkertijd de wiskunde erachter ook belangrijk wordt gevonden.

Buiten dit rapport om heb ik ook enkele gesprekken gevoerd met wetenschappelijk medewerkers van verschillende natuurkunde-vakgroepen aan de VU, die allen met studenten hebben gewerkt in het kader van bachelor- of masterprojecten. Het betreft hier Jan de Munck van de vakgroep Medische Fysica, Sander Verbrugge van de vakgroep Fysica van Complexe Systemen, en Gerhard Raven van het NIKHEF. Hun heb ik gevraagd welke vaardigheden een student nodig heeft op het gebied van statistiek, wanneer hij of zij in de betreffende groep onderzoek zou gaan doen. Het antwoord kwam bij allen op hetzelfde neer: de student zou enige basiskennis moeten hebben van de statistiek, met onder andere een begrip van schatten, bepaling van onzekerheden, fitten. Specifieke kennis en vaardigheden uit de statistiek, die een student bij een specifiek project nodig heeft, kan hij zich dan ter plekke eigen

maken. Inderdaad bleek op basis van de gesprekken dat de statistische methoden die in de vakgroepen gebruikt worden, soms zeer specifiek zijn voor de toepassing.

Uit deze gesprekken komt, naast een onderschrijving van de uitspraken in het rapport van de curriculumcommissie, de extra eis naar voren dat de student in staat moet zijn om zijn kennis van de statistiek uit te breiden indien dat voor een specifiek onderzoek nodig is. Uit het feit dat de studenten in de vakgroepen in aanraking kunnen komen met andere statistische methoden, kan worden afgeleid dat de student genoeg affiniteit met de basis van de statistiek moet krijgen om extra onderwerpen uit de statistiek te kunnen bestuderen op het moment dat hij die in een specifiek onderzoeksproject nodig heeft.

Doel van het statistiekonderwijs

In de paragrafen hierboven zijn de eisen uiteengezet die op het gebied van de statistiek aan een student gesteld moeten worden zodat hij het niveau bereikt dat een fysicus nodig heeft. Deze eisen vormen het doel van het statistiekonderwijs. Samenvattend is het doel van het statistiekonderwijs dat de student in zodanige mate affiniteit krijgt met de statistiek, zowel op intuïtief als op wiskundig niveau, dat hij de statistiek kan gebruiken als hulpmiddel bij het kwantitatief evalueren van de validiteit en betrouwbaarheid van de resultaten van een experiment, en dat hij zijn arsenaal aan statistische methoden kan uitbreiden wanneer dat nodig blijkt te zijn voor een specifiek experiment.

In de rest van dit verslag wordt uitgewerkt hoe dit doel bereikt kan worden.

3 Structuur van het statistiekonderwijs

Nu het doel van het statistiekonderwijs helder is, is het de vraag hoe het onderwijs ingericht moet worden opdat het doel daadwerkelijk bereikt wordt. In dit hoofdstuk ga ik in op die vraag, door verschillende aspecten van het onderwijs te bespreken. Ik zet in dit hoofdstuk mijn belangrijkste didactische keuzes uiteen, met een toelichting waar ik dat nodig acht.

Zoals boven al is aangegeven, is het uiteindelijke doel dat de student een intuïtief gevoel ontwikkelt voor de concepten in de statistiek. Ook heb ik boven al kort aangegeven dat dit mijns inziens niet mogelijk is zonder eerst de statistiek gestructureerd wiskundig te behandelen. De statistiek is een wiskundige theorie die eigenlijk niet zou moeten worden toegepast op een probleem, zonder het probleem eerst netjes wiskundig te modelleren. Pas wanneer je jezelf ervan hebt overtuigd dat je de wiskundige details wel kent, en deze zou kunnen uitwerken indien nodig, is het mogelijk om de precieze wiskundige beschrijving los te laten en over te stappen op een meer intuïtieve beschrijving van het probleem. Een wiskundige aanpak van de stof is ook belangrijk omdat een begrip van de structuur van de statistiek zeer nuttig is om de mogelijkheden en beperkingen van de theorie te overzien. Om deze redenen pleit ik ervoor dat het statistiekonderwijs begint met een gestructureerde wiskundige behandeling van de stof, en dat het onderwijs zich daarna richt op de toepassing van de statistiek in fysische experimenten. Dit advies sluit aan bij de praktische implementatie van de opleiding. Er is immers het vak Statistiek voor Natuurkunde, dat gegeven wordt door een wiskundige, en er is het natuurkundepracticum, waar plaats zou zijn voor de toepassing in de experimentele praktijk.

De wiskundige behandeling van de statistiek, in het vak Statistiek voor Natuurkunde, zou echter niet al te theoretisch moeten zijn: indien te veel op de structuur van de theorie zou worden gefocust, dreigt het gevaar dat de aansluiting op de toepassing in het gedrang komt. Het is juist die aansluiting op de experimentele praktijk die voor een fysicus zo belangrijk is, en ook tijdens de wiskundige behandeling van de theorie zou de student al voortdurend verbanden moeten zien tussen theorie en

toepassing. Door het verschil in aanpak dat wiskundigen en natuurkundigen van nature hebben, is het belangrijk om ervoor te waken dat deze brug in stand blijft wanneer het vak gegeven wordt door een docent van de afdeling wiskunde. Dit verschil in aanpak wordt veroorzaakt door het volgende: Van nature is een wiskundige geneigd om eerst een solide basis voor de theorie te leggen en daarna de theorie stap voor stap op te bouwen. Een natuurkundige denkt eerder vanuit een concreet probleem dat hij heeft, en gaat vanuit dat probleem de theorie in, net zolang totdat hij het probleem voldoende begrijpt. Nadeel van de wiskundige aanpak is dat niet gegarandeerd is dat voldoende stof wordt behandeld om in natuurkundige toepassingen goed toegerust te zijn, nadeel van de natuurkundige aanpak is dat het mogelijk is dat de structuur van de statistiek als geheel niet genoeg in beeld komt. In het vak Statistiek voor Natuurkunde zou derhalve een evenwicht gezocht moeten worden tussen de structuur van de wiskunde en de toepassing in de natuurkunde. In de inhoud van het vak en in de nadruk die op de verschillende onderdelen van de stof wordt gelegd, moet de aansluiting op de toepassing in het oog worden gehouden. Naast de colleges speelt dit een belangrijke rol in de opgaven die studenten moeten maken, en in de keuze voor de vorm van de toetsing van de opgedane kennis.

In het natuurkundepracticum komt de statistiek aan de orde als onderdeel van de experimentele praktijktraining. Het onderwijs heeft hier twee doelen. Enerzijds verdiept de student zijn begrip van de statistiek door het bij Statistiek voor Natuurkunde geleerde toe te passen in de praktijk; hoewel Statistiek voor Natuurkunde op de praktische toepassing gericht is, zal de student daar steeds te maken krijgen met problemen die al betrekkelijk duidelijk geformuleerd zijn, terwijl in het natuurkundepracticum de formulering van het probleem onderdeel is van de opdracht. Anderzijds moet de student in het natuurkundepracticum leren om te communiceren over de statistische methoden en resultaten op een minder formele (intuïtieve) manier dan hij tijdens Statistiek voor Natuurkunde geleerd heeft, in aansluiting op wat in de praktijk van hem verwacht zal worden.

Dit is de globale opzet van het onderwijs. Ik zal dit in de rest van dit verslag verder concreet maken.

4 Ontwerp van het onderwijs

In dit hoofdstuk werk ik in meer detail uit hoe het statistiekonderwijs eruit zou moeten zien. Dit doe ik apart voor Statistiek voor Natuurkunde en voor het Natuurkundepracticum. Voor Statistiek voor Natuurkunde ben ik tamelijk uitgebreid: ik bespreek onder andere leerdoelen en inhoud van het vak en de vorm van het onderwijs. Ook voor het natuurkundepracticum doe ik aanbevelingen hoe de doelen van het onderwijs bereikt kunnen worden, maar daar ben ik iets beknopter over.

4.1 Statistiek voor Natuurkunde

Voor het ontwerpen van onderwijs zijn in de literatuur diverse handreikingen beschikbaar. Ik baseer mij losjes op het werk van Kallenberg e.a. (2006). Zij adviseren te beginnen met het in kaart brengen van de voorkennis van de student, daarna de leerdoelen van het onderwijs en de leerstof vast te stellen, en vervolgens beslissingen te nemen over de vorm van het onderwijs. Deze structuur volg ik ook hier.

Voorkennis

Het is belangrijk om de voorkennis van de studenten in kaart te brengen voordat met het ontwerpen van het onderwijs kan worden begonnen, opdat met het onderwijs op deze voorkennis kan worden aangesloten.

Het vak Statistiek voor Natuurkunde heeft plaats in de derde periode van het tweede jaar van de opleiding. Tot dat punt in de opleiding zijn kansrekening of statistiek niet expliciet aan bod gekomen.

Bij quantummechanica, dat begint in de tweede periode van het tweede jaar, wordt wel gebruik gemaakt van enige begrippen en technieken uit de kansrekening, waaronder het begrip van een kansdichtheidsfunctie (daar in de vorm van een golffunctie) en het berekenen van verwachtingswaarde en variantie van een continue variabele. Op die plaats worden deze onderwerpen ad hoc geïntroduceerd.

Op het VWO is de kansrekening en statistiek niet langer een verplicht onderdeel van de examenstof voor wiskunde B [College voor Examens, 2010]. Bij wiskunde A is dit wel het geval, maar waarschijnlijk zullen de meeste natuurkundestudenten wiskunde A niet in hun pakket hebben gehad. Het is wel goed mogelijk dat bij wiskunde B de kansrekening door de school op eigen initiatief is behandeld.

Het is dus mogelijk dat er studenten zijn die nog niet of nauwelijks in aanraking met de kansrekening zijn gekomen. Ik acht het waarschijnlijk dat alle studenten wel enige intuïtie voor de kansrekening hebben ontwikkeld voordat zij aan dit vak beginnen. De kansrekening mag echter niet bekend worden verondersteld, hier zal tijdens de cursus dus rekening mee moeten worden gehouden.

Leerdoelen

De tweede stap in het ontwerpen van onderwijs, volgens Kallenberg (2006), is het vaststellen van de leerdoelen. In deze paragraaf zet ik uiteen wat ik denk dat de studenten aan het eind van de cursus aan kennis en vaardigheden moeten hebben vergaard. Hierbij speelt de praktische randvoorwaarde, dat voor statistiek in de opleiding maar een beperkte hoeveelheid tijd beschikbaar is, een belangrijke beperkende rol.

Mijn uitgangspunt zal de veronderstelling zijn dat de statistiek voor een fysicus niet een doel op zich is, maar een middel om de resultaten van fysische experimenten kwantitatief te analyseren. Dit uitgangspunt is steeds leidend en is de uiteindelijke motivatie voor alle keuzes die ik in dit onderzoek heb gemaakt. Deze veronderstelling is gebaseerd op de wensen die vanuit de opleiding zijn gesteld (zie aldaar) en op basis van mijn eigen ervaring met de natuurkunde en de statistiek en met collega-natuurkundigen; ik acht het waarschijnlijk dat vrijwel elke fysicus deze veronderstelling zal delen.

Ik zal de leerdoelen formuleren op verschillende niveaus. Hoewel dit niet in de literatuur wordt gesuggereerd, heb ik bij het opstellen van de leerdoelen de behoefte gevoeld om op die manier structuur in de formulering aan te brengen. Het hoogste niveau beslaat de leerdoelen in algemene termen, direct aansluitend op het boven geformuleerde doel van het onderwijs. Daaronder plaats ik een tweede niveau van leerdoelen, waarin de leerdoelen op het hoogste niveau worden gespecificeerd en toegelicht. Bij het formuleren van de leerstof, in de volgende paragraaf, splits ik de leerdoelen in nog meer detail uit; deze uitsplitsing vormt een nog lager niveau van leerdoelen.

Zoals eerder aangegeven heb ik ook een reeks inleveropgaven ontwikkeld die bij de cursus Statistiek voor Natuurkunde gebruikt zouden kunnen worden. Deze opgaven zijn opgenomen in de bijlage. In deze paragraaf geef ik regelmatig een verwijzing naar een opgave, waarin het betreffende leerdoel aan de orde komt.

➤ *Leerdoelen op hoog niveau*

In het licht hiervan, en gezien de wensen vanuit de opleiding, kan het belangrijkste leerdoel van het statistiekonderwijs geformuleerd worden als:

- In praktijksituaties, waar relevant, kunnen gebruiken van enkele veel voorkomende statistische technieken als hulpmiddel bij het analyseren van experimentele resultaten, en bij het kwantitatief evalueren van de validiteit en betrouwbaarheid van conclusies

Welke veel voorkomende statistische technieken bedoeld worden, wordt verderop aangegeven. Merk op dat dit leerdoel veelomvattend is: onder andere is een begrip van de gebruikte methoden noodzakelijk om te kunnen beoordelen of de methode zelf toepasbaar is op de voorliggende situatie.

Een tweede belangrijk leerdoel is:

- Een voldoende basisniveau bereiken van kennis en begrip van de principes achter de statistische technieken, om met behulp van aanvullende literatuur andere statistische technieken te kunnen begrijpen.

De aanvullende literatuur die de student nodig heeft om zich specifieke statistische methoden eigen te kunnen maken, zal in de meeste gevallen uitgaan van een bepaalde basiskennis van de statistiek. Hierover moet de student dus beschikken, wat wordt uitgedrukt in dit leerdoel.

In het kort behelzen deze twee doelen dus begrip van de kansrekening en statistiek en van de gekozen statistische methoden, alsmede de vaardigheid om de geschiktheid van de geleerde methoden in een praktijksituatie te beoordelen en een geschikte methode toe te passen.

► *Leerdoelen op lager niveau*

Het is een belangrijke keuze hoe wiskundig de insteek van het onderwijs is. Ik kies ervoor om deze insteek tamelijk precies te maken: eerst wordt een solide basis in de kansrekening gelegd, en daarna wordt, wiskundig precies, de statistiek opgebouwd. Reden hiervoor is in het kort dat ik denk dat je een begrip in wiskundig detail nodig hebt om de mogelijkheden en beperkingen van de statistiek te kunnen begrijpen. Ik licht dit meer toe aan het einde van dit verslag. Voor nu stel ik vast dat deze keuze tot uitdrukking zal komen in de invulling van de leerdoelen.

Om de ‘leerdoelen op hoog niveau’ te kunnen bereiken, zal de student een begrip van de kansrekening moeten ontwikkelen en daarna de belangrijkste principes van de statistiek moeten leren. Begrip is nodig om de geschiktheid van een methode in een gegeven situatie te beoordelen. Om dit nader te specificeren, formuleer ik nu de leerdoelen op een iets lager niveau.

Hier en daar neem ik termen op als ‘enkele begrippen’, ‘sommige gevallen’. Dit is vaag, maar voor de overzichtelijkheid leg ik die termen hier nog niet verder uit. In de paragraaf over de stofomschrijving doe ik dat alsnog.

Kansrekening

Leerdoel voor de kansrekening is:

- Kennen en in eenvoudige situaties kunnen toepassen van enkele basale begrippen en verdelingen uit de kansrekening (*komt aan de orde in opgavenserie 1, zie bijlage*).

De kansrekening is een precieze wiskundige theorie en het is wenselijk dat de student dat ten minste beseft. Om de statistische teksten te begrijpen is het van belang dat de studenten op een eenvoudig niveau kunnen werken met begrippen uit de kansrekening.

Statistisch modelleren

Leerdoelen:

- Uitleggen dat een statistisch model nodig is om statistiek te kunnen toepassen op een fysisch probleem
- Een statistisch model kunnen opstellen bij een gegeven fysisch probleem, en de validiteit van het model bediscussieren
- De conclusies uit een statistische analyse terugvertalen naar conclusies over de fysische context

Statistiek

Leerdoelen:

- Kunnen schatten van de parameters van een gegeven kansverdeling met behulp van de maximum likelihood methode (indien de kansverdelingsfunctie direct beschikbaar is), en interpreteren van deze schattingen, en het principe van parameterschatten in het algemeen en van de maximum likelihood methode in het bijzonder kunnen omschrijven (*serie 2*)
- Kunnen bepalen en interpreteren van een betrouwbaarheidsinterval voor de schattingen van de parameters (in sommige gevallen, waaronder ML-schattingen bij sommige verdelingen) en kunnen beschrijven volgens welke principes een uitdrukking voor het betrouwbaarheidsinterval is bepaald (*serie 3*)
- Een statistische toets kunnen uitvoeren naar een hypothese over de parameters (in sommige gevallen) en kunnen beschrijven op welke principes de toets en de conclusie is gebaseerd (*serie 3*)
- Kunnen beargumenteren of de geleerde statistische methoden al dan niet geschikt zijn om toe te passen in een gegeven situatie

In welke gevallen men dit precies moet kunnen, moet nader worden gespecificeerd in de te behandelen stof. Hierbij spelen ook overwegingen over de beschikbare tijd een rol. In de paragraaf over de stof kom ik hierop terug.

Regressie

De belangrijkste toepassing is de lineaire regressie: het fitten van een model dat lineair is in de parameters aan de data. Dit komt vaak voor in het practicum. (*Zie opgave 2.4 en serie 4.*)

Hiervan zijn de leerdoelen:

- Kunnen toepassen van de kleinste-kwadratenmethode in een concrete situatie: schatten van de parameters van een lineair model, bepalen van betrouwbaarheidsinterval voor de parameters, hypothesen over de waarde van de parameters toetsen.
- Uitleggen op welke principes de kleinste-kwadratenmethode is gebaseerd.
- Kunnen beschrijven wat de verschillen in aanpak van de regressie-analyse zijn als de standaardafwijking van de meetfouten in het regressiemodel wel of niet gelijk is voor alle meetpunten, en wel of niet precies bekend is; en in al deze situaties een correcte analyse uitvoeren.

Als uitbreiding op de lineaire regressie zijn tevens de volgende leerdoelen te formuleren. Deze liggen minder direct in de lijn van het centrale doel, maar zijn toch ook belangrijk genoeg om er in een inleidende cursus aandacht aan te besteden.

- Een correcte regressieanalyse uitvoeren in een regressiemodel dat niet-lineair in de parameters is, en de principes beschrijven waarop deze analyse gebaseerd is.
- Weten dat parameterschattingen gecorreleerd kunnen zijn, en in zo'n situatie rekening houden met correlaties in berekeningen met de schattingen (foutenvoortplanting).

Stofomschrijving

In dit hoofdstuk geef ik een nadere omschrijving van de leerstof, als verdere precisering van wat de student moet leren. De stof is een afgeleide van de leerdoelen en hangt hier zeer nauw mee samen.

Ik heb ervoor gekozen om de stof te formuleren in de vorm van leerdoelen. Hiervoor heb ik twee redenen. Enerzijds ontstaat zo een duidelijke omschrijving, waarbij ook duidelijk wordt op welk niveau en in hoeveel detail de stof behandeld moet worden. Anderzijds sluit dit aan bij de boven geformuleerde leerdoelen, waarvan het aspect dat de leerinhoud beschrijft nog verder moest worden uitgewerkt. Op deze manier ontstaat een geheel van leerdoelen dat op drie niveaus is geformuleerd: de twee hoogste niveaus zijn behandeld in het hoofdstuk Leerdoelen en hier behandel ik het laagste, meest gedetailleerde niveau.

Ik splits de leerstof hier uit naar de verschillende onderwerpen. De stof volgt in grote mate direct uit de leerdoelen op hogere niveaus; waar ik een toelichting nodig heb geacht is die opgenomen.

► *Kansrekening*

Eerst moet een overzicht gegeven worden van de kansrekening. Dit betreft de basisbegrippen, al deze zaken komen later in de stof terug, toegepast in de statistiek.

Van de volgende begrippen moet de student kunnen aangeven wat het is (in globale termen), de wiskundige definitie gebruiken in opgaven.

- Wat is een kans: mathematische definitie. Reden om dit te behandelen: student moet beseffen dat kans een wiskundig object is.
- Rekenregels voor kansen: bijvoorbeeld complementregel (kans op A is 1 min kans op niet A), kans op A of B is kans op A plus kans op B min kans op A én B (*opgave 1.1*).
- Stochastische variabelen, kansverdelingen, discreet en continu. Kansverdelingsfunctie, kansdichtheidsfunctie.
- Verwachtingswaarde, variantie, standaardafwijking (voor zowel discrete als continue verdelingen). Tevens het onderscheid met corresponderende steekproefgrootheden aangeven.
- Standaard verdelingen, interpretatie van de verdeling in woorden en waar relevant samenhang met andere verdelingen: alternatieve, binomiale, geometrische, poisson verdeling; homogene, normale, exponentiële verdeling (*opgave 1.6*).
- De kansverdelingsfunctie geven van deze verdelingen, eventueel na opzoeken.
- Afhankelijkheid van stochasten, covariantie, correlatie.
- Wet vd grote aantallen, centrale limietstelling: herkennen wiskundige notatie, uitleggen in woorden wat het betekent. CLS tevens gebruiken om kansen te benaderen (*opgave 1.4*).

► *Modelleren*

De stap van de fysische context naar een statistisch-mathematisch model is belangrijk en de student moet deze kunnen maken. In aansluiting op het gestelde leerdoel worden in de cursus een aantal voorbeelden gegeven van de vertaling van context naar model. Dit oefent de student zelf ook in een aantal opgaven. De leerdoelen zijn:

- In eigen woorden uitleggen dat de geschiktheid van een statistisch model voor het beschrijven van de fysische situatie, van invloed is op de validiteit van de conclusies die uit de statistische analyse getrokken worden (*bijvoorbeeld in opgaven 2.3h, 3.1f*).
- Op basis van een omschrijving van een experiment in informele termen, een geschikt statistisch model voor de waarnemingen opstellen, en aangeven welke aannames daarbij (expliciet en impliciet) worden gedaan (*bijvoorbeeld 2.1*).
- De validiteit van gedane aannames in een statistisch model bediscussiëren in relatie tot de fysische context en tot de waarnemingen zelf, waar relevant met hulp van een histogram of andere grafische representaties van de data (*bijvoorbeeld 3.2a*).

► *Schatten*

Een korte algemene inleiding over schatten en schatters: wanneer is een schatter ‘goed’? De maximum likelihood schatter is vaak ‘goed’, deze wordt dan ook vaak gebruikt en kan volgens een algemene methode worden berekend, daarom wordt deze methode behandeld.

- Aangeven wat een schatter is en uitleggen welke eigenschappen een goede schatter moet hebben.
- De parameters van een normale verdeling kunnen schatten met steekproefgemiddelde en steekproefstandaarddeviatie, en aangeven dat deze schatters zuiver zijn (*bijvoorbeeld 3.2a*).
- In eigen woorden beschrijven wat een maximum likelihood schatter voorstelt, en deze analytisch kunnen berekenen voor een steekproef uit een gegeven verdeling (wanneer de berekeningen niet te ingewikkeld worden) (*opgaven 2.1, 2.2, 2.3*).

➤ *Betrouwbaarheidsintervallen*

De student moet weten wat de precieze interpretatie is van een betrouwbaarheidsinterval, deze is immers vrij subtiel, en waarom een betrouwbaarheidsinterval nuttige informatie geeft. Betrouwbaarheidsintervallen worden in het algemeen berekend aan de hand van zogenaamde pivots, dit moet de student dus ook kunnen doen. Tevens moet hij deze procedure uit kunnen voeren (met een theorieboek ernaast is voldoende) voor de maximum-likelihood schatter.

- Uitleggen waarom een (punt)schatting van een parameter niet voldoende informatie geeft, en aangeven dat de missende informatie met een betrouwbaarheidsinterval kan worden bepaald
- In eigen woorden een correcte interpretatie van een betrouwbaarheidsinterval beschrijven (*opgave 3.5*)
- In eenvoudige gevallen een betrouwbaarheidsinterval opstellen voor een parameter (met name van de normale verdeling) met behulp van een geschikte pivot – niet uit het hoofd maar met een voorbeeld erbij (*opgaven 3.2, 3.3*)
- Een maximum-likelihood betrouwbaarheidsinterval bepalen met behulp van de waargenomen fisher-informatie (*opgave 3.4*)

➤ *Toetsen*

De student moet bij toetsen omtrent de parameters van een normale verdeling, een geschikte toets kunnen kiezen (in het geval van een normale verdeling met verwachting μ en standaardafwijking σ : toetsen voor μ als σ bekend is, voor μ als σ onbekend is, voor σ als μ onbekend is, en voor twee steekproeven uit mogelijk verschillende verdelingen). Tevens de Wilcoxon toets als voorbeeld van een verdelingsvrije toets (dit is een toets voor twee steekproeven die niet de aanname doet dat de steekproeven uit een normale verdeling komen) opdat de student weet dat de normaliteitsaanname niet noodzakelijk is.

- In eigen woorden omschrijven wat het principe van een statistische toets is
- Voor de volgende toetsen aangeven op welk statistisch model (met aannames) ze gebaseerd zijn, welke hypothesen er getoetst worden en met welke toetsingsgrootte ...
- ... en de toetsen quasi-handmatig kunnen uitvoeren met behulp van een naslagwerk: gauss-toets (normale verd., toets voor μ , σ bekend), t-toets (normale verd., toets voor μ , σ onbekend), chikwadraattoets (voor σ^2 van een normale verdeling als μ onbekend is), gepaarde t-toets (toets voor de μ 's van twee steekproeven uit normale verdeling, waarbij twee keer aan hetzelfde individu een meting is gedaan), twee steekproeven t-toets (analoog voor ongepaarde waarnemingen), wilcoxon rangtekentoets (zie boven) (*opgaven 3.1, 3.3*).
- Bij een gegeven context en probleemstelling, van de bovenstaande toetsen aangeven of ze geschikt zijn om te gebruiken; de uitkomst van de toets correct interpreteren en vertalen naar de context.

➤ *Lineaire regressie*

Dit is de belangrijke toepassing die het meest gebruikt wordt in de praktijk. De student moet met name weten wat er gebeurt als hij op de fit-knop drukt in een computerprogramma. Hij moet de verschillen begrijpen tussen fitten met bekende of onbekende fouten, en inzien hoe gewichten / ongelijke fouten de resultaten beïnvloeden.

- Bij een gegeven fysisch lineair model een standaard statistisch lineair regressiemodel opstellen voor de waarnemingen en aangeven welke aannames in dit model worden gedaan (*opgaven 4.1, 4.2*)
- Uitleggen hoe met behulp van de kleinste kwadratenmethode schatters voor de regressieparameters van dit model kunnen worden gevonden

- Aantonen dat in het geval van normaal verdeelde meetfouten, de kleinste-kwadratenschatters voor de regressieparameters ook de maximum likelihood schatters zijn
- Kleinste kwadraten schattingen uitrekenen met behulp van een computerpakket (*opgaven 4.1, 4.2*)
- Aangeven dat de verdeling van de kleinste-kwadratenschatters normaal is als de meetfouten normaal verdeeld zijn
- De normaliteitsaannname van de meetfouten bediscussiëren op basis van de residuen van de fit (*opgave 4.2e*)
- Een betrouwbaarheidsinterval voor de regressieparameters opstellen (*opgave 4.2g*)
- Aangeven wat de standaardfout, de t-waarde en de p-waarde betekenen, die door een statistisch computerpakket vaak bij een regressie-analyse worden berekend.
- In eigen woorden, op informele wijze, uitleggen hoe ongelijkheid van de standaardafwijking van de verdeling van de respectievelijke meetfouten de schattingen van de regressieparameters beïnvloedt (regressie met gewichten). Ook: de gewogen regressie uitvoeren in een computerpakket (*opgave 4.3*).
- Aangeven dat de standaardafwijking van de meetfouten kan worden geschat als deze niet van tevoren bekend is, en deze schatten met behulp van een computerpakket (*opgave 4.2d*)
- Aangeven dat de standaardafwijking van de meetfouten vaak van tevoren bekend is door een onafhankelijke bepaling: als betaald wordt door deze extra informatie aan het model toe te voegen, krijg je extra mogelijkheden, zoals hieronder blijkt (*opgave 4.2j*).
- Aangeven dat een goodness-of-fit chikwadraat toets kan worden uitgevoerd als de standaardafwijking van de meetfouten van tevoren bekend is; aangeven welke hypothese dan getoetst wordt ("het model beschrijft de data correct") en op welke toetsingsgrootte deze toets gebaseerd is, en de toets uitvoeren (*opgave 4.2h*).
- Aangeven hoe de bepaling van een betrouwbaarheidsinterval en het toetsen van een hypothese op de waarde van de regressieparameters veranderen, in het geval dat de standaardafwijkingen van de meetfouten van tevoren bekend zijn (toets en betrouwbaarheidsinterval worden dan gebaseerd op een normale verdeling in plaats van een t-verdeling) (*opgave 4.2j*).

➤ *Niet-lineaire regressie*

Omdat dit ook regelmatig voorkomt, is het wenselijk dat de student ook hier weet hoe deze methode door de computer wordt uitgevoerd.

- In eigen woorden globaal omschrijven hoe de parameters van een niet-lineair statistisch model geschat kunnen worden.
- Omschrijven hoe onzekerheden in de parameters in dit geval worden bepaald
- Deze berekeningen uitvoeren met behulp van een computerpakket

➤ *Computervaardigheden*

De student moet de statistische methoden kunnen uitvoeren met behulp van de computer. Er zijn statistische softwarepakketten beschikbaar, waarbij alleen de data ingevoerd hoeven te worden, en met een druk op de juiste knoppen worden dan allerlei analyses op de data uitgevoerd. Het gevaar van die pakketten is dat het meer draait om de interpretatie van de resultaten dan om het begrip waar de resultaten vandaan komen. Het gebruik van een dergelijk pakket is voor de natuurkundestudent niet nodig en ook niet wenselijk.

Nuttig is wel om bepaalde berekeningen niet met de hand te hoeven uitvoeren: het rekenwerk zelf moet door de computer worden gedaan. Ook mogen bepaalde processen best geautomatiseerd worden uitgevoerd. Programma's die geschikt zouden kunnen zijn, zijn R en Mathematica. Vooral met het

programma Mathematica werkt de student ook op andere momenten in de bacheloropleiding. In de laatste versies zijn de mogelijkheden van dit programma op het gebied van de statistiek flink uitgebreid en ik pleit er dan ook voor dat dit programma als hulpmiddel bij het statistiekonderwijs wordt gebruikt. Praktisch probleem is dat de huidige docent (Marianne Jonker) heeft gezegd dat zij te weinig affiniteit heeft met Mathematica en dus niet optimaal hulp kan bieden als studenten zouden vastlopen tijdens het gebruik ervan. Zij heeft gezegd dat zij om deze reden de voorkeur geeft aan het gebruik van het programma R. Op zich is dat inderdaad een nuttig programma, maar nadeel is dat het voor de meeste studenten nieuw is en dat zij dus veel tijd zouden moeten besteden aan het onder de knie krijgen van dit programma. Ik neig ernaar te denken dat die tijd beter aan het leren van de statistiek besteed zou kunnen worden.

Op het natuurkundepracticum wordt veel gebruik gemaakt van Origin Pro, voor het grafisch weer-geven en analyseren van data. In dit programma zijn uitgebreide mogelijkheden opgenomen voor het automatisch fitten van diverse functies aan de data. De studenten zullen hier dus ook mee moeten leren werken: zeker bij lineaire regressie moeten ze weten hoe de resultaten van statistische berekeningen weergegeven kunnen worden.

Vorm van het onderwijs

In dit hoofdstuk bediscussieer ik enkele mogelijkheden om het onderwijs vorm te geven, en beargu-menteer ik welke keuze mijns inziens de beste is.

Bij het bepalen van de vorm van het onderwijs spelen, naast didactische, ook praktische factoren een rol. Het gaat dan onder meer om de tijd die de docent beschikbaar heeft en de tijd die beschikbaar is om de leerstof te behandelen. Voor het vak Statistiek voor Natuurkunde zijn drie studiepunten ingeroosterd, wat neerkomt op een studielast van 84 uur. Het vak wordt in een maand tijd gegeven, wat overeenkomt met halftijdse besteding. Dit betekent dat de student in relatief weinig tijd veel stof moet verwerken. Dat beperkt onder andere het aantal colleges en het aantal huiswerkopgaven dat kan worden opgegeven.

➤ *Colleges*

Hoorcolleges zijn geschikt om op een gestructureerde manier de leerstof te behandelen. Het is mo-ge-lijk om de hele stof te behandelen, of om alleen de hoofdpunten en de structuur te behandelen en de studenten de details zelf te laten bestuderen. Het grote nadeel van een hoorcollege is dat daar slechts een beperkt gedeelte van het leerproces aan de orde komt: alleen de fase actieve begripsvorming uit de leercyclus van Kolb (zie Kallenberg, 2006) wordt gebruikt. Het volledige leren van de stof kan dus niet met alleen hoorcolleges worden bereikt. De student wordt hiermee bovendien niet getraind om de statistische methoden toe te passen, wat toch het belangrijkste leerdoel is. Hoorcolleges alleen zijn dus niet voldoende.

Werkcolleges zijn geschikt om onder begeleiding aan opgaven te werken. Hiermee wordt, in ter-men van de leercyclus van Kolb, gezorgd dat de studenten actief oefenen met de beschikbare theorie en concrete ervaringen opdoen. Ook kan de student hier de leerstof nog eens bestuderen en er vragen over stellen. Een demonstratie tijdens het werkcollege door de docent of door een andere student kan helpen om de fase van het observeren door te lopen. In werkcolleges is ook meer ruimte voor studen-ten met verschillende leerstijlen om zich de stof op hun eigen manier eigen te maken.

➤ *Toetsing*

Uiteindelijk is het wenselijk om te onderzoeken of de student de leerdoelen heeft bereikt. Voor een gedeelte kan dit direct worden getoetst, voor een gedeelte is dit moeilijker. Hierbij speelt het feit dat de toetsing moet leiden tot een objectief meetbare prestatie (een cijfer) een belangrijke beperkende rol.

De leerdoelen die op toepassing in een praktijkprobleem gericht zijn, kunnen gedeeltelijk worden getoetst door een fysische context te presenteren en daarover opgaven te geven. Hiervoor kunnen voorbeelden gebruikt worden van daadwerkelijke toepassingen van de statistiek in onderzoek van een van de vakgroepen. Om de opgaven geschikt te maken voor een toets, moet echter wel een duidelijke definitie van de context worden gegeven en is er maar beperkt ruimte om zaken aan de eigen interpretatie van de student over te laten. Dit is dus anders dan veel situaties die de student in de praktijk tegen zal komen, waarbij hij zelf extra vragen kan (en zal moeten) stellen over de aard van de verkregen metingen, extra onderzoek kan doen naar de betrouwbaarheid van het experiment en dergelijke. Dit aspect van de leerdoelen kan dus niet direct worden getoetst met behulp van opgaven. Eventueel is het wel mogelijk om het handelen in een praktijksituatie (experimenten op het practicum) te beoordelen op toepassing van de statistiek.

De leerdoelen betreffende het bereiken van een basisniveau aan kennis, op basis waarvan in een later stadium de kennis kan worden uitgebreid aan de hand van literatuur, kunnen moeilijker worden getoetst. In de tijd die voor het vak beschikbaar is, is het immers juist niet het doel dat studenten hun kennis buiten de leerstof gaan uitbreiden, en een opdracht van deze strekking kan dus niet worden gebruikt om te toetsen of de student het leerdoel heeft bereikt. Het is wel mogelijk om een noodzakelijke voorwaarde voor het bereiken van dit leerdoel te toetsen, namelijk het hebben van een voldoende begrip van de stof. Begrip kan getoetst worden door (theorie)opdrachten aan de student te geven die een bepaalde mate van begrip vereisen om te kunnen worden uitgevoerd.

In de toetsing moeten deze aspecten dus aan de orde komen: theorievragen waarbij wordt gevraagd om begrip van de leerstof, en toepassingsvragen waarbij de student de geleerde methodes moet toepassen op een concreet geschetst probleem.

► *Inleveropgaven*

In 2010 werden bij Statistiek voor Natuurkunde inleveropgaven gebruikt om de student te toetsen. Ik stel voor om dat zo te houden, omdat inleveropgaven naast de toetsfunctie ook, mits goed geconstrueerd, een leerfunctie kunnen hebben, en omdat een schriftelijk of mondeling tentamen niet goed aansluit op de leerdoelen. Ik licht dit hieronder toe.

Met opgaven die de student als huiswerk zelfstandig maakt en daarna schriftelijk inlevert, kan goed getoetst worden of de student de leerdoelen heeft bereikt. Hierin kunnen zowel theorievragen worden gesteld: de student moet bijvoorbeeld laten zien dat hij kan werken met de wiskundige notatie en hiermee een eenvoudige afleiding kan maken. De opgaven 1.1, 1.5, 1.6b en 2.1a in de bijlage zijn voorbeelden van theorievragen die gesteld kunnen worden. Daarnaast zijn ook toepassingsvragen goed mogelijk. Hierbij kan de student een praktische situatie worden voorgeschoteld, op basis waarvan hij een statistisch model moet opstellen en daarna een geschikte methode moet gebruiken om het gevraagde te berekenen.

Belangrijk is dat de opgaven zo geconstrueerd worden, dat er tijdens het maken van de opgaven nog iets te leren valt, dat de opgaven niet alleen maar gericht zijn op toetsing van al bekende begrippen en vaardigheden. Dit kan bijvoorbeeld gerealiseerd worden door in een opgave een nieuw begrip of een klein nieuw stukje theorie te presenteren, waarmee direct gewerkt moet worden; door in de opgave een schijnbare tegenstrijdigheid te verwerken, waar de student even over moet nadenken voor hij het probleem kan oplossen (bijvoorbeeld: a) Bepaal de maximum-likelihoodschatter, b) Bepaal de maximum-likelihoodschatting.); door te vragen naar het essentiële verschil tussen twee methoden, als men dat pas kan inzien als er even mee is geoefend. De opgaven moeten dus aanzetten tot nadenken en moeten eventueel uitnodigen tot het in meer detail bestuderen van de theorie die al in het college behandeld is. Gevaar van deze dubbele functie van de inleveropgaven is wel dat een student het kan opgeven als hij het probleem niet kan oplossen – hij heeft dan het leerdoel van die opgave niet bereikt.

Een bijkomend voordeel van inleveropgaven in deze vorm is dat de student met andere studenten kan overleggen. Dit kan het leereffect van de opgaven vergroten doordat de andere studenten met een ander perspectief naar het probleem kunnen kijken, en de student dus gedwongen wordt 'reflectief te observeren' wat anderen doen: een fase uit de leercyclus van Kolb die in andere aspecten van de onderwijsvorm niet veel wordt belicht.

Een tentamen in de gebruikelijke vorm is in dit geval een minder geschikte methode om te toetsen of de student de leerdoelen heeft bereikt. Dit komt vooral doordat de vorm van een tentamen niet compatibel is met de aard van de leerdoelen. Bij een tentamen wordt de beschikbare tijd en het beschikbare naslagmateriaal beperkt; die beperkingen zouden er juist niet moeten zijn als men de leerdoelen, die meer op toepassing in de praktijk gericht zijn, wil toetsen. Daar is het immers mogelijk om zonder tijdsdruk over een probleem na te denken en hulpmiddelen als literatuur, internet of een computerprogramma te gebruiken.

Als bijlage bij dit verslag heb ik de inleveropgaven opgenomen, zoals ik voorstel dat die bij Statistiek voor Natuurkunde worden opgegeven. Ik heb deze opgaven gedeeltelijk zelf opgesteld, en gedeeltelijk overgenomen van opgaven die de docent heeft aangedragen, met grotere of kleinere aanpassingen (de opgaven die ik zelf heb opgesteld zijn 1.3 t/m 1.6, 2.1, 2.2, 3.4, 3.5, 4.1 t/m 4.3; wel zij opgemerkt dat ik de context van 4.2 en 4.3 overgenomen heb van de docent, en dat veel van deze opgaven losjes gebaseerd zijn op het werk van anderen). De inleiding die boven de opgavenseries staat, is niet de mijne en ik onderschrijf de opmerking "dat je ook een beschrijving van de context en de vraagstelling moet geven" niet.

► *Conclusie vorm van het onderwijs*

Alle mogelijkheden afwegend lijkt een combinatie van de verschillende onderwijsvormen het meest geschikt voor het vak Statistiek voor Natuurkunde:

- In hoorcolleges wordt de theorie besproken. Hier kunnen ook wiskundige bewijzen en dergelijke, waar relevant, worden gepresenteerd. De theorie kan, waar het iets toevoegt voor het begrip, worden geïllustreerd met computerdemonstraties.
- In werkcolleges gaat de student aan de slag met theorieopgaven en oefent hij met de computervaardigheden, onder begeleiding van de docent.
- Met de inleveropgaven gaat de student echt aan de slag om de theorie toe te passen in concrete praktijkproblemen. Een groot deel van het leerproces vindt hier plaats. Naar aanleiding van deze opgaven wordt de student gemotiveerd om de theorie zelf te bestuderen uit boek of dictaat.

4.2 Statistiek op het natuurkundepracticum

Het statistiekonderwijs in het natuurkundepracticum is een vervolg op het vak Statistiek voor Natuurkunde. Het doel in het practicum, zoals uiteengezet in paragraaf 3, is niet om nieuwe stof te behandelen, maar om het begrip van de in Statistiek voor Natuurkunde behandelde stof te verdiepen en om te leren op een meer informele manier te communiceren over de statistiek, in aansluiting op wat van een fysicus verwacht wordt. Lineaire regressie is een voorbeeld van een onderwerp waarover de student informeel zal moeten leren communiceren met collega's en begeleiders. In Statistiek voor Natuurkunde zijn deze onderwerpen ook aan bod gekomen, en als de student de stof echt begrijpt zal hij er weinig moeite mee hebben om de theorie in eigen woorden te formuleren. Uiteindelijk zal de student in eigen woorden moeten uitleggen volgens welke procedure de regressieparameters worden geschat, welke invloed de ligging van de meetpunten heeft op de nauwkeurigheid van de schattingen, dat je een t-toets kunt gebruiken om te onderzoeken of de regressieparameters gelijk zijn aan een bepaalde (literatuur)waarde. Daarnaast moet de student ook begrijpen en kunnen uitleggen dat de precieze

standaarddeviatie van de meetfouten niet van tevoren bekend is, maar wel geschat kan worden, en dat het mogelijk is om als extra aanname aan het regressiemodel toe te voegen dat de werkelijke standaarddeviatie van de meetfouten gelijk is aan de schattingen. Deze extra aanname levert de mogelijkheid op om te toetsen of het lineaire regressiemodel past bij de metingen, zorgt voor een nauwkeuriger bepaling van de regressieparameters en scherpt dus ook de mogelijkheid tot vergelijking met de literatuurwaarde aan (men kan een Gauss-toets in plaats van een t-toets gebruiken). Ook dit moet de student kunnen bediscussiëren met collega's. Het oefenen van het informeel communiceren op de zojuist geschetste manier is voor de fysicus in opleiding een belangrijke stap naar het in een onderzoek gebruiken van de geleerde statistische methoden. Deze stap zou in het practicumonderwijs moeten worden ondersteund.

Om deze doelen te bereiken zal er in het practicum expliciet ruimte moeten worden gemaakt voor de statistiek. Doordat de sturing van het leerproces in het practicum voor een groot deel bij de student ligt en de opdracht die de student op het practicum krijgt zeer open is, is er het gevaar dat de student aan de statistiek te weinig aandacht besteedt. De student zal dus meer moeten worden gestuurd dan bij sommige andere leeractiviteiten op het natuurkundep practicum. Dit kan bijvoorbeeld vorm krijgen door te eisen dat de student in minstens één experiment (of zelfs twee) het onderzoek zo uitvoert, en zijn onderzoeksvraag zo kiest, dat hij een zinvolle statistische analyse op de experimentele gegevens kan uitvoeren. Op die manier heeft de student de mogelijkheid om te leren om in een zeer open praktijk situatie een statistische analyse toe te passen.

Direct gekoppeld aan het uitvoeren van de statistische analyse is de communicatie erover. Het is om twee redenen belangrijk om ook hieraan tijdens het practicumonderwijs aandacht te besteden. Ten eerste moet de student leren om te communiceren met collega's en begeleiders over de gebruikte statistische methoden. Van de bachelor wordt verwacht dat hij kan communiceren in een taal die minder formeel is dan de wiskundige beschrijving van de statistiek. Hij moet dus, net als op andere gebieden van de fysica, de wiskundige concepten ook in woorden kunnen uitdrukken, zonder dat de precisie van de discussie er al te veel onder lijdt. Dit is een kunst die moet worden geoefend en de student zal daarbij ook feedback moeten ontvangen. Ten tweede is het informeel communiceren over de statistiek ook van direct nut voor de ontwikkeling van het begrip van de statistische leerstof. Door erover te communiceren wordt de student immers gedwongen om de theorie in eigen woorden te herformuleren, en om dat te kunnen moet men de theorie goed begrijpen.

De student moet zowel mondeling als schriftelijk over de statistische analyse kunnen communiceren. Aan de mondelinge communicatie kan worden gewerkt tijdens de individuele gesprekken die de student heeft met de practicumbegeleider, en in de, meer gestructureerde, werkplanbespreking. Ik adviseer om bij beide gelegenheden tijdens het onderwijs inderdaad aandacht te besteden aan het informeel communiceren over de statistische methoden die geschikt zijn voor toepassing in het huidige experiment. In het verslag oefent de student in de schriftelijke communicatie. Ook daarbij zou de student moeten worden getraind, voornamelijk door terugkoppeling achteraf, op de communicatie over de statistiek.

De mondelinge en vooral de schriftelijke verslaglegging speelt ook een belangrijke rol in de toetsing of de student de leerdoelen bereikt heeft. De aspecten die zouden moeten worden beoordeeld zijn direct gerelateerd aan de leerdoelen. Voor het toetsen van het leerdoel van begripsvorming wordt de inhoudelijke kwaliteit van de discussie over de statistische analyse beoordeeld: is de analyse correct, is de analyse valide, dat wil zeggen toepasbaar in de voorliggende situatie, en is de analyse uitputtend, dat wil zeggen dat alle mogelijke analysemethoden die relevant zijn, ook daadwerkelijk gebruikt zijn. Het leerdoel van de communicatie wordt getoetst door te beoordelen of de verslaglegging genoeg

informeel is, dat wil zeggen dat niet te veel in wiskundige taal wordt geformuleerd, maar tegelijkertijd ook duidelijk genoeg is.

5 Evaluatie Statistiek voor Natuurkunde in 2011

In deze paragraaf geef ik een korte evaluatie van de cursus Statistiek voor Natuurkunde die in januari 2011 gegeven is, en maak ik een vergelijking met de cursus van januari 2010. Ik besteed ook aandacht aan hoe het advies dat ik in dit verslag beschrijf, in de cursus van 2011 is overgenomen. Historisch gezien vond mijn onderzoek plaats nadat de cursus van 2010 was gegeven, en (grotendeels) voordat de cursus van 2011 werd gegeven. Tijdens het onderzoek heb ik regelmatig contact gehad met de docent van Statistiek voor Natuurkunde, Marianne Jonker, en zij heeft een gedeelte van mijn voorstellen overgenomen en heeft ook gebruik gemaakt van de opgaven die ik heb gemaakt. Tijdens en na de cursus van 2011 heb ik met enkele studenten en met de docent gesproken over de cursus, de bevindingen uit die gesprekken heb ik hieronder samengevat. Bij het interpreteren van de opmerkingen hieronder, dient de lezer in gedachten te houden dat deze evaluatie niet op een gestructureerde of wetenschappelijk verdedigbare manier is uitgevoerd.

Belangrijkste om te vermelden is wat de docent in de cursus van 2011 anders heeft gedaan dan het advies dat ik hier neerleg. De docent heeft gekozen voor een meer wiskundig rigoureuze aanpak dan die ik adviseer. Dit blijkt onder andere uit de presentatie van de kansrekening. Het verschil is subtiel, de inhoud van de stof zou vrijwel hetzelfde zijn maar de benadering is anders en ik verwacht dat een minder wiskundige aanpak iets minder tijd zou kosten, waardoor mogelijk tijd overblijft om meer stof te behandelen. Anderzijds moet worden opgemerkt dat de docent de cursus minder wiskundig heeft gemaakt dan de cursus die zij normaliter aan een publiek van wiskundestudenten geeft.

De docent heeft een aantal onderwerpen niet behandeld, waarvan ik in dit verslag adviseer om ze wel te behandelen. Dit betreft de vier gebieden, die ik hier opsom.

- De lineaire regressie is alleen behandeld voor de situatie dat de meetfouten alle gelijke standaardafwijking hebben. De gewogen regressie, waar verschillende meetpunten een verschillend gewicht krijgen, is niet behandeld.
- Niet-lineaire regressie is niet aan bod gekomen.
- Het concept covariantie is mogelijk rakelings aan bod gekomen, dat weet ik niet zeker. Er is verder niets met de toepassing ervan gedaan: dat de geschatte parameters van een regressie gecorreleerd zijn is niet besproken en de interpretatie daarvan dus ook niet.
- Het bepalen van de onzekerheid in een maximum-likelihoodschatting, met behulp van de fisher-informatie, is niet behandeld. Dit vergt redelijk wat theorie als men dit goed wil begrijpen, dus is het de vraag of dit überhaupt deel van de cursus zou moeten uitmaken. De essentie van de methode zou echter in relatief weinig tijd kunnen worden behandeld, wat wel wenselijk is.

Een ander punt waarop de cursus van 2011 verschildte van 2010 betreft de opgegeven huiswerkopgaven. Het maken van opgaven is mijns inziens cruciaal voor het leerdoel van kunnen toepassen van de statistiek in concrete problemen. De opgavenseries in de bijlage heeft de docent met enkele inkortingen en aanpassingen overgenomen en opgegeven. Bij de cursus van 2010 waren er minder verplichte inleveropgaven dan in 2011, er waren toen wel meer niet-verplichte opgaven. De verschillende series inleveropgaven van 2011 waren langer dan het jaar ervoor, waardoor de studenten gedwongen werden meer met de stof bezig te zijn. De opgaven waren ook meer dan vorig jaar gericht op de toepassing, en er waren weinig theorievragen. Als voorbeeld noem ik serie 1. In 2010 bestond deze serie uit een redelijk eenvoudige theorievraag die op papier gemaakt moest worden, en twee computeropgaven waarin de wet van de grote aantallen en de centrale limietstelling werden gecontroleerd aan de hand

van een voorbeeld. De theorievraag is nu behouden (als opgave 1.1), de opgave over de wet van de grote aantallen is ingekort (1.2), de opgave over de centrale limietstelling is weggehaald, en verder zijn er in de eerste opgavenserie nog een aantal toegepaste kansrekeningvragen opgenomen (1.3 t/m 1.6).

Nog een verschil lag in het gebruik van het computerpakket R. Waar relevant heb ik geprobeerd in de opgavenseries R code te geven in plaats van de student er zelf op te laten komen. Het bleek dat dit vaak dezelfde hints betrof en dus heb ik een extra handleiding voor R gemaakt voor enkele veel voorkomende situaties. Ik pleit er echter, zoals eerder opgemerkt, voor om het onderwijs aan de hand van Mathematica te geven omdat de studenten met die omgeving al bekend zijn. (Ook daarbij zou dan een korte handleiding gegeven moeten worden voor de statistische functies.)

In de gesprekken met studenten tijdens de cursus in 2011, die ik op informele basis heb gevoerd, bleek dat zij het nut van het vak over het algemeen wel inzagen. Zij gaven aan moeite te hebben met het gebruik van R en met de verslaglegging van de gemaakte opgaven, die te veel tijd zou kosten. De cijfers die de studenten voor de inleveropgaven haalden liepen uiteen, maar uit de spreiding van de cijfers en de opgegeven tijdsbesteding (gemiddeld zo'n 10 uur per opgavenserie), heb ik de indruk gekregen dat de opgaven niet te makkelijk of te moeilijk waren. Direct aansluitend op de cursus Statistiek voor Natuurkunde begon het Natuurkundepracticum 2. Enkele studenten bleken spontaan aan de slag te gaan met de statistiek in hun experimenten.

Het is moeilijk om een vergelijking te maken met de ervaring van studenten van de cursus van 2010. Reden is dat de studentengroepen van 2010 en 2011 gemiddeld van elkaar blijken te verschillen en dus naast de verandering van de inhoud en vorm van de cursus ook andere factoren invloed hebben op het leerresultaat van de groep van 2010.

6 Afsluiting

In dit verslag heb ik betoogd dat een wiskundige aanpak van het statistiekonderwijs essentieel is om de student en experimentator te dwingen de aannames van de gebruikte statistische analysemethoden expliciet te maken, wat belangrijk is voor het beoordelen van de validiteit van de analyse. Tegelijk moet het onderwijs voldoende gericht zijn op de toepassing en niet verzanden in te veel wiskundige details. In de opgaven die ik heb ontwikkeld heb ik geprobeerd hierin een evenwicht te zoeken. Ook voor de rest van het statistiekonderwijs heb ik geschetst hoe dat evenwicht er mijns inziens uit zou moeten zien en hoe het bereikt kan worden.

Met veel plezier heb ik aan dit project gewerkt. Ik hoop dat mijn werk, zoals in dit verslag gerapporteerd, ten positieve bijdraagt aan het effect van het statistiekonderwijs aan bachelorstudenten natuurkunde aan de VU.

Referenties

- College voor Examens, *Examenprogramma Wiskunde B vwo*, 2010.
- A.J. Kallenberg, L. van der Grijspaarde, A. ter Braak, C.J. van Horzen, *Leren (en) Doceren in het hoger onderwijs*. Lemma (Utrecht), 2006.
- G.J.L. Wuite, *Natuurkunde in het licht van leven en energie. Vernieuwing van het curriculum van de bacheloropleiding natuurkunde aan de VU*. Rapport van de Curriculum commissie, 2008.

Statistiek voor Natuurkunde

Opgavenserie 1: Kansrekening

Inleveren: nog te beslissen

Afspraken

Overleg is toegestaan, maar iedereen levert zijn eigen werk in. Overschrijven of kopiëren van antwoorden is niet toegestaan en zal worden bestraft.

Werk de opdrachten uit en schrijf een overzichtelijk verslag. Het verslag moet te begrijpen zijn voor iemand die de opdrachten niet gelezen heeft. Dat betekent dat je ook een beschrijving van de context en de vraagstelling moet geven. De R-code komt in een bijlage. Ik wil een uitgeprinte versie; inleveren per e-mail is niet mogelijk!

Opgave 1.1 *Elektrische schakeling*

In deze opgave ga je enkele kansen berekenen. Gebruik hiervoor een nette notatie, en geef duidelijk aan welke rekenregels je gebruikt.

- (a) Beschouw een systeem met 10 in serie geschakelde componenten. De componenten kunnen onafhankelijk van elkaar kapot gaan. Het systeem werkt alleen als alle componenten werken. De kans dat een gegeven component werkt is p . Bereken de kans dat het systeem werkt.
- (b) Beschouw nu een systeem met 10 parallel geschakelde componenten. De kans dat een component werkt is opnieuw p en de componenten kunnen onafhankelijk van elkaar kapot gaan. Het systeem werkt als één of meer van de componenten werkt. Bereken de kans dat het systeem werkt.

TODO: HIER PLAATJE INVOEGEN

- (c) Beschouw nu het systeem dat hierboven getekend is. Het systeem werkt als component 1 werkt en tevens alleen component 2, alleen component 3, of zowel component 2 als component 3 werkt. De kans dat een gegeven component werkt is opnieuw p , onafhankelijk van de andere componenten. Bereken de kans dat het systeem werkt, uitgedrukt in p . Wat is deze waarde als $p = 0.8$?

Opgave 1.2 *Wet van de grote aantallen*

Als een fysicus middels een experiment herhaaldelijk de waarde van een constante bepaalt, vindt hij niet steeds dezelfde waarde. Een voorbeeld hiervan is een experiment om de lichtsnelheid te bepalen. Veronderstel dat we een proefopstelling hebben waarmee we de lichtsnelheid kunnen meten. Bij elke meting zullen we een kleine meetfout maken. Dit leidt tot een waarneming die afwijkt van de werkelijke lichtsnelheid ($299.7925 \cdot 10^6$ m/s). We veronderstellen dat de gemaakte meetfouten realisaties zijn van onderling onafhankelijke en normaal verdeelde stochastische variabelen met verwachting 0 en standaarddeviatie $0.080 \cdot 10^6$. De metingen van de lichtsnelheid zijn dan realisaties van onderling onafhankelijk normaal verdeelde stochastische variabelen met verwachting $299.7925 \cdot 10^6$ m/s (de lichtsnelheid) en standaarddeviatie $0.080 \cdot 10^6$ m/s (ga na). Voor het gemak in de notatie vergeten we vanaf nu de factor 10^6 .

In deze opgave bootsen we een experiment na waarin de lichtsnelheid herhaaldelijk gemeten wordt. De lichtsnelheid schatten we door de gemeten snelheden te middelen. We onderzoeken het effect van het aantal waarnemingen op onze schatting.

- (a) Genereer een reeks van 500 metingen van de lichtsnelheid door een steekproef ter grootte 500 uit de normale verdeling met verwachting 299.7925 en standaardafwijking 0.080 te nemen. Bereken de gemiddeld gemeten lichtsnelheid op basis van de eerste 1, 2, 3, ... metingen. Plot deze gemiddelden tegen het aantal gedane metingen.

R-tips: Een steekproef uit de normale verdeling kan worden gegenereerd met het R-commando `rnorm`. Gebruik `help(rnorm)` om uitleg te krijgen over dit commando. Bekijk ook het document *Enkele vaardigheden in R*. Het lopend gemiddelde over een vector kun je bijvoorbeeld krijgen door de functie `cumsum` te gebruiken, en het resultaat daarvan te delen door een vector `1:n`.

- (b) Leg uit wat je ziet en formuleer je conclusie.

Bij een simulatiestudie als deze, die gebaseerd is op (pseudo)random getallen, is het verstandig om de simulatie te herhalen met andere waarden voor de random getallen. Zo kun je meer zekerheid krijgen dat de geobserveerde effecten niet zijn verdraaid doordat je toevallig net de ‘verkeerde’ random getallen hebt gekregen.

- (c) Herhaal opgave (a) vijf malen, en beschrijf hoe je resultaten verschillen tussen de verschillende runs.

We onderzoeken nu het effect van de gedane aannames.

- (d) Varieer de standaardafwijking en beschrijf en verklaar welk effect dit heeft op je resultaat. Laat plaatjes zien. (Maak ook hier verschillende runs van de simulatie.)

Opgave 1.3 Meetfouten

Bij het meten van de lengte van een tafel met een meetlint, zijn er veel factoren die de meting vertroebelen: optische fouten, de ijking van het meetlint, het bewegen van je hand, afrondingsfouten, enzovoort. De totale fout in de meting van de lengte is een optelling van deze effecten. Het blijkt dat meetfouten vaak normaal verdeeld zijn; omdat deze bestaan uit de optelling van onafhankelijke vertroebelende factoren, hangt dit samen met de Centrale Limietstelling. (Ga dit na.) Als er geen systematische fouten zijn, is de verwachtingswaarde van de totale meetfout gelijk aan 0.

In deze opgave bestuderen we de meetfout in de gemeten lengte van een tafel. We zullen veronderstellen dat deze normaal verdeeld is met verwachting 0 en variantie σ^2 . In opgaven (a) t/m (d) zullen we $\sigma = 2$ mm nemen.

Maak bij deze opgave gebruik van de R-functies `pnorm` en `qnorm`. (N.B.: Zelf integralen uitrekenen is niet de bedoeling.)

- (a) Bereken de kans dat de meetfout ligt tussen $-\sigma$ en $+\sigma$. (Doe dit dus voor $\sigma = 2$ mm.)
- (b) Bereken de kans dat de meetfout ligt tussen 0 en $+\sigma$. Vergelijk met vraag (a) en verklaar wat je ziet.
- (c) Bereken, zonder opnieuw gebruik te maken van R, de kans dat de meting meer dan σ afwijkt van de werkelijke lengte van de tafel.

- (d) Bereken voor welke waarde van x de kans dat de meting minder dan x van de werkelijke waarde afwijkt, gelijk is aan 0.95.
- (e) Welke van je antwoorden veranderen als σ verandert? Druk in die gevallen het antwoord uit in σ .

Opgave 1.4 Muntworpen

Iemand geeft je een munt en beweert dat deze zuiver is (bij een zuivere munt is bij een worp de kans op kop gelijk aan $1/2$). Om die bewering te onderzoeken, werp je de munt 1000 keer. Het blijkt dat je 538 keer kop hebt gekregen.

- (a) Bereken exact (met behulp van R) de kans dat je 538 keer of vaker kop gooit, als je dit experiment zou herhalen met een munt die zeker zuiver is. Definieer hierbij de stochastische variabele die je gebruikt, welke verdeling deze heeft, en welke aannames over het experiment je hierbij doet.
- (b) Gebruik de Centrale Limietstelling om de bij (a) gevraagde kans te benaderen. Beschrijf duidelijk welke stappen je maakt. Geef commentaar op de kwaliteit van de benadering.
- (c) Leg uit of je er vertrouwen in hebt dat de munt die je hebt gekregen, inderdaad zuiver is.
- (d) Leg uit waarom je bij vraag (b) de Centrale Limietstelling kon gebruiken. Dat wil zeggen: leg uit of aan de voorwaarden van de CLS voldaan is, en waarom je al dan niet kunt verwachten dat de benadering die de CLS geeft daadwerkelijk een *goede* benadering is.

Opgave 1.5 Levensduur

Beschouw opnieuw een elektrisch systeem. In dit systeem bevindt zich een kritiek onderdeel: zonder dit onderdeel functioneert het systeem niet. Het onderdeel heeft echter een beperkte levensduur, die exponentieel verdeeld is met parameter μ . Bekend is dan dat de verwachtingswaarde van de levensduur van het systeem gelijk is aan $1/\mu$. (Ga dit voor jezelf na.)

Het is erg duur om het onderdeel te laten vervangen. Daarom plaatst men in het systeem twee exemplaren van dit onderdeel. De levensduren van beide exemplaren gaan tellen vanaf het begin – het is dus niet zo dat het tweede exemplaar pas wordt ingeschakeld als de eerste kapot gaat. De levensduren van beide exemplaren zijn onafhankelijk van elkaar, en zijn nog steeds exponentieel verdeeld met parameter μ . Het systeem stopt met functioneren als beide exemplaren het begeven hebben. (Veronderstel dat de rest van het systeem niet kapot kan.) We zijn geïnteresseerd in de totale levensduur van het systeem; we noemen deze X .

- (a) Bepaal de kansverdelingsfunctie $x \mapsto P(X \leq x)$.
(*Hint*: Bedenk dat de levensduur van het totale systeem gelijk is aan het maximum van de levensduren van de twee exemplaren van het onderdeel.)
- (b) Bereken de verwachtingswaarde van X in termen van μ . Vergelijk de uitkomst met de verwachtingswaarde van de levensduur van het systeem met slechts één exemplaar van het onderdeel.

Opgave 1.6 Interpretatie van kansverdelingen

In deze opgave leer je hoe de verschillende standaard-kansverdelingen met elkaar samenhangen.

Binomiale en poissonverdeling

De binomiale verdeling en de poissonverdeling hangen nauw met elkaar samen: als n heel groot wordt en p heel klein, gaat de binomiale verdeling steeds meer lijken op de poissonverdeling. Als toepassing hiervan kunnen we nemen een meting gedurende vijf minuten aan een radioactief monster met een grote halfwaardetijd (dat wil zeggen, zeer veel groter dan vijf minuten). Het aantal atoomkernen in een dergelijk monster dat in een periode van vijf minuten vervalst, hangt onder andere af van het totaal aantal actieve kernen (n), en de kans dat een gegeven kern in de meetperiode van vijf minuten vervalst (p). Het aantal vervallen kernen is dan binomiaal verdeeld met parameters n en p . De parameter n is zeer groot, en door de aanname van een grote halfwaardetijd is p zeer klein – het blijkt dat dan het aantal vervallen kernen ook goed kan worden gemodelleerd met een poissonverdeling met parameter $\lambda = np$.

- (a) Laat zien, dat de kansdichtheidsfunctie van de binomiale verdeling met parameters n en p steeds meer gaat lijken op die van de poissonverdeling met parameter $\lambda = np$, als $n \rightarrow \infty$ en $p \rightarrow 0$. Doe dit door de kansdichtheidsfuncties te plotten (gebruik `plot` en `points`), voor een vaste waarde van λ en steeds grotere n . Het is niet de bedoeling dat je dit met simulatie doet.

R-tips: Als je meerdere plots naast elkaar wilt laten zien, kun je het R-commando `par(mfrow=c(m,n))` gebruiken. (Zie de Handleiding R.) Er wordt dan een m bij n raster gemaakt; de volgende plots worden dan naast en/of onder elkaar weergegeven. Je kunt een titel bij een plot zetten door het argument `main=` op te geven; de assen kun je labelen met de argumenten `xlab` en `ylab`.

Je kunt het samenvallen van de binomiale en de poissonverdeling voor grote n ook formeel laten zien.

- (b) Stel X is binomiaal(n, p) verdeeld, en Y is poisson verdeeld met parameter $\lambda := np$. Laat zien, als $n \rightarrow \infty$ (en dus $p = \frac{\lambda}{n} \rightarrow 0$ zodat λ constant blijft), dat $P(X = k) \rightarrow P(Y = k)$, voor $k = 0, 1, \dots, n$.

Gebruik de standaardlimiet $\lim_{n \rightarrow \infty} (1 + \frac{x}{n})^n = e^x$, en de regel dat “de limiet van een eindig product gelijk is aan het product van de limieten”.

Som van normale verdelingen

Niet alle appels zijn even zwaar. Veronderstel dat de massa van een appel normaal verdeeld is met verwachting $\mu = 220\text{g}$ en standaardafwijking $\sigma = 30\text{g}$. We gaan 8 willekeurig te kiezen appels kopen, en vragen ons af wat de verdeling is van de totale massa van deze 8 appels. We zijn dus geïnteresseerd in de verdeling van een som van normaal verdeelde grootheden.

- (c) Stel X_i , $i = 1, \dots, n$ zijn onderling onafhankelijk en normaal verdeeld: $X_i \sim N(\mu, \sigma^2)$.
- (i) Laat door middel van simulatie in R zien, dat de som $Z = \sum_{i=1}^n X_i$ zelf ook normaal verdeeld is. Doe dit voor het getallenvoorbeeld dat hierboven gegeven is.
- (ii) Bepaal, mede op basis van de resultaten van je simulatie, ook de parameters μ_{som} en σ_{som} van de verdeling van Z , en druk deze uit in μ en σ van de verdeling van de individuele X_i .

Hints: Gebruik de paragraaf over simulatie uit het document *Enkele vaardigheden in R*. Genereer een aantal realisaties van Z en teken hiervan een histogram; trek op basis daarvan je conclusie.

De som van normaal verdeelde grootheden is dus weer normaal verdeeld. Op grond van de Centrale Limietstelling mocht je inderdaad verwachten dat dit bij benadering waar is. Echter, in het geval van normaal verdeelde X_i is de som zelfs exact normaal verdeeld, dus niet alleen bij benadering.

Exponentiële verdeling en Poisson verdeling

Op het natuurkundepracticum worden vaak oscilloscopen gebruikt. Zo af en toe gaat er een oscilloscoop kapot. Men is nu bezig met de planning voor het komende collegejaar. Veronderstel dat op 1 september de oscilloscopen allemaal functioneren. Het duurt dan een tijdje totdat de eerste oscilloscoop van het jaar kapot gaat. Deze tijd is stochastisch en noemen we Y_1 . De kapotte oscilloscoop wordt onmiddellijk vervangen en daarna duurt het een tijd Y_2 totdat er weer eentje kapot gaat, enzovoort. Tijdsduur Y_i is dus de tijd tussen het kapot gaan van de $i - 1$ -ste en i -de oscilloscoop. Dit gaat zo door tot 31 augustus van het volgende jaar. (Voor het gemak veronderstellen we dat de oscilloscopen het hele jaar door even intensief worden gebruikt, ook in de zomermaanden.) Veronderstel dat $Y_1, Y_2, \dots$, onderling onafhankelijk en exponentieel verdeeld zijn met intensiteitsparameter $\theta = 6$ per jaar. Het aantal oscilloscopen dat in dit jaar kapot gaat noemen we X .

- (d) Laat met behulp van simulatie (gebruik `rexp`) zien dat X een poissonverdeling heeft met parameter $\theta = 6$. Gebruik de informatie in het document *Enkele vaardigheden in R*.

(*Hints:* Simuleer realisaties van X en teken een histogram. Gebruik bij het histogram het argument `breaks=(min(xvector)-0.5):(max(xvector)+0.5)`), waarbij `xvector` de gesimuleerde resultaten voor X bevat. Bedenk zelf waarom deze specificatie van `breaks` handig is. Teken in hetzelfde plaatje de kansdichtheidsfunctie (met `lines` of `points`) van de poisson verdeling verdeling met parameter $\theta = 6$.)

Vat voor jezelf je bevindingen van opgave 1.6 nog eens samen (dit hoeft je niet in te leveren).

Statistiek voor Natuurkunde

Opgavenserie 2: Schatten en lineaire regressie

Inleveren: nog te beslissen

Afspraken

Overleg is toegestaan, maar iedereen levert zijn eigen werk in. Overschrijven of kopiëren van antwoorden is niet toegestaan en zal worden bestraft.

Werk de opdrachten uit en schrijf een overzichtelijk verslag. Het verslag moet te begrijpen zijn voor iemand die de opdrachten niet gelezen heeft. Dat betekent dat je ook een beschrijving van de context en de vraagstelling moet geven. De R-code komt in een bijlage. Ik wil een uitgeprinte versie; inleveren per e-mail is niet mogelijk!

Opgave 2.1 *Eiwerpen*

Met een vriend gooi je een rauw ei over, totdat het ei breekt. Er is geen verschil tussen een worp of een vangst van jou en die van je vriend. Je vraagt je af wat de kans is dat het ei breekt tijdens een gegeven worp (inclusief vangen). Om dit te onderzoeken gebruik je een doos met 10 eieren die allemaal even stevig zijn. Je telt voor elk ei het aantal keer dat je het ei kunt overgooien voordat het breekt. Hier zijn de resultaten:

3 2 1 2 1 2 4 7 1 8

Om de breekkans te schatten, stel je het volgende statistisch model op. Voor elk ei zijn de worpen onafhankelijk van elkaar en bij elke worp is de kans dat het ei breekt gelijk aan p . Definieer X_i ($i = 1, \dots, 10$) als het aantal worpen totdat het i -de ei breekt. Zoals je zelf kunt nagaan, geldt dan voor elke i :

$$P(X_i = k) = (1 - p)^{k-1} p. \quad (k = 1, 2, \dots)$$

- (a) Bepaal de maximum likelihood schatter voor p . Beschrijf netjes hoe je dit doet.
- (b) Bepaal de maximum likelihood schatting voor p .

Opgave 2.2 *Machtswetverdeling*

In de bergen vinden regelmatig sneeuwlawines plaats, grote maar ook kleine. De grootte van een lawine volgt bij benadering een machtswetverdeling: de kansdichtheid p_α wordt voor zekere $\alpha > 1$ gegeven door

$$p_\alpha(s) = \begin{cases} C s^{-\alpha} & \text{voor } s \geq 1, \\ 0 & \text{anders} \end{cases} \quad (1)$$

Hierin is C een nog onbekende constante ($C > 0$).

- (a) Bepaal C , zodanig dat p_α een kansdichtheid is.

We zijn geïnteresseerd in het bepalen van de onbekende parameter α .

In het bestand `lawines.dat` op de website bij dit vak, bevindt zich een serie van $n = 732$ (gesimuleerde) metingen van lawinegroottes, opgegeven in willekeurige eenheden. We vatten deze metingen op als realisaties van $X_1, X_2, \dots, X_n$, waarbij $X_1, X_2, \dots, X_n$ een steekproef vormt uit de machtswetverdeling met dichtheid p_α .

- (b) Leid een uitdrukking af voor de maximum-likelihoodschatter voor α . Bereken vervolgens de maximum likelihoodschatting voor α voor de gegeven dataset.

R-tip: Om de gegevens in R in te lezen, download je eerst het bestand van de website naar je computer (bijvoorbeeld naar de map `C:\My Documents\StatVoorN\`). Daarna geef je het commando `x=scan("C:/My Documents/StatVoorN/lawines.dat")`. Let op de aanhalingstekens en gebruik `/` in plaats van `\`.

Opgave 2.3 Hoekverdeling

Antimuonen, die werden ontdekt in de kosmische straling die onze dampkring binnenkomt, kunnen vervallen in een positron (positief geladen antideeltje van het elektron), een elektron-neutrino en een muon-antineutrino:

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu.$$

In een detector wordt een bundel antimuonen gecreëerd en men bestudeert het bovenstaande proces. Men meet de hoek ϕ waaronder het positron wordt uitgezonden (ten opzichte van de richting waarin de antimuonen bewegen). Er geldt $0 \leq \phi \leq \pi$. De hoek ϕ is niet altijd hetzelfde en heeft een verdeling met dichtheidsfunctie

$$f_\alpha(x) = \frac{1 + \alpha x}{2} \quad -1 \leq x \leq 1 \text{ en } -1 \leq \alpha \leq 1$$

met $x = \cos \phi$. De parameter α is gerelateerd aan de magnetische polarisatie van de bundel antimuonen. Het standaardmodel van de subatomaire fysica voorspelt dat bij een volledig gepolariseerde bundel α gelijk is aan $1/3$. In deze opgave zullen we α als onbekende parameter beschouwen en schatten op basis van onze waarnemingen.

We onderzoeken twee manieren om de onbekende parameter α te schatten: de momentenschatter en de maximum likelihood schatter.

- (a) Veronderstel dat we waarnemingen $X_1, \dots, X_n$ hebben. Leid de loglikelihood vergelijking af voor de maximum likelihood schatter van α . Laat zien dat er geen expliciete uitdrukking voor de maximum likelihood schatter voor α bestaat.

Dat er geen expliciete uitdrukking voor de schatter gevonden kan worden, betekent niet dat de maximum likelihood schatter niet bestaat. Om toch de maximum likelihood schatting te kunnen uitrekenen gaan we de (log) likelihood numeriek maximaliseren. Dit kan met de R-functie `optim`. Bestudeer de documentatie over deze R-functie die je krijgt door `help(optim)` in te typen in je R-commando-window.

Als argument van de functie moet een startwaarde meegegeven worden. Deze waarde nemen we gelijk aan de zogenaamde momentenschatting.

- (b) Veronderstel dat $X_1, \dots, X_n$ een steekproef is uit de hoekverdeling, de verdeling met dichtheid f_α . Bereken de verwachting $E(X_i)$ en laat zien dat $\alpha = 3 E(X_i)$.

Volgens de wet van de grote aantallen convergeert het gemiddelde $\bar{X}$ (in kans) naar zijn verwachting $E(X_i)$. Dat betekent dat als het aantal waarnemingen niet te klein is, $\bar{X}$ dicht bij $E(X_i)$ zal liggen (zie opgave 1.2). Door $E(X_i)$ in bovenstaande uitdrukking voor α te vervangen door $\bar{X}$ krijgen we de zogenaamde momentenschatter van α : $\tilde{\alpha} = 3\bar{X}$.

We gaan nu gesimuleerde data trekken uit de hoekverdeling met parameter $\alpha = 0.3$.

- (c) Simuleer 1000 waarnemingen voor $x = \cos \phi$ uit de hoekverdeling met parameter $\alpha = 0.3$. Dit kan als volgt worden gedaan:

```
alpha = 0.3
y = runif(1000)
```

```
x = (-1+sqrt(1-2*alpha*(1-alpha/2-2*y)))/alpha
```

(Als je benieuwd bent wat hier gebeurt, kijk dan eens op Wikipedia bij *Inverse Transform Sampling*.)

Op basis van de gesimuleerde waarnemingen gaan we α schatten met behulp van de momentenschatter en de maximum likelihood schatter. Als de schatters nauwkeurig zijn, zullen we waarden in de buurt van 0.3 vinden; immers dit was de waarde van α waarbij de data werden gesimuleerd.

- (d) Bereken het gemiddelde van de 1000 gesimuleerde waarnemingen en bepaal vervolgens de momentenschatting voor α . Ligt de schatting in de buurt van 0.3?
- (e) Bereken vervolgens de maximum likelihood schatting voor α met de R-functie `optim`. Neem als beginwaarde de momentschatting. Ligt de maximum-likelihood-waarde in de buurt van 0.3?

Hier volgt de R-code die je zou kunnen gebruiken:

```
alpha = 0.3
dmuon = function(x,alpha) (1+alpha*x)/2

muonlik = function(alpha,x) {
  logl = sum(log(dmuon(x,alpha)))
  return(-logl) }

y = runif(1000)
x = (-1+sqrt(1-2*alpha*(1-alpha/2-2*y)))/alpha
alphamom = 3*mean(x)
zz = optim(alphamom,muonlik,x=x,method="L-BFGS-B",lower=-1,upper=1)
alphamle = zz$par
```

- (f) Herhaal de vragen (c), (d) en (e) 1000 maal en sla alle schattingen op. Dit kun je in R doen door een `for`-lus te gebruiken. Maak eerst twee vectoren ter lengte 1000; één voor de momentenschattingen en één voor de maximum likelihood schattingen. Bij de i -de simulatie zet je de gevonden waarden op de i -de plek in je vectoren.
- (g) Teken histogrammen van de 1000 momentenschattingen voor α . Doe hetzelfde voor de 1000 maximum likelihood schattingen. Welke verdeling zouden de moment- en maximum likelihood schatters kunnen hebben? Schat de parameters van de twee verdelingen. Schaal de histogrammen met `prob=T` en teken de kansdichtheidsfuncties in hetzelfde plaatje. Op basis van welke limietstelling uit de kansrekening kom je ook tot deze verdeling voor de momentenschatter?
- (h) Schat de verwachting en de variantie van de gevonden verdelingen voor de momentenschatter en de maximum likelihood schatter (deze heb je, als het goed is in de vorige opdracht bepaald). Welke van de twee schatters vind je in dit geval beter? Leg uit.

Opgave 2.4 Lineaire regressie

In de data-set “Zoon.txt” en “Dochter.txt” op de website van dit vak staan in de eerste kolom de lengtes (in centimeter) van 100 volwassen zonen en 100 volwassen dochters, respectievelijk, en de tweede kolom de gemiddelde lengte van hun ouders. Lees de data in zoals staat beschreven in de tweede opgave van deze serie. In deze opgave gaan we regressie-modellen schatten om de eindlengte voor jongens en meisjes

te voorspellen op basis van de gemiddelde lengte van hun ouders. De eindlengte is de maximale lengte die een persoon haalt; de lengte als een persoon is uitgegroeid.

- (a) Teken een scatter-plot waarin de lengte van de zoon wordt uitgezet tegen de gemiddelde lengte van zijn ouders. Doe hetzelfde voor de dochters. Zet de scatter-plots in je verslag en beschrijf wat je ziet. Is een lineair regressie model zinvol? Leg uit?
- (b) Stel voor beide data-sets een lineair regressie model op en schat voor beide modellen de regressie-parameters met de kleinste kwadraten schatters. Gebruik hiertoe in R de functie `lm`; in het document *Enkele vaardigheden in R* vind je informatie over het gebruik van deze functie.
- (c) Onderzoek de geschatte meetfouten $(y_i - \hat{\alpha} - \hat{\beta}x_i, i = 1, \dots, n)$. Is het redelijk te veronderstellen dat de meetfouten normaal verdeeld zijn? Met verwachting 0?
- (d) Bepaal nu één lineair regressie model voor de gecombineerde data-set voor jongens en meisjes. Je moet nu dus een variabele opnemen die aangeeft wat het geslacht van het kind is. Hoe kies je deze? Leg uit.
- (e) Artsen gebruiken de formule “verwachte eindlengte = $4.5 + \text{gemiddelde lengte ouders} + 6.5$ ” voor jongens en “verwachte eindlengte = $4.5 + \text{gemiddelde lengte ouders} - 6.5$ ” voor meisjes. Schrijf deze voorspelde eindlengte in de vorm van een lineair regressie model en vergelijk de waarden van de regressiecoëfficiënten met de waarden die jij hebt gevonden. Verschillen de waarden veel?
- (f) Vergelijk nu de voorspellingen van dit model met de voorspellingen van jouw geschatte model. Past het model dat gebruikt wordt door artsen ook goed bij de data? Leg uit.

Statistiek voor Natuurkunde

Opgavenserie 3: Toetsen en betrouwbaarheidsintervallen

Inleveren: nog te beslissen

Afspraken

Overleg is toegestaan, maar iedereen levert zijn eigen werk in. Overschrijven of kopiëren van antwoorden is niet toegestaan en zal worden bestraft.

Werk de opdrachten uit en schrijf een overzichtelijk verslag. Het verslag moet te begrijpen zijn voor iemand die de opdrachten niet gelezen heeft. Dat betekent dat je ook een beschrijving van de context en de vraagstelling moet geven. De R-code komt in een bijlage. Ik wil een uitgeprinte versie; inleveren per e-mail is niet mogelijk!

Geef bij elke toets die je uitvoert tenminste:

1. de nulhypothese en de alternatieve hypothese;
2. de toetsingsgrootheid;
3. de aannamen die je maakt (zijn de aannamen acceptabel?);
4. de overschrijdingskans of het kritieke gebied;
5. je conclusie.

Neem de onbetrouwbaarheid van elke toets gelijk aan 0,05.

Opgave 3.1 *Ratten*

De datavectoren `ozon` en `control` op de website <http://www.math.vu.nl/sto/onderwijs/statistiekvoorNatuurkunde> bevatten gegevens van een onderzoek naar de effecten van ozon (een bestandsdeel van smog) op gewichtstoename van ratten. De vector `ozon` bevat de gewichtstoename (in gram) van 22 ratten in 7 dagen tijd. Gedurende deze 7 dagen verbleven de ratten in een ozon-rijke omgeving. De vector `control` bevat de gewichtstoename van de controlegroep: van 23 ratten die 7 dagen in een ozon-vrije ruimte zijn gehouden. In beide groepen waren alle ratten 70 dagen oud bij het begin van de studie. In deze opdracht kun je gebruikmaken van de functies `t.test` en `wilcox.test`. Bestudeer voordat je de functies gebruikt de uitleg van deze functies met de commando's `help(t.test)` en `help(wilcox.test)`.

- (a) Geef van beide datasets een samenvatting (bepaal gemiddelde, mediaan, variantie, standaardafwijking, minimum, maximum). Onderzoek de verdelingen van de gewichtstoenames van de ratten in de controlegroep en in de ozongroep afzonderlijk. Teken histogrammen. Geef plaatjes en beschrijf wat je ziet.
- (b) Toets met een t-toets of 70 dagen oude ratten in een ozon-rijke ruimte toenemen in gewicht. Doe hetzelfde voor ratten in een ozon-vrije ruimte.
- (c) Toets met een t-toets of 70 dagen oude ratten in een ozon-rijke omgeving minder in gewicht toenemen dan 70 dagen oude ratten in een ozon-vrije omgeving.
- (d) In de controlegroep zit een rat met een gewichtstoename van $-16,9$ gram. Deze waarde wijkt erg af van de overige waarden. Mogelijk is dit een foute waarde. We verwijderen de waarde om te onderzoeken wat haar effect is op de toetsingsuitslagen. Voer de toetsen in de onderdelen b en c opnieuw uit. Wat je bevindingen samen in een paar regels. Zijn de verschillen groot?

- (e) Toets opnieuw of ratten in een ozon-vrije omgeving meer in gewicht toenemen dan ratten in een ozon-rijke omgeving (met en zonder de “foute” waarde) met de Wilcoxon toets voor twee steekproeven (ook wel de Mann-Whitney toets genoemd). Gebruik het commando `wilcox.test`, maar lees eerst de uitleg die je vindt met het commando `help(wilcox.test)` en voorbeeld 5.33 in het dictaat.
- (f) Vergelijk de resultaten van de twee twee-steekproeven toetsen. Naar welke toets gaat je voorkeur uit? Waarom?

Opgave 3.2 *Schatten en betrouwbaarheidsintervallen*

Als een fysicus middels een experiment herhaaldelijk de waarde van een constante bepaalt, vindt hij niet steeds dezelfde waarde. Een voorbeeld hiervan zijn de metingen van de lichtsnelheid gevonden door Michelson in 1882 (zie de dataset “lichtsnelheid.txt” op de homepage van het vak <http://www.math.vu.nl/sto/onderwijs/statistiekvoorNatuurkunde>). De getallen, vermenigvuldigd met 10^6 , zijn 100 metingen van de lichtsnelheid in meters per seconde. Tegenwoordig weten we dat de lichtsnelheid gelijk is aan $299.7925 \cdot 10^6$ meter per seconde. We veronderstellen dat de metingen realisaties zijn van onderling onafhankelijke en identiek verdeelde stochastische variabelen $X_1, \dots, X_{100}$. De meetfouten die door Michelson gemaakt worden, zijn dan realisaties van $X_1 - \mu, \dots, X_{100} - \mu$ met $\mu = 299.7925 \cdot 10^6$, de lichtsnelheid. Trek van elk van de 100 waarnemingen in de data-set 299.7925 af; de nu gevonden waarnemingen zijn de gemaakte meetfouten.

- (a) Teken een histogram (`hist`) van de meetfouten, en ga met een histogram na of het redelijk is te veronderstellen dat de meetfouten uit een normale verdeling afkomstig zijn. Schat de verwachting en variantie van deze normale verdeling met hun maximum likelihood schatters.
- (b) Bepaal een 95% betrouwbaarheidsinterval voor de verwachting van de verdeling van de meetfouten.
- (c) Heb je het idee dat de metingen van Michelson systematisch te hoog of te laag zijn? Waarom?

Opgave 3.3 *Chi-kwadraat toets*

Een fabrikant van weegschalen beweert dat door hen gefabriceerde weegschalen een massa van (ongeveer) 1000mg meten met een nauwkeurigheid van 2mg. Dat betekent dat als X een meting voorstelt van een object van (ongeveer) 1000mg, de variantie van X gelijk is aan $2^2 = 4\text{mg}^2$. We willen onderzoeken of de fabrikant gelijk heeft. Daartoe bepalen we de massa van het object met behulp van 100 willekeurige weegschalen van de fabrikant. De verschillende metingen worden genoteerd met $X_1, \dots, X_{100}$. We veronderstellen dat de waarnemingen $X_1, \dots, X_{100}$ onderling onafhankelijk en normaal verdeeld zijn met onbekende verwachting μ en onbekende variantie σ^2 (μ ligt in de buurt van 1000 mg). De gevonden waarnemingen staan op de website <http://www.math.vu.nl/sto/onderwijs/statistiekvoorNatuurkunde>.

- (a) Ga met behulp van een QQ-plot na of de normaliteitsaannname acceptabel is.
- (b) Construeer een 95% betrouwbaarheidsinterval voor σ^2 . Welke pivot heb je gekozen? Laat zien welke stappen je hebt genomen! Als hulp voor deze opgave kun je Voorbeeld 5.29 bestuderen. De kwantielen van de chi-kwadraat-verdeling kunnen worden gevonden met de functie `qchisq`. Let op dat je het juiste aantal vrijheidsgraden meegeeft.
- (c) Toets of σ^2 significant afwijkt van de door de fabrikant opgegeven onnauwkeurigheid van $2^2 = 4\text{mg}^2$. Neem de onbetrouwbaarheidsdrempel gelijk aan $\alpha = 0,05$. Deze toets mag je

uitvoeren met behulp van het betrouwbaarheidsinterval in het vorige onderdeel, een kritiek gebied of een overschrijdingskans. Geef duidelijk aan welke stappen je neemt. Voorbeeld 5.29 kan opnieuw hulp bieden. Bij deze opgave kan de functie `pchisq` van pas komen.

Opgave 3.4 *Machtswetverdeling, vervolg*

In Serie 2 stond een opgave over sneeuwlawines: In de bergen vinden regelmatig sneeuwlawines plaats, grote maar ook kleine. De grootte van een lawine volgt bij benadering een machtswetverdeling:

$$p_\alpha(s) = \begin{cases} (\alpha - 1) s^{-\alpha} & \text{voor } s \geq 1, \\ 0 & \text{anders} \end{cases} \quad (1)$$

voor zekere $\alpha > 1$ onbekend. In die opgave hebben jullie afgeleid dat de maximum likelihood schatter voor α op basis van een steekproef $X_1, \dots, X_n$ uit deze verdeling gelijk is aan $\hat{\alpha} = 1 + n/(\sum_{i=1}^n \log X_i)$. In deze opgave bepalen we een bijbehorend 95% betrouwbaarheidsinterval voor α .

- Bepaal de Fisher informatie voor α op basis van een enkele waarneming.
- Welke verdeling heeft de maximum likelihood schatter $\hat{\alpha}$ bij benadering, als het aantal waarnemingen groot is?
- Schat de Fisher informatie met de waargenomen informatie of de plug in schatter.
- Bepaal een 95% betrouwbaarheidsinterval voor α ; eerst theoretisch, daarna numeriek.

Opgave 3.5 *Meetonzekerheden*

We willen de stroomsnelheid van water door een buis bepalen. De experimentele methode is echter niet erg nauwkeurig. Daarom herhalen we de meting van de stroomsnelheid zes keer. We vinden achtereenvolgens de volgende waarden in m/s: 3.7, 3.5, 2.4, 3.9, 2.9, 3.4.

We kunnen een meting van de stroomsnelheid opvatten als een realisatie van een stochastische grootte X , die normaal verdeeld is met als verwachting de werkelijke stroomsnelheid. Zouden we een waarde willen rapporteren voor de werkelijke stroomsnelheid, dan moeten we de verwachting van de verdeling van X schatten en de onzekerheid in die schatting bepalen. Als schatter voor de werkelijke stroomsnelheid kiezen we het gemiddelde van zes metingen, $\bar{X}_6$. Op basis van de bovenstaande waarnemingen kunnen we een realisatie van deze schatter berekenen: 3.3 m/s. Om de onzekerheid in de schatting te bepalen, moeten we de standaardafwijking van de verdeling van $\bar{X}_6$ schatten.

- Schat de standaardafwijking van de verdeling van $\bar{X}_6$, dus van de verdeling van het gemiddelde van zes metingen van de stroomsnelheid.

In sommige gevallen zijn we niet geïnteresseerd in de verdeling van $\bar{X}_6$, maar in de verdeling van X zelf. Deze situatie doet zich onder andere voor wanneer men een foutvlag moet opgeven in een regressie – dit zal in opgavenserie 4 aan de orde komen.

- Schat de standaardafwijking van de verdeling van X , dus van de verdeling van één meting van de stroomsnelheid.

Statistiek voor Natuurkunde

Opgavenserie 4: Lineaire regressie

Inleveren: nog te beslissen

Afspraken

Overleg is toegestaan, maar iedereen levert zijn eigen werk in. Overschrijven of kopiëren van antwoorden is niet toegestaan en zal worden bestraft.

Werk de opdrachten uit en schrijf een overzichtelijk verslag. Het verslag moet te begrijpen zijn voor iemand die de opdrachten niet gelezen heeft. Dat betekent dat je ook een beschrijving van de context en de vraagstelling moet geven. De R-code komt in een bijlage. Ik wil een uitgeprinte versie; inleveren per e-mail is niet mogelijk!

Opgave 4.1 *Snelheid van een speelgoedautootje*

We bestuderen met behulp van een fotocamera de beweging van een speelgoedautootje over een baan, die zodanig is geplaatst dat de snelheid van het autootje constant is. Langs de baan is een meetlint bevestigd; de afstand die het autootje heeft afgelegd kan worden bepaald door een foto te nemen en op de foto de positie van het autootje op het meetlint af te lezen. Het doel van het experiment is om de snelheid van het autootje te bepalen.

Op vaste tijdstippen maken we een foto van het autootje en bepalen we zijn positie. Het tijdstip van de eerste foto noemen we $t = 0$. De afgelegde afstand s op tijdstip t bepalen we door de afgelezen positie op tijdstip $t = 0$ steeds af te trekken van de afgelezen positie op tijdstip t . Als we de snelheid van het autootje noteren met v , wordt het verband tussen s en t gegeven door

$$s = vt.$$

De waarnemingen zijn in onderstaande tabel weergegeven.

tijd (s)	0.4	1.0	1.1	3.1	3.3	4.1	5.0	7.5
afstand (cm)	10.2	18.6	28	54.2	64.2	73.2	91.3	135.2

- Stel een lineair regressiemodel op voor de waarnemingen; geef daarbij onder andere duidelijk aan welke aannames je over de meetfouten doet.
- Schat de snelheid van het autootje met de kleinste-kwadratenmethode. Teken de meetpunten en de gefitte regressielijn in een grafiek. Bepaal ook de onzekerheid in de schatting van de snelheid.

Op het eerste gezicht lijkt het gebruikte model goed te passen bij de data. In een plot van de residuen van de regressie tegen de tijd, lijkt er echter een systematisch effect zichtbaar te zijn in de residuen: op vroege tijdstippen zijn de residuen gemiddeld groter dan op latere tijdstippen. Dit kan veroorzaakt zijn door toeval, maar het zou ook kunnen dat er een systematische meetfout is gemaakt, bijvoorbeeld doordat de positie op $t = 0$ niet goed is bepaald. We kunnen de mogelijke aanwezigheid van een systematische meetfout onderzoeken door een intercept aan het regressiemodel toe te voegen:

$$s = s_0 + vt,$$

waarbij de regressieparameter s_0 de systematische meetfout modelleert.

- (c) Voer opnieuw een regressie uit, nu met het regressiemodel met intercept. Onderzoek de aanwezigheid van een systematische meetfout door de nulhypothese dat de interceptparameter s_0 gelijk is aan 0 te toetsen; doe dit bij onbetrouwbaarheid 0.05. Formuleer je conclusie.
- (d) Bediscussieer kort welk model het beste geschikt is om de snelheid van het autootje te bepalen, en geef de geschatte waarde van de snelheid met onzekerheid.

Stel dat we dit experiment over zouden kunnen doen, waarbij we de tijdstippen waarop de foto's worden genomen zelf mogen kiezen (het is mogelijk om enkele foto's zeer kort na elkaar te nemen). Je kunt echter slechts 8 foto's maken. De keuze van de tijdstippen is van invloed op de conclusies die we uit het experiment kunnen trekken.

- (e) Beredeneer hoe je de tijdstippen waarop de foto's worden genomen zou moeten kiezen als je wilt onderzoeken of de snelheid van het autootje inderdaad constant is.
- (f) Veronderstel dat je zeker weet dat het autootje met constante snelheid beweegt, en dat je geen systematische fout maakt in de bepaling van de beginpositie. Leg uit hoe je de tijdstippen dan moet kiezen om de snelheid zo nauwkeurig mogelijk te bepalen.

Veronderstel nu dat bij nadere inspectie van de experimentele methode blijkt dat de tijdstippen toch niet nauwkeurig bekend zijn, dus dat de toevallige meetfout in de tijdstippen niet kan worden verwaarloosd. Gegeven is nu dat de meetonzekerheid in de positie gelijk is aan 1.0 cm, dat wil zeggen dat de standaardafwijking van de verdeling van de meetfout in de positie steeds 1.0 cm bedraagt. Verder is nu gegeven dat de onzekerheid in de tijdstippen 0.2 seconde is.

- (g) Voer nu lineaire regressie uit voor deze situatie, waarin dus beide variabelen een meetfout hebben. Bepaal de waarde van de snelheid met onzekerheid.

Opgave 4.2 *Foto-elektrisch effect*

Het foto-elektrisch effect is het verschijnsel dat sommige metalen elektronen uitzenden als zij met licht worden bestraald. De verklaring voor dit verschijnsel is dat elektronen uit de geleidingsband van het metaal kunnen ontsnappen doordat het metaal een foton absorbeert. Als W de potentiaal is waarmee het elektron aan het metaal is gebonden, en hf de energie van het geabsorbeerde foton (met f de frequentie van het licht en h de constante van Planck), dan kan het elektron ontsnappen indien $hf > W$. Het elektron komt dan vrij met kinetische energie

$$E_{\text{kin}} = hf - W.$$

TODO: PLAATJE INVOEGEN

Het foto-elektrisch effect kan worden onderzocht met een opstelling zoals geschetst in bovenstaande figuur. In deze opstelling worden de bij de anode (A) vrijkomende elektronen afgeremd in een tussen de anode en kathode (K) aangelegd elektrisch veld. Door de spanning V steeds groter te maken worden de elektronen steeds meer afgeremd, en zullen uiteindelijk zelfs de snelste elektronen (die, waarvoor de bindingspotentiaal W het kleinst is) de kathode niet meer bereiken. Bij die spanning zal de elektrometer E plotseling geen elektronen meer detecteren.

Laat, voor vast gekozen frequentie f van het licht, V_0 de spanning zijn waarbij de elektrometer geen elektronen meer detecteert, en schrijf W_0 voor de bindingsenergie van de minst gebonden elektronen in het metaal. Dan geldt

$$eV_0 = hf - W_0,$$

waarbij e de elektronlading is. Experimenteel kan door gebruik te maken van monochromatische lichtbronnen V_0 gemeten worden als functie van f . Met zo'n experiment kan de zogenaamde uittree-energie W_0

van het betreffende metaal bepaald worden, alsmede de waarde van de constante van Planck, als we de waarde van de elementaire lading e als gegeven beschouwen.

In een experimentele opstelling zoals hierboven geschetst meet men bij licht van verschillende golflengtes de spanning V_0 . De resultaten staan in onderstaande tabel.

golflengte (nm)	V_0 (V)
330	1.80
389	1.28
410	1.08
434	0.92
486	0.60
502	0.51
588	0.14

De frequentie van het licht f is te bepalen uit de relatie $f = c/\lambda$ met λ de golflengte en c de lichtsnelheid, $2.99872 \cdot 10^8$ m/s.

- (a) Zet de metingen uit in een grafiek, met V_0 op de verticale as tegen f op de horizontale as. Vergeet niet een label bij de assen te plaatsen.

De golflengtes zijn bepaald met een verwaarloosbare onnauwkeurigheid. De onnauwkeurigheid in V_0 is niet te verwaarlozen. In deze opgave worden achtereenvolgens twee verschillende situaties beschouwd met betrekking tot de onzekerheden in de metingen van V_0 . We onderzoeken de invloed van deze situaties op de resultaten van de regressie-analyse.

Situatie I: variantie in meetfouten gelijk en onbekend

Als eerste zullen we aannemen dat de variantie van de verdeling van de meetfout in V_0 bij elke meting even groot is, maar een *onbekende* waarde heeft. Deze aanname zullen we **situatie I** noemen.

- (b) Definieer een geschikt lineair regressiemodel voor de metingen. Geef onder andere aan wat de verdeling van de meetfouten is.
- (c) Gegeven is $e = 1.6021773 \cdot 10^{-19}$ C. Schat de parameters W_0 en h met de kleinste-kwadraten-methode. Vergeet niet om de eenheden te vermelden. Teken de gefitte lijn in de grafiek van opgave (a).
R-tips: Volg de aanwijzingen in het document *Enkele vaardigheden in R*. Gebruik de functie `lm`. De gefitte lijn kun je gemakkelijk tekenen door de functie `abline` toe te passen op het object dat de functie `lm` uitvoert. (Zie eventueel `help(abline)`, bij het argument `reg.`) Je kunt de lijn opmaken door aan `abline` het extra argument `lty="solid"` mee te geven (zie *Enkele vaardigheden in R*).
- (d) Bepaal de onzekerheden (standaardafwijking) in de schatters voor W_0 en h .
- (e) Plot in een nieuwe grafiek de (ongeschaalde) residuen op de verticale as, tegen f op de horizontale as. Is er een systematisch effect te zien? Ga ook na of het redelijk is om te veronderstellen dat de meetfouten normaal verdeeld zijn.
- (f) Toets of de gevonden waarde voor h significant afwijkt van de uit de literatuur bekende waarde van de constante van Planck, $h = 6.626069 \cdot 10^{-34}$ Js. Neem de onbetrouwbaarheid gelijk aan 0.05. Geef de nul- en alternatieve hypothese, de waarde van de toetsingsgrootheid, de overschrijdingskans en je conclusie.

- (g) Bepaal een 90% betrouwbaarheidsinterval voor W_0 .
- (h) Leg uit of het mogelijk is om een goodness-of-fit toets uit te voeren, zoals die besproken is in het college. Zo ja, voer de toets uit. Geef dan de waarde van de toetsingsgrootte, de overschrijdingskans en je conclusie.

Situatie II: variantie in meetfouten gelijk en bekend

We doen nu een andere aanname over de variantie van de meetfouten in de metingen van V_0 . Deze is nog steeds voor elke meting van V_0 hetzelfde, maar heeft nu een bekende waarde, bijvoorbeeld doordat de fabrikant van de voltmeter de onnauwkeurigheid van de meter heeft opgegeven: de standaardafwijking van de verdeling van de meetfout in V_0 is steeds gelijk aan 0.03 V (de variantie is dus $(0.03 \text{ V})^2 = 0.0009 \text{ V}^2$). Deze aanname over de meetfouten noemen we **situatie II**.

Een bekende meetonzekerheid wordt in een grafiek vaak weergegeven door een foutvlag. De lengte van een foutvlag heeft een precieze interpretatie: de halve lengte van de foutvlag geeft de standaardafwijking van de verdeling van de meetfout in één meting weer (vergelijk opgave 3.5!). In deze opgave is dus de halve lengte van de foutvlaggen bij de metingen van V_0 gelijk aan 0.03 V.

Bij vraag (a) heb je de metingen uitgezet in een grafiek. Omdat de meetfouten in situatie I onbekend waren, heb je daar geen foutvlaggen getekend.

- (i) Teken nu, in dezelfde figuur als die je bij vraag (a) hebt gemaakt, de foutvlaggen in die horen bij situatie II. In R kun je hiervoor de volgende code gebruiken:

```
errorbars<-function(x,y,ebl,ebu=ebl,length=0.08,...) {
  arrows(x,y+ebu,x,y-ebl,angle=90,code=3,length=length,...) }
errorbars(x,y,0.03) # errorbars met onzekerheid 0.03 in y
```

- (j) Voer nu, met deze nieuwe aanname van bekende standaardafwijking van de verdeling van de meetfouten, de onderdelen (b) t/m (h) van deze opgave nog een keer uit. *Zorg ervoor dat in het verslag dat je inlevert steeds ondubbelzinnig duidelijk is over welke situatie je het hebt!*
- (k) Vergelijk in een paar regels tekst de situaties I en II met elkaar. Welke onderdelen van de regressie-analyse worden beïnvloed door de aannames over de variantie van de meetfouten?

Opgave 4.3 Foto-elektrisch effect (vervolg)

Deze opgave is een vervolg op de voorgaande opgave over het foto-elektrisch effect.

In de vorige opgave werden twee verschillende aannames over de variantie van de meetfouten gedaan. In deze opgave zullen we nog twee andere aannames doen over de variantie van de meetfouten, en onderzoeken we opnieuw de invloed die deze aannames hebben op de resultaten van de regressie-analyse.

De experimentele situatie zoals beschreven in de vorige opgave blijft hetzelfde, op een ding na. Waar in de vorige opgave bij elke golflengte de meetonzekerheid in V_0 dezelfde waarde had is dat nu niet zo. De metingen van V_0 die bij verschillende golflengtes gedaan worden, hebben nu een verschillende onzekerheid, zoals aangegeven in onderstaande tabel. Merk op dat de meetwaarden van de golflengtes en de V_0 niet zijn veranderd.

golflengte (nm)	V_0 (V)	onzekerheid in V_0 (V)
330	1.80	0.03
389	1.28	0.04
410	1.08	0.06
434	0.92	0.03
486	0.60	0.02
502	0.51	0.02
588	0.14	0.03

De precieze interpretatie van het concept onzekerheid wordt hieronder gegeven.

Situatie III: variantie in meetfout ongelijk en bekend

We doen nu de volgende aanname over de meetfouten: de standaardafwijking van de meetfouten in V_0 is bij elke golflengte anders, en is bekend: de waarde wordt gegeven in de tabel hierboven. Dit noemen we **situatie III**.

- (a) Teken in een nieuwe figuur een grafiek van V_0 tegen f . Teken de foutbalken voor V_0 ook in.
R-tip: Aan de functie `errorbars` van opgave 4.2(a) geef je nu als derde argument de vector met onzekerheden mee.
- (b) Beantwoord nu, voor situatie III, de vragen (b) t/m (h) van opgave 4.2. Let op: gebruik bij vraag (e) de geschaalde residuen, dat wil zeggen de residuen $y_i - \hat{\alpha} - \hat{\beta}x_i$ gedeeld door de standaardafwijking.

Situatie IV: variantie in meetfout ongelijk en onbekend

De laatste situatie die we beschouwen, betreft de volgende aanname over de meetfouten. De standaardafwijking van de meetfouten in V_0 is niet bekend. Wel is de *verhouding* tussen de standaardafwijkingen van de meetfouten bekend: de standaardafwijkingen van de verdeling van de meetfouten verhouden zich als gegeven in de tabel hierboven. Deze situatie noemen we **situatie IV**. Voorbeeld: de standaardafwijking van de meetfout in V_0 bij de eerste meting is de helft (0.03/0.06) van de standaardafwijking van de meetfout in V_0 bij de derde meting.

- (c) Beantwoord nu, voor situatie IV, de vragen (b) t/m (h) van opgave 4.2. Gebruik bij vraag (e) de geschaalde residuen, dat wil zeggen de residuen $y_i - \hat{\alpha} - \hat{\beta}x_i$ gedeeld door de gewichten horend bij de standaardafwijking.
- (d) Beschrijf je conclusies van de opgaven 4.2 en 4.3. Vergelijk de situaties III en IV met elkaar en met situaties I en II. Geef aan welke onderdelen van de regressie-analyse worden beïnvloed door de aannames over de variantie van de meetfouten.