

DIE METODOLOGIE VAN AKADEMIESE PRESTASIEVOORSPELLING

3.1 INLEIDING

In hoofstuk twee is aangetoon dat akademiese prestasie voorspel word op grond van kennis aangaande die veranderlikes wat daarmee verband hou. Dit is die doel met hierdie hoofstuk om kennis te neem van die metodes wat aangewend kan word om die beste voorspellers van akademiese prestasie te identifiseer. Die klem val daarom op die basiese beskrywing en rasionaal van beskikbare metodes. Die doel is nie om die besondere metodes en statistiese tegnieke in tegniesse terme te beskryf nie. Dit word wel met geselekteerde tegnieke in hoofstuk ses gedoen.

Die metodologie van akademiese prestasievoorspelling behels 'n verskeidenheid prosedures en relevante statistiese metodes waarvan kennis geneem moet word. Die prosedures hang saam met die verloop van navorsing oor akademiese prestasievoorspelling, en was in 'n mindere of meerdere mate ter sprake in verteenwoordigende studies in hierdie veld soos dié van Monteith (1983), Van der Watt (1982) en Venter (1983). Dit word ook in besonderhede beskryf in erkende handboeke oor navorsingsmetodologie soos dié van Kerlinger en Pedhazur (1973). Die metodologiese kwessies het te make met die seleksie van voorspeller veranderlikes en van proefpersone. Heel aanvanklik word beskikbare veranderlikes en metings daarvan bestudeer. In hierdie stadium word 'n seleksie uit die beskikbare data gemaak. Hierna word 'n keuse van proefpersone gemaak; 'n prosedure wat noodsaaklik is, aangesien daar vir homogene groeperings van leerlinge meer akkuraat voorspel kan word (Zedeck, 1971a:295). Die laaste prosedure wat gevolg word, is die identifisering van beste voorspellers. In die geval van elk van hierdie prosedures bestaan statistiese tegnieke waarvan kennis geneem moet word.

Om die doel met die hoofstuk te bereik, word daar eerstens kennis geneem van die wyses waarop die aanvanklike seleksie van

voorspeller veranderlikes plaasvind (paragraaf 3.2). Tweedens word besin oor die verdeling van proefpersone vir wie die seleksie gedoen moet word, in homogene groepe (paragraaf 3.3). Derdens word die wyses beskryf waarop die seleksie van beste voorspellers van akademiese prestasie gedoen word (paragraaf 3.4). Laastens word kennis geneem van die faktore wat die seleksie van voorspellers en dus die akkuraatheid van voorspelling beïnvloed.

3.2 DIE AANVANKLIKE SELEKSIE VAN VOORSPELLERVERANDERLIKES

3.2.1 INLEIDING

In die voorspelling van akademiese prestasie is dit vanselfsprekend dat die veranderlikes op grond waarvan voorspel word, noukeurig geselekteer word. Hiervoor bestaan daar 'n aantal besondere prosedures en statistiese tegnieke wat verband hou met die stadium van seleksie. Daar word normaalweg onderskei tussen **aanvanklike** en **uiteindelike** seleksie van voorspeller veranderlikes.

Die wesenlike verskil tussen die twee stadia van seleksie, is dat daar by die **aanvanklike seleksie** met die voorspeller veranderlikes as sodanig gewerk word, en dat dit by die **uiteindelike seleksie** gaan om die identifisering van veranderlikes wat met 'n kriterium verband hou. By die **aanvanklike seleksie** word die verbande tussen potensiële voorspeller veranderlikes onderling ondersoek, en by die **uiteindelike seleksie** die verbande tussen voorspeller veranderlikes en 'n kriterium.

Die **aanvanklike seleksie** van voorspeller veranderlikes vind in die beginstadium van navorsing plaas wanneer die beskikbare spektrum van voorspeller veranderlikes (nog voordat die kriterium van voorspelling ter sprake is), bestudeer word. Dit vind tipies plaas wanneer 'n navorser byvoorbeeld oor 'n stel data beskik bestaande uit metings van verwante veranderlikes soos belangstelling, houding en gemotiveerdheid. So 'n spektrum is

problematies in dié sin dat oorvleueling in meting kan voorkom (die probleem van multikolineariteit), en dat daar onduidelikheid oor die konstrukte onderliggend aan die data kan wees.

Die doel met die aanvanklike seleksie is om eksploratief te werk te gaan en breë tendense in data, met hoofsaaklik twee oogmerke, te ondersoek. In hierdie stadium kan die navorser òf a) spesifiek hipoteses toets oor die bestaan al dan nie van onderliggende konstrukte ¹ in 'n datastel, òf b) hy kan vasstel of daar hoegenaamd enige konstrukte wel moontlik in 'n gegewe beskikbare stel data aanwesig is (Nunnally, 1967:289). Laasgenoemde ondersoek is verkennend van aard en daar is gewoonlik nie sprake van spesifieke hipoteses nie (anders as in die geval van die uiteindelijke seleksie van voorspeller veranderlikes).

Die uiteindelijke seleksie van veranderlikes vind plaas wanneer dit gaan om die spesifieke identifisering van beste voorspellers van 'n kriterium. Hieroor word in paragraaf 3.4 gehandel. In hierdie paragraaf val die klem verder op die aanvanklike seleksie van veranderlikes.

3.2.2 OORWEGINGS BY DIE AANVANKLIKE SELEKSIE VAN VOORSPELLERVER- ANDERLIKES

Daar is 'n aantal teoretiese en praktiese oorwegings by die aanvanklike seleksie van voorspeller veranderlikes. Hierdie oorwegings hang spesifiek saam met die aard van die datastel waarmee gewerk word, en die situasies waarin veranderlikes geselekteer word.

Die aanvanklike seleksie van voorspeller veranderlikes kom ter sprake wanneer daar met 'n groot hoeveelheid veranderlikes gewerk

1 Met konstrukte word hier bedoel bepaalde begrippe wat deur verbande tussen spesifieke metings van veranderlikes gedefinieer word.

word. In sodanige gevalle poog die navorser om ter wille van teoretiese oorwegings die datastel sinvol te vereenvoudig en meer bruikbaar te maak sonder verlies aan te veel inligting, en om ter wille van praktiese oorwegings met minder veranderlikes te werk.

Die teoretiese oorwegings hou verband met die mate van oorfleueling in metings van veranderlikes (vergelyk Bloom, 1976:193). 'n Meetinstrument van byvoorbeeld akademiese prestasie is nie net 'n aanduiding van die kognitiewe vermoëns van leerders nie, maar in 'n mindere of meerdere mate ook van veranderlikes soos houding teenoor die vak, studiegewoontes, en so meer. Wanneer so 'n meetinstrument saam met byvoorbeeld 'n meetinstrument wat houding teenoor 'n skoolvak meet, gebruik word, kan verwag word dat daar 'n mate van oorfleueling in meting voorkom.

Die probleem van oorfleueling in metings staan bekend as multikolineariteit. Dit dui op 'n gebrek aan ortogonaliteit en beteken dat veranderlikes interverwant is (Chatterjee en Price, 1977:143). Om hierdie probleem op te los, is dit nodig dat gerigte prosedures gevolg moet word ten einde veranderlikes sō te selekteer dat oorfleueling in metings tot die minimum beperk word. Dit het ensyds die voordeel dat 'n datastel teoreties vereenvoudig word en die navorser duidelikheid kry oor die strukture onderliggend aan sy data (McCornack, 1970:257). Andersyds word aan 'n belangrike vereiste vir doeltreffende voorspelling voldoen deurdat gepoog word om veranderlikes te kies wat onderling betreklik ongekorreleerd is.

Benewens bogenoemde oorwegings vir die aanvanklike seleksie van voorspeller veranderlikes, voer navorsers soos McCornack (1970:257), Schoeman (1976:16) en ander ook nog aan dat dit praktiese voordele het wanneer die aantal voorspeller veranderlikes verminder word en selfs tot die minimum beperk word. Schoeman (1976:16) beweer byvoorbeeld dat hoe meer veranderlikes as voorspellers in voorspelling studies gebruik word, hoe meer neem die akkuraatheid van voorspelling af. 'n Beperkte hoeveelheid veranderlikes vergemaklik ook die

moontlikheid van kruisgeldigheidsbepaling met ander steekproewe van proefpersone.

Ten slotte dien dit gemeld te word dat die prosedure om voorspeller veranderlikes in die beginstadium van navorsing te selekteer, algemene praktyk onder navorsers is (vergelyk die werk van Monteith, 1983 en Scott, 1984). Hiervoor is daar soos in hierdie paragraaf beredeneer, teoretiese en praktiese oorwegings en die doel is om 'n datastel te vereenvoudig deur veranderlikes te selekteer wat die minste oorvleuel en terselfdertyd die beste aanduiding van die konstrunkte onderliggend aan die data is.

3.2.3 DIE GEBRUIK VAN FAKTORANALISE VIR DIE AANVANKLIKE SELEKSIE VAN VOORSPELLERVERANDERLIKES

3.2.3.1 Inleiding

In hierdie paragraaf word faktoranalise beskryf as metode met behulp waarvan die aanvanklike seleksie van veranderlikes waargeneem kan word. Die doel is om die metode as sodanig te beskryf sonder om tegniese raak. Om hierdie doel te bereik, word faktoranalise allereers as metode getipeer en daarna word die doel en gebruike beskryf.

3.2.3.2 Omskrywing en tipering van faktoranalise

Volgens Child (1976:3 e.v.) is die metode van faktoranalise reeds in die negentiende eeu deur Galton en later deur Pearson, Spearman en Thurstone ontwikkel en toegepas. Die aanvanklike oogmerk was om wiskundige modelle in die Psigologie toe te pas, veral met betrekking tot intelligensie- en persoonlikheidsteorieë (Child, 1976:5).

Faktoranalise is 'n metode waardeur die interkorrelasies tussen veranderlikes bestudeer word. Dit is 'n metode waarmee die interrelasies tussen veranderlikes op 'n akkurate wyse weergegee

word om sodoende as hulp te dien in die pogings van navorsers om 'n begrip van die eienskappe van die mens te vorm (Gorsuch, 1974:2; Cattell, 1966:175).

Cattell (1966:174) beskryf faktoranalise as 'n metode waardeur die navorser tot 'n verminderde aantal abstrakte veranderlikes kom in ooreenstemming met die struktuur ("structural indications") in die data self; 'n metode om as't ware konsepte te skep en data op te som (vergelyk ook Frane en Hill, 1976:489).

Deur data aan 'n faktoranalise te onderwerp, word faktore geïdentifiseer wat sekere eienskappe van byvoorbeeld leerders voorstel. Sulke faktore is nie 'n direkte meting van 'n eienskap nie, maar word geïnterpreteer in terme van die veranderlike wat dit wel direk meet (Cureton en D'Agostino, 1983:3). Sodanige faktore word gewoonlik afgelei van die interkorrelasies tussen veranderlikes. Uit 'n gegewe groep veranderlikes vorm dié wat hoog met mekaar korreleer, 'n faktor wat gewoonlik 'n aanduiding van 'n teoretiese konstruk is wat die data verteenwoordig (en van die variansie wat die meetinstrumente deel) (Kerlinger en Pedhazur, 1973:361; Frane en Hill, 1976:489).

Ten opsigte van die faktore wat deur faktoranalise geïdentifiseer word, wys Irvine (1979:303) daarop dat daar minstens in metaforiese terme aan die faktore gedink moet word, omdat dit bloot verbeeldingryke ("imaginative") en idiosinkratiese pogings is om oënskynlik nie-verbandhoudende veranderlikes te groepeer. Die faktore kan hierteenoor hoogstens geldige beskrywings wees van verwantskappe wat tussen veranderlikes bestaan, en gee geen aanduiding van moontlike oorsaaklike verwantskappe nie.

3.2.3.3 Doel met en gebruike van faktoranalise

Die hoofdoel met die gebruik van faktoranalise is om die aantal veranderlikes in 'n datastel te verminder, terwyl daar terselfdertyd die maksimum hoeveelheid inligting in die data behoue bly. Die oorspronklike stel veranderlikes word verminder tot 'n kleiner stel wat die meeste van die betroubare variansie

van die oorspronklike stel van veranderlikes verklaar. Die kleiner stel veranderlikes kan dan gebruik word as "operational representatives" van die konstrakte onderliggend aan die volledige stel veranderlikes (Gorsuch, 1974:3). (Vergelyk ook Cattell, 1966:174, Nunnally, 1967:289 e.v. en Ebel, et al., 1969:986.)

Benewens die doel om die aantal veranderlikes te verminder, word daar volgens Everett en Entekin (1980:178) met die metode van faktoranalise nog twee verdere doeleindes nagestreef. Eerstens word dit aangewend om die struktuur of die dimensies van 'n stel data, by wyse van eksplorering en bevestiging te ondersoek. (In gevalle waar 'n navorser nie vooraf seker van die dimensies in sy data is nie, word faktoranalise gebruik om te eksploreer, maar in gevalle waar wel vermoedens aangaande die dimensies in 'n datastel bestaan, word dit gebruik om te bevestig.) Tweedens word gepoog om sinvolle liniêre samestellings van veranderlikes op te lewer wat as aanduiders (of "summary measures") in die plek van die groter aantal veranderlikes kan dien (vergelyk ook Gorsuch, 1974:3).

Die doel met faktoranalise is voorts om die probleem van multikolineariteit op te los (Velicer, et al., 1982:371; Everett en Entekin, 1980:165). Dit word gedoen deurdat faktore opgelewer word wat met mekaar ongekorreleerd is, en waarin die variansie wat meetinstrumente deel, ook gereflekteer word.

Uit bostaande blyk dit dus dat faktoranalise as metode geskik vir die aanvanklike seleksie van veranderlikes is. Soos beskryf, kan dit aangewend word vir die vermindering van die aantal veranderlikes, en om 'n datastel as sodanig te bestudeer om vas te stel watter struktuur onderliggend daaraan is.

3.2.3.4 Metodes van faktoranalise

Uit die werk van verskeie skrywers blyk dat daar hoofsaaklik twee basiese metodes van faktoranalise bestaan, te wete die algemene (of eenvoudige) faktoranalisemetode, en die hoofkomponentemetode

(Cattell, 1966:176; Frane en Hill, 1976:490; Gorsuch, 1974:12 en Velicer, et al., 1982:371).

Die ooreenkoms tussen hierdie metodes is daarin geleë dat die doeleindes met faktoranalise, soos beskryf in die voorafgaande paragraaf, deur beide bereik kan word. Die verskil is geleë in die wyse waarop die variansie in 'n stel data in faktore verreken word (Cureton en D'Agostino, 1983:296 e.v.). In die geval van hoofkomponentanalise word die totale variansie geanaliseer en elke faktor of hoofkomponent is 'n funksie van al die veranderlikes. In die geval van algemene faktoranalise word die gemeenskaplike variansie geanaliseer en voorsiening word ook gemaak vir die unieke variansie in 'n stel data.

In terme van hierdie onderskeid word algemene faktoranalise as 'n "oop model" beskou, omdat dit voorsiening maak vir swakker of minder duidelike faktore waarvoor verdere navorsing nodig is. Daarteenoor word hoofkomponentanalise as 'n "geslote model" beskou omdat die totale variansie in die hoofkomponente verreken word (vergelyk Cattell, 1966:176).

Verskeie navorsers is van mening dat die metode van hoofkomponentanalise in 'n meerdere mate deur navorsers in die gedragswetenskappe gebruik word (vergelyk Velicer, et al., 1982:286 en Von Mollendorf, 1978:47). Schoeman (1981:58) wys in hierdie verband daarop dat die resultate van die twee metodes in elke geval toenemend ooreenstem soos wat die kommunaliteite van veranderlikes (die variansie wat met ander veranderlikes gedeel word), 1,0 nader. Hoofkomponentanalise word egter die meer onderskeidende metode wanneer die kommunaliteite laag is of baie varieer (vergelyk Gorsuch, 1974:120 e.v.).

Die twee metodes van faktoranalise wat hierbo beskryf is, is wesenlik metodes waarvolgens faktorekstraksie onderneem word (Frane en Hill, 1976:496). Daarmee word bedoel dat beide gebruik word om faktore uit 'n datastel te onttrek. Vanweë die feit dat daar heelwat ooreenkomste tussen die twee metodes bestaan en dat daar met beide in die doel geslaag kan word om 'n datastel te reduseer, is dit volgens Frane en Hill (1976:497) 'n minder

belangrike probleem om te kies tussen algemene faktoranalise en hoofkomponentanalise.

Dit is veral 'n minder belangrike probleem wanneer daar met "goeie" ² data gewerk word (Frane en Hill, 1976:497). Dit is 'n groter probleem en moeiliker keuse, volgens die skrywers, om 'n beslissing oor die aantal faktore wat uit 'n datastel onttrek moet word, te maak. Die keuse is moeilik aangesien dit juis die doel met faktoranalise is om die aantal veranderlikes te verminder en dat daarom nouliks met 'n arbitrêre besluit oor die aantal faktore volstaan kan word.

In hierdie verband is dit normale praktyk dat daar van 'n kriterium gebruik gemaak word by die keuse van die aantal faktore waartoe 'n datastel gereduseer word. Die twee bestaande kriteria is Cattell (1966:206) se Scree-toets en Kaiser (Child, 1976:43) se kriterium. Hierdie twee word in paragraaf 6.7.2.2 omskryf. Verder is dit ook praktyk om veral met die oog op voorspelling, faktore te identifiseer wat sowel teoreties sinvol, as prakties bruikbaar is (vergelyk Kerlinger en Pedhazur, 1973:364).

In die uiteensetting tot dusver is die twee basiese metodes van faktoranalise beskryf. Dit blyk daaruit dat beide in wese wyses is waarop faktore uit 'n stel veranderlikes onttrek kan word, en dat hoofkomponentanalise die meer gebruikte (en onderskeidende) metode is. In die verbygaan is ook verwys na die kriteria vir die seleksie van die aantal faktore wat in die praktyk enigsins problematies is, en wat in paragraaf 6.7.2.2 bespreek word.

'n Laaste faset van faktoranalise wat by beide die beskrewe metodes ter sprake kom, is naamlik die wyses waarop faktore geroteer word. Frane en Hill (1976:497) wys daarop dat die keuse van faktore in elk geval gewoonlik nie gemaak word alvorens faktore geroteer is nie.

2 Die skrywers bedoel hiermee waarskynlik data wat aan die eise van betroubaarheid en geldigheid voldoen.

Die rotasie van faktore is basies metodes wat gebruik word om te verseker dat die beste passing van faktorasse verkry word (Steyn, 1983:10). Dit kom daarop neer dat faktore in meetkundige terme in die meerdimensionele subruimtes wat die veranderlikes voorstel, sodanig geroteer word dat dit die beste voorstelling van groepe ("clusters") van veranderlikes bied (Frane en Hill, 1976:499).

Die twee rotasiemetodes wat basies onderskei word, is die ortogonale en skuinsrotasiemetodes (Gorsuch, 1974:191,192). Deur ortogonale rotasie word ortogonale faktore gelever; faktore bestaande uit gekorreleerde veranderlikes, maar wat ongekorreleerd ten opsigte van ander faktore is. Skuinsrotasie daarenteen, lewer faktore op wat nie heeltemal ongekorreleerd is nie. Van hierdie twee rotasiemetodes is die ortogonale en spesifiek die varimax-tipe die mees gebruikte (Frane en Hill, 1976:498).

3.2.3.5 Die gebruik van faktoranaliseresultate in die voorspelling van akademiese prestasie

Soos aangetoon in die voorafgaande paragrafe, lewer die metodes van faktoranalise resultate in die vorm van 'n geselekteerde aantal faktore (of hoofkomponente) op waarmee die variansie in 'n dataset in 'n mindere of meerdere mate verklaar word. Daar bestaan verskeie wyses waarop hierdie resultate vir voorspellingsdoeleindes aangewend word. Afhangende van die doel daarmee, kan daar hoofsaaklik twee wyses onderskei word.

Die eerste wyse kom daarop neer dat daar vir die geïdentifiseerde hoofkomponente of faktore hoofkomponenttellings (of faktortellings) bereken word en dat hierdie tellings in meervoudige regressieanalises gebruik word. Hierdie prosedure is deur navorsers soos Schoeman (1976), Von Mollendorf (1978) en Schoeman (1981) gevolg. Die hoofdoel in hierdie studies was hoofsaaklik om optimale voorspellings in die vorm van 'n hoogs moontlike bepaaldheidskoëffisiënt (R^2) te lewer.

In die tweede moontlike prosedure word gebruik gemaak van faktorveranderlikes ³ om geïdentifiseerde faktore te verteenwoordig. Dit is 'n prosedure wat deur navorsers soos Monteith (1983), Venter (1983), Scott (1984) en ander aangewend is. Hierdie prosedure verskil van die voorgenoemde daarin dat daar nie van faktor- of hoofkomponenttellings gebruik gemaak word om te voorspel nie, maar dat een of meer veranderlike per faktor geselekteer word om die faktor te verteenwoordig. (Die presiese motivering hiervoor en wyse waarop dit gedoen word, word in paragraaf 7.3.3 verduidelik.)

Die genoemde twee prosedures verskil van mekaar omdat daar by elk verskillende oorwegings geld. Die hoofoorwegings, soos afgelei kan word uit die genoemde skrywers se werk, verskil in die volgende opsigte:

- a) In die geval van die hoofkomponente-prosedure word die relatiewe gewig van elke veranderlike in elke hoofkomponent gereflekteer. Dit bring mee dat hoofkomponente moeilik geïnterpreteer kan word en selfs grootliks oninterpreteerbaar is (vergelyk Gorsuch, 1974).

In die geval van die faktorveranderlike-prosedure word 'n hoë premie op die sinvolheid van faktore gestel. Die veranderlikes wat geselekteer word om faktore voor te stel, moet vir die praktyk van voorspelling sinvol en bruikbaar wees. Voorspellings op grond van die hoofkomponente-prosedure is in die geheel gesien meer van teoretiese waarde (vergelyk Schoeman, 1976:44).

- b) 'n Volgende baie belangrike oorweging by die hoofkomponente-prosedure is die waarde wat geheg word aan die variansie wat van 'n afhanklike veranderlike verklaar word. In studies

3 Die term faktorveranderlikes dui vir die doel van die bespreking op veranderlikes wat geselekteer is om geïdentifiseerde faktore te verteenwoordig. In paragraaf 6.7.2.3 word 'n uiteensetting van hierdie prosedure gegee.

waar hierdie prosedure gebruik is, is die hoofmerk deurgaans die berekening van 'n hoogs moontlike R^2 . Schoeman (1976:43) wys in hierdie verband byvoorbeeld daarop dat regressies op hoofkomponente waar veranderlikes met hoofkomponenttellings vervang word, beter skattings aan parameters gee as die klassieke voorspellingsmodelle (modelle op grond van veranderlikes as sodanig).

In die geval van die faktorveranderlike-prosedure word ook 'n hoë premie op 'n hoë R^2 geplaas, maar die klem word ook op die eenvoud en bruikbaarheid van die resultate gelê (vergelyk Venter, 1983:255 e.v. en Monteith, 1983:33 e.v.).

3.2.4 SAMEVATTING

In die voorafgaande paragrawe is daar gehandel oor die aanvanklike seleksie van voorspeller veranderlikes. Sodanige seleksie is veral gepas wanneer 'n navorser met 'n omvattende en dikwels onbekende stel data werk, wanneer duidelikheid gekry moet word van die onderliggende konstrakte in die data, en van die oorvleueling in metings wat moontlik mag bestaan. Seleksie in hierdie stadium word onderskei van uiteindelijke seleksie waar dit gaan om die identifisering van veranderlikes wat met die kriteriumveranderlike verband hou. (Meer besonderhede hieroor volg in paragraaf 3.4.)

Daar is hoofsaaklik twee metodes aan die hand waarvan die aanvanklike seleksie van voorspeller veranderlikes plaasvind, te wete die metodes van faktoranalise en hoofkomponentanalise. Beide die metodes is daarop gemik om 'n groter aantal veranderlikes (m) te bestudeer en 'n kleiner aantal (n) te selekteer sodat $n \ll m$ met die minimum verlies aan inligting.

3.3 DIE SELEKSIE EN GROEPERING VAN LEERLINGE

3.3.1 INLEIDING

Dit word algemeen aanvaar dat vir die doel van akademiese prestasievoorspelling, geldiger resultate verkry kan word deur die verdeling van proefpersone in homogene groepe (Zedeck, 1971a:295). Dit spreek vanself dat inligting aangaande voorspeller veranderlikes wat vir 'n homogene groep leerlinge ingesamel is, 'n hoër mate van geldigheid sal vertoon as dié vir meer heterogene groepe.

Die keuse van proefpersone deur middel van aanvaarde metodes van steekproefneming, is een metode waarvolgens min of meer homogene groepe uit 'n populasie geïdentifiseer kan word. Die metode is geskik in normale praktyk en gevalle waar vóóraf kennis van die kriteria gedra word waarvolgens 'n populasie in strata (vir gestratifiseerde steekproefneming), of bundels (vir bundelsteekproefneming), verdeel word. Op hierdie wyse kan ook meer homogene steekproewe verkry word.

Dit gebeur dat navorsers om bepaalde (dikwels praktiese) oorwegings 'n prosedure van steekproefneming volg wat 'n heterogene groep proefpersone oplewer. Dit gebeur ook dat navorsers wel steekproefprosedures volg wat min of meer homogene groepe oplewer, maar dan ter wille van geldiger voorspellings nog meer homogene groepe wil selekteer. In sodanige gevalle word gebruik gemaak van moderator veranderlikes waarvolgens die verdeling van proefpersone gedoen kan word. Hierdie algemeen gebruikte metode is soos aangetoon gaan word, geskik vir navorsing oor prestasievoorspelling waar dit belangrik is dat voorspelling steeds vir alle proefpersone gedoen moet kan word (vergelyk paragraaf 3.3.5). Ander tegnieke soos die tegniek van diskriminantanalise wat veral in Psigologie gebruik word, is meer gesofistikeerd en geskik vir die bestudering van groepverskille

en die waarskynlikheid dat proefpersone aan bepaalde groepe toegewys kan word (Van der Watt 1982:148).

Vir die doel van hierdie studie word vervolgens kennis geneem van die moontlikhede van moderator tegnieke. 'n Omskrywing hiervan word allereers in paragraaf 3.3.2 aangebied. Daarna word die prosedures vir die identifisering van moderator veranderlikes in paragraaf 3.3.3 beskryf. Ten slotte word op enkele probleme en voorbehoude by die gebruik van sodanige veranderlikes gewys.

3.3.2 MODERATORVERANDERLIKES: OMSKRYWING EN GEBRUIKE

Moderator veranderlikes word beskou as sodanige veranderlikes waarvolgens 'n groep proefpersone vir voorspellingsdoeleindes in subgroepe verdeel word. Sodanige verdeling lei daartoe dat onafhanklike veranderlikes beter voorspellings vir die subgroepe oplewer as vir die hele groep (Fourie, 1967:236).

Die term moderator verwys na bepaalde beheer- of populasieveranderlikes, na interaksies tussen veranderlikes, en/of na verskille in voorspelbaarheid van veranderlikes (Zedeck, 1971a:308). Hierdie gebruike van die term vind plaas in situasies waar die verhoging van die geldigheid van die resultate van 'n ondersoek ter sprake is (Dunette, 1963:318; Schoeman 1976:26).

Volgens Fourie (1967:212) gebruik navorsers veral **affektiewe veranderlikes** as moderator veranderlikes. Veranderlikes soos geslag en ouderdom word dikwels ook met vrag gebruik, en dit is veral menslike eienskappe wat met sukses as moderator veranderlikes gebruik kan word. Dunette (1963:318) sluit hierby ook veranderlikes soos onderwyspeil, vorige prestasies, belangstelling, en persoonlikheid in.

Daar is volgens Zedeck (1971a:308 e.v.) drie gebruike van moderator veranderlikes. Eerstens word dit gebruik waar **differensiële geldigheid** ter sprake is. Dit geld in gevalle waar die invloed van 'n derde veranderlike op die verwantskap tussen

die voorspeller (X) en die kriterium (Y) ondersoek word. Op die basis van die derde veranderlike word 'n steekproef in subgroepe verdeel en vir elke subgroep word die korrelasie tussen X en Y dan bereken.

Tweedens word moderatorveranderlikes met die oog op differensiële voorspelbaarheid gebruik. Groepe word ook verdeel en wel op grond van verskille in voorspelbaarheid (vergelyk die metode van Ghiselli wat in paragraaf 3.3.3 beskryf word). Derdens word moderatorveranderlikes gebruik in wat bekend staan as gemodereerde regressieonderseoeke. Hier is nie sprake van die subgroepering van individue nie, maar die invoeging in die regressievergelyking van 'n waarde gelyk aan dié van die modererende veranderlike (Saunders, 1956, aangehaal deur Schoeman, 1976:22).

Om die belangrikheid van moderatorveranderlikes in voorspellingstudies te illustreer, haal Schoeman (1976:27 e.v.) vir Williams (1959) aan wat die teorie van heterogene regressie beskryf. Hiervolgens word verskillende regressievergelykings vir verskillende subgroepe van 'n populasie bereken, en moderatorveranderlikes word gebruik om die verskille tussen die subgroepe uit te lig. Dit blyk daaruit dat die grootste voordeel van moderatorveranderlikes steeds is om groepe te identifiseer waarvan die metingsfout verskil en waar dieselfde veranderlikes geldiger voorspellings kan lewer.

Moderatorveranderlikes kan samevattend beskryf word as sodanige veranderlikes wat die geldigheid van navorsingsresultate beïnvloed en wat daarom aangewend word om 'n ondersoekpopulasie in meer homogene groepe te verdeel. Dit blyk dat leerlingeïenskappe soos geslag, ouderdom en belangstelling onder meer as moderatorveranderlikes in navorsing oor prestasievoorspelling kan dien.

3.3.3 DIE IDENTIFISERING VAN MODERATORVERANDERLIKES

Daar is verskeie wyses waarop potensiële moderatorveranderlikes geïdentifiseer kan word. Fourie (1967:212 e.v.) bespreek twee hiervan. Die eerste moontlike wyse is om die veranderlike(s) as voorspellers te gebruik. 'n Veranderlike wat as 'n goeie voorspeller van die kriterium uitgewys word, kan ook as moontlike moderatorveranderlike dien.

Die tweede wyse waarop moderatorveranderlike volgens Fourie (1967:213) geïdentifiseer kan word, is die metode wat Ghiselli (1960a:3) beskryf. Dit kom daarop neer dat die waarde van $D = Z_p - Z_k$ (waar Z_p die standaardpunttellings van die voorspellerveranderlikes en Z_k die standaardpunttellings van die kriteriumveranderlikes is), vir alle proefpersone bereken word. Hierna word die potensiële moderatorveranderlikes met die D-tellings van proefpersone gekorreleer, en indien die korrelasie hoog is, word die groep volgens sodanige moderatorveranderlikes verdeel. Sodoende word meer homogene subgroepe, waarvoor afsonderlik verder voorspellings gedoen word, verkry.

Volgens Banas (1964, aangehaal deur Zedeck, 1971a:306) word voorts 'n rasionele en 'n empiriese benadering tot die identifisering van moderatorveranderlikes onderskei.

Met die rasionele benadering maak die navorser staat op hipoteseformulering en op die identifisering van 'n veranderlike op 'n intuïtiewe wyse. Dit behels (a) die stel van 'n hipotese, (b) die intensiewe bestudering van 'n paar gevalle en die uitsoek van veranderlikes waarvan die interpretasie afhang van ander veranderlikes, (c) 'n soeke na subgroepe waarbinne korrelasies betekenisvol verskillend is, en (d) 'n toets van die interaksievariansie van een paar items teenoor 'n kriterium (Saunders, 1956, aangehaal deur Zedeck, 1971a:306). Hierdie benadering het volgens Zedeck (1971a:306) die nadeel dat dit tydrowend en

onekonomies is, omdat so 'n groot hoeveelheid moontlike hipoteses kan bestaan.

Die empiriese benadering behels volgens Banas (1964, aangehaal deur Zedeck, 1971a:306) ondersoek na werklike korrelasiekoëffisiënte tussen veranderlikes, asook 'n analise van responsvariasies (intra-individueel) en respons-inkonsekwentheid ("respons inconsistency") in die onderskeie meetinstrumente. Hierdie benadering het volgens Zedeck (1971a:306) die voordeel van objektiwiteit, maar is tog nie baie doeltreffend nie, omdat een veranderlike per keer bestudeer word.

Vanweë die probleme wat die genoemde twee benaderings tot die identifikasie van moderatorveranderlikes meebring, is navorsers dit eens dat sistematiese prosedures ondersoek moet word, veral wanneer met groot stelde data gewerk word (Zedeck, 1971a:306). Verskeie navorsers het reeds met prosedures geëksperimenteer en spesiale rekenaarprogramme geskryf waarin diskriminantanalise en kovariansieanalise byvoorbeeld aangewend is, maar telkens het kruisvalidasies bewys dat die gebruik van geselekteerde moderatorveranderlikes problematies is (Zedeck, 1971a:306). Oor die besondere voorwaardes vir die gebruik van moderatorveranderlikes, word in paragraaf 3.3.4 verder gehandel.

Uit die voorafgaande uiteensetting blyk dat daar basiese benaderings en metodes bestaan waarvolgens moderatorveranderlikes geïdentifiseer word. Vir die doel van hierdie paragraaf word kennis van die metodes as sodanig geneem, en ook van die feit dat die metodes bepaalde leemtes het waarmee rekening gehou moet word. In die empiriese ondersoek wat in hoofstuk ses beskryf word, kom die wyse waarop moderatorveranderlikes geïdentifiseer word, weer ter sprake wanneer 'n keuse gestel en gemotiveer word.

3.3.4 ENKELE PROBLEME MET EN VOORBEHOUDE BY DIE GEBRUIK VAN MODERATORVERANDERLIKES

In die praktyk lewer die gebruik van moderatorveranderlikes oënskynlik teleurstellende resultate, soos blyk uit die navorsing

van Zedeck (1971a) en De Villiers (1983). Laasgenoemde navorser het gebruik gemaak van die nie-intellektuele persoonlikheidseienskappe en belangstellings van Indiërleerlinge as moderatorveranderlikes en die proefpersone daarvolgens in drie groepe verdeel. Die meervoudige korrelasieonderzoek vir elk van die drie groepe ten opsigte van aanleg (soos gemeet deur die SATISA) en eksamenvakprestasies, het egter geen duidelike verskille in die resultate van die drie homogene subgroepe opgelewer nie. Op grond van hierdie resultate kan daar volgens De Villiers (1983:201), egter geen betekenisvolle gevolgtrekkings oor die gebruik van moderatorveranderlikes vir voorspelling gemaak word nie.

Die teleurstellende resultate is waarskynlik te wyte aan verskeie faktore of kombinasies van faktore. In hierdie verband meld Zedeck (1971:30) dat dit belangrik is dat die korrekte beslissings by die praktiese aanwending van moderatorveranderlikes geneem moet word. By die keuse van 'n geskikte veranderlike, moet 'n navorser byvoorbeeld sekerheid oor die mate van differensiële geldigheid wat opgelewer gaan word, hê en ook of voorspellers wel van nut sal wees vir een geïdentifiseerde groep en vir 'n ander nie. Verder moet ook besin word oor die aantal subgroepe wat geïdentifiseer word. Die aantal kan nie onbeperk wees nie en word bepaal deur die steekproefgrootte en die stabiliteit van die verskille tussen die groepe. Laastens moet die moontlike rol van 'n potensiële moderatorveranderlike as onafhanklike veranderlike ook ondersoek word. Sodanige insluiting in die gewone meervoudige regressievergelyking kan resultate oplewer waarin die moontlike invloed van die veranderlike wel verreken is.

Benewens die bogenoemde faktore, is daar ook nog ander maatreëls wat oorweeg kan word om die gebruikswaarde van moderatorveranderlikes te verhoog (vergelyk Zedeck, 1971a:307; Schoeman, 1976:21). Die nut en bruikbaarheid van hierdie veranderlikes kan byvoorbeeld verhoog word wanneer kontinue en nie diskrete veranderlikes nie, gebruik word om 'n heterogene populasie te verdeel. Die ideaal is dan ook dat die dusdanige verdeling in homogene subgroepe so moet plaasvind dat

veranderlikes geldige voorspellers vir alle subgroepe is. Dit gebeur byvoorbeeld dikwels dat die gebruik van 'n moderatorveranderlike om 'n populasie in subgroepe te verdeel, meebring dat voorspelling vir een subgroep daardeur baie verbeter en vir ander nie. Die implikasie is dan om swak of ongeldige voorspellers weg te laat of om nuwes by te kry. Eersgenoemde implikasie is oneties veral as daar met moderators soos geslag, ras, of geloof gewerk word. Om hierdie probleem te ondervang, pleit Zedeck (1971a:307) vir 'n soeke na geldige voorspellers vir alle groepe in 'n populasie (vergelyk ook Anastasi, 1968:142).

'n Verdere wyse waarop die nut en bruikbaarheid van moderatorveranderlikes verhoog kan word, is volgens Zedeck (1971a:307) die gelyktydige gebruik van 'n paar moderatorveranderlikes. Navorsers het egter met hierdie prosedure ook beperkte sukses behaal omdat soorte veranderlikes, steekproefgrootte, asook die korrelasie tussen moderatorveranderlikes, die resultate van voorspelling beïnvloed (Zedeck, 1971a:307; Schoeman, 1976:20). Hierteenoor blyk dit uit die werk van Klein, Rock en Evans (1968) dat die gesamentlike gebruik van moderatorveranderlikes soos ouderdom, belangstelling, vader se beroep en beroepskeuse, wel sinvol is. Die navorsers het aangetoon dat dit vir voorspellingsdoeleindes meer doeltreffend is om van 'n aantal moderatorveranderlikes saam gebruik te maak om proefpersone in meer homogene groepe te verdeel.

3.3.5 SAMEVATTING EN EVALUERING

Die doel met die voorafgaande uiteensetting was om kennis te neem van die wyses waarop leerlinge met die oog op die voorspelling van akademiese prestasies, in homogene groepe verdeel kan word. Dit blyk dat die hoofoorweging vir die gebruik van moderatorveranderlikes geleë is in die ideaal om sover moontlik akkurate en geldige voorspellings te kan maak. Die argumentasie berus op die aanname dat voorspelling akkurate is indien dit gebaseer is op inligting wat vir homogene subgroepe van leerlinge ingesamel is.

In die oorsig wat aangebied is, is gemeld dat veranderlikes soos geslag, ouderdom, belangstelling en nie-intellektuele persoonlikheidseienskappe as moderatorveranderlikes kan dien. Sodanige veranderlikes word op verskillende wyses geïdentifiseer en is gewoonlik ook goeie voorspellers van die kriterium.

Ten opsigte van die doeltreffendheid van die gebruik van moderatorveranderlikes, is daar nie eenduidige resultate nie. Dit wil voorkom asof die gebruik van meer as een moderatorveranderlike om 'n populasie in homogene subgroepe te verdeel, meer doeltreffend is. Die sukses hiervan hang egter steeds van die teoretiese oorwegings en van werklike korrelasiekoëffisiënte af. Ander probleme wat by die gebruik van moderatorveranderlikes kan ontstaan, is dat voorspelling nie vir alle groepe ewe doeltreffend is nie, en dat basiese veranderlikes soos ras, geslag of geloof vir voorspelling by sommige groepe ter wille van akkurater voorspellings weggelaat moet word; 'n praktyk wat navorsers as oneties beskou.

Verskeie prosedures vir die identifisering en gebruike van moderatorveranderlikes is beskryf. Omdat dit buite die bestek van hierdie studie val, word daar verder aan in die empiriese ondersoek nie geëksperimenteer met die verskillende wyses waarop moderatorveranderlikes geïdentifiseer word nie. In hoofstuk ses (vergelyk paragraaf 6.7.3) word die keuse van een basiese metode beskryf en gemotiveer.

Ten slotte is dit van pas om die gebruik van moderatortegnieke vanuit opvoedkundige oogpunt te evalueer. Fundamenteel beskou, is die doel met die gebruik van hierdie tegnieke aanvaarbaar. Die ideaal bly steeds, en so is dit ook in hoofstuk twee beredeneer, dat voorspellings so akkuraat moontlik gemaak moet word, en dat die inligting op grond waarvan voorspel word, geldig moet wees. Moderatortegnieke kan op die oog af 'n besliste bydrae lewer tot die bereiking van hierdie ideaal, omdat die gebruik daarvan die voorspellingsgeldigheid van resultate kan verbeter.

In die praktyk lewer die moderatortegnieke resultate wat in die lig van opvoedkundige oorwegings problematies is. Die gebruik van moderatorveranderlikes om 'n populasie in subgroepe te verdeel, kan meebring dat die subgroepe radikaal kan verskil ten opsigte van die mate van sukses waarmee akademiese prestasie voorspel word. Gestel byvoorbeeld die populasie (Afrikaanssprekende hoërskoolleerlinge in 'n nywerheidsgebied) word verdeel volgens geslag en persoonlikheid, en daar word voorspellingsformules vir subgroepe op grond van dieselfde stel voorspeller veranderlikes bereken. Dit kan dan gebeur dat dieselfde veranderlikes met 'n hoër mate van geldigheid na net een subgroep (sê dié wat seuns is en ekstroverte persoonlikhede het) se akademiese prestasie gedoen kan word. In die geval van die oorblywende subgroep kan dan met dieselfde veranderlikes minder geldige voorspellings gemaak word. Hierdie verskil in voorspelbaarheid by die twee subgroepe is vir die onderwyser/opvoeder problematies.

In die opvoedingsituasie kan die onderwyser nie anders as om hom met alle leerlinge te bemoei nie. Dit is nie opvoedkundig verantwoord om ter wille van die eis van geldige voorspelling vir een deel van 'n populasie met meer sukses te kan voorspel as vir 'n ander nie. Dit is veral nie verantwoord om veranderlikes soos geslag of geloof vir een groep weg te laat as voorspellers en vir 'n ander groep nie. Soos in hoofstuk twee aangetoon is, bemoei die onderwyser/opvoeder hom met alle leerlinge en moet hy vir almal doeltreffend kan voorspel. Terselfdertyd bly dit 'n vereiste dat voorspellings so akkuraat moontlik moet wees. Daarom is dit tog wel sinvol dat moderatorveranderlikes in hierdie studie gebruik word met die oog op die moontlikheid van akkurater en geldiger voorspellings. Daar moet egter nie uit die oog verloor word dat die eis van akkurate voorspelling vir alle subgroepe gestel moet word.

3.4 DIE IDENTIFISERING VAN BESTE VOORSPELLERS VAN 'N KRITERIUM

3.4.1 INLEIDING

In die inleiding tot hierdie hoofstuk (paragraaf 3.1) is aangevoer dat daar in antwoord op die vraag hoe akademiese prestasie voorspel word, allereers uitsluitel oor twee sake gekry moet word. Dit is naamlik die wyses waarop die aanvanklike seleksie van voorspeller veranderlikes kan plaasvind (vergelyk paragraaf 3.2) en leerlinge in homogene groepe verdeel kan word (vergelyk paragraaf 3.3).

Soos ook aangetoon, is kennis aangaande die aanvanklike seleksie van voorspeller veranderlikes en die verdeling van leerlinge in homogene groepe vir die doel van hierdie hoofstuk onontbeerlik. Die beskrewe metodes is basies vir die voorspelling van akademiese prestasie: alvorens die beste voorspellers van 'n kriterium geïdentifiseer kan word, moet 'n aanvanklike seleksie van veranderlikes plaasvind en duidelikheid oor die homogeniteit van die ondersoekgroep gekry word. Die redes hiervoor is reeds in paragraaf 3.1 aangevoer.

Die volgende metodologiese vraagstuk is die identifisering en seleksie van die deelversameling van beste voorspellers⁴ van 'n kriterium. Die basiese metode wat gevolg word, is om meervoudige verbande tussen voorspeller veranderlikes en 'n kriterium te ondersoek. Die metode van meervoudige regressieanalise is hiervoor geskik en word daarom in paragraaf 3.4.2 bespreek. Die

4 Met die "deelversameling van beste voorspellers" word daardie groep voorspeller veranderlikes bedoel wat die variansie in Y, die afhanklike veranderlike, die beste verklaar of wat 'n "beter passing" gee as ander groeperings van onafhanklike veranderlikes (De Wet, et al., 1981b:240).

twee besondere variasies van die basiese metode is die alle-moontlike- en die stapsgewyse-meervoudige regressietegnieke (vergelyk De Wet, et al., 1981b:240) wat onderskeidelik in paragrawe 3.4.3 en 3.4.4 bespreek word. In paragraaf 3.4.5 volg 'n evaluering van die genoemde metodes.

3.4.2 DIE METODE VAN MEERVOUDIGE REGRESSIEANALISE

3.4.2.1 Rasionaal en aannames

Meervoudige regressieanalise is by uitstek dié basiese statistiese metode wat gebruik word om akademiese prestasie te voorspel. Dit is geskik omdat daarmee die meervoudige verband tussen 'n aantal onafhanklike veranderlikes (of voorspeller veranderlikes soos in hierdie studie ook daarna verwys word), en 'n afhanklike of kriteriumveranderlike ondersoek kan word (Kerlinger en Pedhazur, 1973:3).

Meervoudige regressieanalise is basies 'n statistiese metode wat verbande met behulp van 'n wiskundige model ondersoek (Stoker, 1985:24). Die basiese liniêre model ⁵ sien soos volg daaruit:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 \dots, b_kX_k \text{ waar}$$

a en b_1 tot b_k konstantes is en ook regressiekoëffisiënte genoem word, en

Y die voorspelde waarde van die afhanklike veranderlike is (De Wet, et al., 1981b:236).

Die meervoudige regressievergelyking hierbo dui op die mate waarin Y in terme van $X_1 \dots, X_k$ uitgedruk kan word. Die waardes

5 Hierdie is die tipiese voorspellingsmodel wat in die praktyk aangewend word om Y-waardes (soos akademiese prestasie) te voorspel.

van b_1 ..., b_k dui die relatiewe gewig aan wat aan elke onafhanklike veranderlike (X_1 ..., X_k) in die voorspelling van Y toegeken word. Die meervoudige korrelasiekoëffisiënt R is 'n maatstaf van die liniêre verband wat daar tussen die voorspelde Y en die waargenome Y bestaan. Die R -waarde dui dus die korrelasie aan tussen die afhanklike veranderlike en die liniêre kombinasie $a + b_1X_1 + \dots + b_kX_k$ van die onafhanklike veranderlikes gesamentlik (De Wet, et al., 1981b:237).

Die vierkant van die meervoudige korrelasiekoëffisiënt, naamlik R^2 word die bepaaldheidskoëffisiënt genoem en dui die proporsionele deel van die variansie in Y aan wat deur 'n liniêre model gebaseer op die veranderlikes X_1 , X_2 , ..., X_k , verklaar word (Stoker, 1985:25).

Die gebruik van die model van meervoudige regressieanalise soos hierbo beskryf, berus op 'n aantal aannames. Eerstens veronderstel dit 'n liniêre verband tussen die voorspellers en die kriterium. Kerlinger en Pedhazur (1973:47, 48) stel hierby dat dit aanvaar word dat Y -tellings normaal versprei is. Hierdie aanname word gestel aangesien dit bepalend vir die geldigheid van die F -toets vir statistiese beduidendheid is. Tweedens word aanvaar dat die Y -tellings gelyke variansies vertoon by elke waarde van X .

Met hierdie uiteensetting van die rasionaal en aannames van meervoudige regressieanalise word volstaan. Dat dit as besondere statistiese tegniek vir die voorspelling van akademiese prestasie geskik is, word in die volgende paragraaf aangetoon.

3.4.2.2 Gebruik van meervoudige regressieanalise

Volgens McNemar (1969:185) het meervoudige regressieanalise twee hoofgebruik. In die eerste plek maak die metode dit moontlik dat die "optimale" kombinasie van voorspeller veranderlikes geïdentifiseer kan word. Daarmee word bedoel dat 'n stel uit die beskikbare voorspeller veranderlikes geselekteer word wat saam die hoogste meervoudige korrelasie met die kriterium oplewer. In die

tweede plek bied die metode ook die moontlikheid dat die variasie in die kriterium Y en die faktore wat daarvoor verantwoordelik is, ontleed kan word.

Deur meervoudige regressieanalise word daar dus bepaalde gewigte aan voorspeller veranderlikes toegeken op grond waarvan die beste moontlike beraming van 'n afhanklike veranderlike verkry kan word, en op grond waarvan die variansie in die kriterium verklaar kan word (Brislin, et al., 1973:290).

Die metode van meervoudige regressieanalise maak dit voorts ook moontlik dat die relatiewe bydrae van veranderlikes tot R^2 bepaal kan word (Kerlinger en Pedhazur, 1973:296). Deur sekere voorspeller veranderlikes konstant te hou, kan die proporsie van die variansie wat 'n ander veranderlike, of wat subversamelings van veranderlikes van die afhanklike veranderlike verklaar, vasgestel word. Sodoende kan die relatiewe bydrae van voorspeller veranderlikes ook vergelyk word (Felice, 1973:719).

Samevattend kan dit gestel word dat meervoudige regressieanalise vir die voorspelling van akademiese prestasie geskik is. Daarmee kan die meervoudige verbande tussen voorspeller veranderlikes en 'n kriterium ondersoek word, en ook volgens bepaalde metodes die stel of deelversameling van beste voorspellers geïdentifiseer word.

Die twee metodes wat gevolg kan word om die beste stel voorspeller veranderlikes te identifiseer, word onderskeidelik in paragrawe 3.4.3 en 3.4.4 bespreek.

3.4.3 STAPSGEWYSE REGRESSIEANALISE

Stapgewyse regressieanalise is 'n metode waarmee die beste voorspellers van 'n kriterium geselekteer kan word. In hierdie geval word voorspeller veranderlikes individueel oorweeg deur 'n proses van byvoeging tot of weglating uit die regressieanalise (Steyn, 1983:27, 28).

Volgens Kerlinger en Pedhazur (1973:285 e.v.) is daar twee wyses waarop individuele veranderlikes vir seleksie oorweeg word. Die eerste is deur voorwaartse seleksie en dit kom daarop neer dat veranderlikes volgens die grootte van die korrelasie met die afhanklike veranderlike in die regressieanalises gebruik word. Aanvanklik word die veranderlike wat die hoogste met die kriterium korreleer, in die analise ingesluit en daarna dié veranderlikes wat telkens die grootste toename in R^2 teweegbring (nadat die veranderlike(s) wat reeds in die vergelyking is, in berekening gebring is). Dit is dan ook dié veranderlikes wat die hoogste gekwadreerde semiparsiële korrelasie ⁶ met die afhanklike veranderlike het. Die toevoeging van veranderlikes tot die regressievergelyking word gestaak indien daar geen statisties-beduidende verhoging in R^2 is nie, of indien dit nie meer teoreties sinvol is nie.

Die tweede wyse van oorweging is volgens Kerlinger en Pedhazur (1973:286) die terugwaartse seleksie. Dit kom daarop neer dat alle onafhanklike veranderlikes aanvanklik in die analises ingesluit word en dat dit een vir een weggelaat word volgens dieselfde kriteria van statistiese beduidendheid en/of teoretiese sinvolheid wat by die voorwaartse seleksie gebruik word.

Die beskrewe twee wyses waarop veranderlikes vir seleksie oorweeg word, kom basies daarop neer dat regressieanalises stap vir stap gedoen word, en dat veranderlikes een vir een by die regressievergelyking gevoeg of daaruit weggelaat word totdat 'n bevredigende oplossing verkry word (vergelyk ook Draper en Smith, 1966:171).

Die rasionaal vir die gebruik van stapsgewyse seleksiemetodes, verklaar Bowles en Levin (1968:15 e.v.) baie duidelik aan die hand van die model $Q = X_1 + X_2$ waar Q = kriteriumprestasie en X_1

6 Die semiparsiële korrelasie tussen veranderlikes is daardie korrelasie tussen byvoorbeeld veranderlike X en Y nadat die invloed van ander veranderlikes wat met X en/of Y korreleer, deur berekening verwyder is (Gouws, et al., 1979:225).

en X_2 die onafhanklike veranderlikes is. Indien X_1 30 persent van die variansie in Q verklaar en $X_1 + X_2$ 40 persent, dan verklaar X_2 alleen slegs tien persent. Dit is egter slegs waar as die twee onafhanklike veranderlikes totaal onafhanklik (of ortogonaal) is.

Voorts poneer Bowles en Levin (1968:15) dat indien die onafhanklike veranderlikes wat in 'n stapsgewyse prosedure gebruik word, wel met mekaar korreleer, dit in gedagte gehou moet word dat beide 'n sekere hoeveelheid van die gesamentlike verklarende krag deel. Die gedeelde proporsie variansie in prestasie (Q) wat verklaar word deur δ of X_1 δ of X_2 , sal altyd toegeskryf word aan daardie veranderlike wat eerste in die regressieanalise gebruik is. Sodanige veranderlike X_1 se verklarende krag word oorskat, terwyl X_2 se verklarende krag onderskat word wat dan beskou word as 'n nadeel van die metode van stapsgewyse regressieanalise.

3.4.4 DIE ALLE-MOONTLIKE-DEELVERSAMELINGS MEERVOUDIGE REGRESSIEANALISE

Hierdie metode is 'n variasie van die reeds beskrewe meervoudige regressieanalise en behels die vergelyking en evaluering van alle moontlike kombinasies van onafhanklike veranderlikes as voorspellers (Hocking, 1972:967). Wanneer 'n navorser byvoorbeeld wil vasstel watter kombinasie van 'n gegewe tien onafhanklike veranderlikes die beste voorspelling van 'n kriterium lewer, word alle moontlike kombinasies van tien veranderlikes ($2^{10}-1 = 1023$ moontlike kombinasies) aan regressieanalises onderwerp. Verskeie kriteria word dan aangewend om te beslis watter kombinasie die "beste" voorspelling lewer.

Twee voor die hand-liggende kriteria wat aangewend kan word, is die bepaaldheidskoeffisiënt (R^2 -waarde) en die aangepaste bepaaldheidskoeffisiënt (R^2_a -waarde) (Steyn, 1983:30). Beide gee 'n aanduiding van die persentasie wat 'n bepaalde deelversameling van die variansie in 'n kriterium verklaar. Wanneer byvoorbeeld

ondersoek ingestel word na alle moontlike kombinasies van 'n gegewe tien veranderlikes, word dié kombinasie met die hoogste R^2 - of R^2a -waarde as die beste kombinasie beskou.

Aangesien die waarde van R^2 beïnvloed word deur die aantal veranderlikes (Steyn, 1983:30) en ook aangesien dit moontlik is om met 'n kombinasie van minder veranderlikes 'n groter persentasie van die variansie in die kriterium te verklaar, is dit ook sinvol om R^2a as kriterium te neem. R^2a is in elke geval aangepas vir die hoeveelheid veranderlikes in die deelversamelings en kan daarom 'n beter kriterium vir die seleksie van die deelversameling van beste voorspellers wees (Steyn, 1983:30).

Benewens die R^2 en R^2a -waardes, word daar ook gebruik gemaak van eksterne kriteria vir die keuse van beste deelversamelings. Een sodanige kriterium wat vry algemeen gebruik word, is die C_p -kriterium van Mallows (Hocking, 1972:969). 'n C_p -waarde word volgens 'n bepaalde formule (vergelyk paragraaf 6.7.2.4) vir elke moontlike deelversameling bereken. Die deelversameling met die kleinste C_p -waarde word dan beskou as die beste om 'n kriterium te voorspel.

3.4.5 EVALUERING VAN DIE METODES VAN STAPSGEWYSE EN DIE ALLE-MOONTLIKE-DEELVERSAMELINGS MEERVOUDIGE REGRESSIEANALISE

Met die uiteensetting in paragrawe 3.4.3 en 3.4.4 is aangetoon dat beide die metodes ter sprake aangewend kan word om beste voorspellers van 'n kriterium te selekteer. Die vraag is nou hoe hierdie metodes teenoor mekaar opweeg en watter een die mees geskikte vir die seleksie van die beste voorspellers van akademiese prestasie is.

In die geheel gesien blyk dit dat stapsgewyse regressieanalise betreklik minder analyses vereis (vergelyk ook Steyn, 1983:27). Die metode het egter die nadeel dat veranderlikes individueel oorweeg word en dat die moontlike kombinasies as sodanig nie geëvalueer word nie (Van der Watt, 1982:249). Daarby is dit 'n

nadeel dat 'n veranderlike wat deur voorwaartse of terugwaartse analises geselekteer is, nie weer weggelaat of bygevoeg kan word wanneer dit saam met ander veranderlikes 'n swak bydrae tot voorspelling blyk te lewer nie (Steyn: 1983:27).

Voorts blyk dit dat die metode van alle-moontlike-deelversamelings meer voor- as nadele inhou. Die enigste nadeel is waarskynlik geleë in die groot hoeveelheid berekenings wat vereis word, 'n nadeel wat in die lig van die beskikbaarheid van statistiese programppakkette op die rekenaar nie problematies is nie.

Die belangrikste voordeel van die metode van alle-moontlike-deelversamelings is daarin geleë dat kombinasies van veranderlikes ook in terme van sinvolheid oorweeg kan word (Van der Watt, 1982:251). Vir die opvoedkundige is dit van uiterste belang dat hierdie voordeel bestaan. Dit is, met die oog op die gebruik van die resultate van voorspelling in die praktyk, belangrik dat sinvolle keuses van kombinasies van veranderlikes gemaak kan word. Die voordeel van die metode van alle-moontlike-deelversamelings is juis geleë in hierdie kriterium van sinvolheid wat benewens die statistiese kriteria soos R^2 en C_p , aangelê kan word.

Uit die evaluering van die twee seleksiemetodes kan tot die gevolgtrekking gekom word dat die alle-moontlike-deelversamelings regressieanalise die verkieslike is. Navorsers soos Steyn (1983:32) en Van der Watt (1982:247 e.v.) kom na 'n soortgelyke evaluering tot dieselfde gevolgtrekking. Van der Watt (1982:247) baseer sy gevolgtrekking ook op 'n empiriese vergelyking van analises op dieselfde data. Sy bevinding was dat die twee metodes presies dieselfde R^2 -waardes oplewer, maar dat daar 'n klein variasie in die veranderlikes wat geselekteer word, is. Dit is daarom verkieslik om die alle-moontlike-deelversamelingsmetode te gebruik aangesien die sinvolheid van die deelversamelings en die veranderlikes wat daarin opgeneem is, ook oorweeg kan word.

3.5 FAKTORE WAT DIE IDENTIFISERING VAN BESTE VOORSPELLERVERANDERLIKES EN DUS DIE AKKURAATHEID VAN VOORSPELLING BEÏNVLOED

3.5.1 INLEIDING

Daar is 'n aantal faktore wat medebepalend vir die akkurate en doelmatige identifisering van beste voorspeller veranderlikes is. Hiervan word vervolgens kennis geneem aangesien daarmee rekening gehou moet word by die voorspelling van akademiese prestasie.

Die faktore onder bespreking word vir die doel van hierdie paragraaf gegroepeer in steekproeffaktore, veranderlike- en meetinstrumentfaktore, en ander faktore.

3.5.2 STEEKPROEFFAKTORE

Die steekproeffaktore het hoofsaaklik met die grootte en foute van steekproewe wat getrek word, te make. Volgens Kerlinger en Pedhazur (1973:442) veroorsaak klein steekproewe (100 en minder) 'n groot mate van fluktuasie in en onbetroubaarheid van regressiegewigte (dit is die relatiewe gewig wat veranderlikes in die voorspellingsmodel dra). Hierdie probleem kan opgelos word deur met groter groepe proefpersone te werk; groepe van meer as 100 en verkieslik meer as 200. Dit word aanvaar dat hoe groter die steekproef is, hoe meer stabiel is die resultate van voorspellingstudies (Kerlinger en Pedhazur, 1973:282).

Benewens die grootte van die steekproef, kan ander steekproeffoute ook 'n invloed op resultate hê. Ebel, et al. (1969:988) wys in hierdie verband daarop dat foute in die basiese prosedures vir wetenskaplike steekproefneming beperkte variasie in veranderlikes tot gevolg kan hê. Beperkteid in die variasie van byvoorbeeld die vermoëns van 'n groep individue kan die

grootte van korrelasiekoëffisiënte en dus die akkuraathed van voorspelling beïnvloed (Ebel, et al., 1969:988).

3.5.3 VERANDERLIKE- EN MEETINSTRUMENTEFAKTORE

Die veranderlikes en meetinstrumente waarmee gewerk word, is die basiese elemente vir voorspelling en daarom moet rekening met bepalende faktore in hierdie verband gehou word. Die faktore hier ter sprake, het te doen met die operasionalisering en meting van veranderlikes, die mate van oorfleueling in meting, en die aantal veranderlikes wat vir voorspelling geselekteer word.

'n Basiese faktor wat die seleksie van veranderlikes en voorspelling op grond daarvan beïnvloed, is die wyse waarop die veranderlikes as sodanig geoperasionaliseer word (Draper en Smith, 1966:163). Dit moet aanvaar word dat enige meetinstrument aan metingsfoutte onderhewig is en as sodanig onvolmaakte operasionaliserings van veranderlikes is. Hierdie probleem is aktueel in gevalle waar veranderlikes wat individueel goeie voorspellers is, neig om saam met ander voorspeller veranderlikes minder goed te voorspel as gevolg van kovariasie of oorfleueling in metings (Foshay en Misanchuk, 1981:352). In sulke gevalle moet bepaal word of die oorfleueling in meting ook oorfleueling in die veranderlikes self (of die basiese konstrakte ter sprake), verteenwoordig.

In die hantering van bogemelde probleem, moet gestreef word na ortogonaliteit. Dit word gewoonlik gedoen deur veranderlikes wat met mekaar korreleer en laag of glad nie met ander nie, as 'n groep te sien en as sodanig te hanteer. Hiervoor is, soos aangetoon in paragraaf 3.2.2, die metode van faktoranalise by uitstek geskik.

Dit spreek voorts vanself dat meetinstrumente wat gebruik word, 'n hoë mate van betroubaarheid en geldigheid moet hê (Ebel, et al., 1969:988).

Benewens die feit dat die probleme in verband met meetinstrumente die akkuraatheid van voorspelling kan beïnvloed, moet daar ook rekening met die aantal veranderlikes gehou word. Draper en Smith (1966:163) wys daarop dat 'n toename van die aantal voorspeller veranderlikes kan lei tot 'n afname in die akkuraatheid van voorspelling. Hierby is dit ook so dat die weglating van veranderlikes uit die voorspellingsmodel waardevolle inligting verlore kan laat gaan. In hierdie verband bestaan daar nie kriteria vir die aantal veranderlikes wat vir voorspelling gebruik moet word nie. Wat wel geld, is dat die navorser hom moet laat lei deur die teoretiese sinvolheid vir die insluiting van veranderlikes vir voorspellingsdoeleindes (Draper en Smith, 1966:163).

3.5.4 ANDER FAKTORE

Ander faktore wat die akkuraatheid van voorspelling beïnvloed, het te doen met die model ⁷ op grond waarvan voorspel word. Schoeman (1976) het byvoorbeeld aangetoon dat enkelvoudige en meervoudige voorspellingsmodelle nie met dieselfde mate van sukses na akademiese prestasie voorspel nie. Meervoudige voorspellingsmodelle lewer beter resultate omdat akademiese prestasie waarna voorspel word, kompleks is en daarom deur middel van meer as een veranderlike in 'n model voorspel word (vergelyk paragraaf 6.7.2.4). Daar word egter steeds met variasies van meervoudige voorspellingsmodelle geëksperimenteer ten einde met 'n toenemende mate van sukses te kan voorspel (vergelyk Schoeman, 1976:19 e.v.).

7 'n Model word gedefinieer as 'n voorstelling van bepaalde beginsels of hipoteses wat gedrag onderlê of verklaar (Gouws, et al., 1979:189). Dit word ook gesien as 'n vereenvoudigde struktuur wat gebruik word om verskynsels te verklaar; 'n soort replika van die konkrete en abstrakte werklikheid (Van Dalen, 1979:55).

Ongeag die variasie van tipes voorspellingsmodelle wat ontwikkel word, is daar twee verdere vereistes wat gestel word. Eerstens is dit noodsaaklik dat modelle geldig moet wees. Schoeman (1981:254) wys daarop dat dit standaardpraktyk is dat kruisvalidasie van resultate van navorsing gedoen word, soos dit blyk uit byvoorbeeld die werk van Schoeman (1976:178), Monteith (1983:54) en ander. Dit word aanvaar dat geldige resultate in 'n groter mate veralgemeenbaar en in die praktyk bruikbaar is.

Ten slotte dien dit daarop gewys te word dat voorspellingsvergelykings of -modelle altyd ruimte vir 'n foutfaktor laat (Kerlinger en Pedhazur, 1973:163). Daarby is dit ook so dat daar altyd 'n hoeveelheid onverklaarde variansie is, en dat 'n model of vergelyking nie al die variansie in 'n kriterium kan verklaar nie.

3.6 SAMEVATTING

Die doel met hierdie hoofstuk was om duidelikheid oor die wyse waarop akademiese prestasie voorspel word te kry. Om hierdie doel te bereik, is kennis van besondere metodes en statistiese tegnieke geneem. 'n Beskrywing is hiervan gegee en waar toepaslik, is metodes teen mekaar opgewees.

Die eerste vraagstuk met betrekking tot die wyse waarop akademiese prestasie voorspel word, behels die aanvanklike seleksie van voorspellerveranderlikes. Sodanige seleksie vind plaas wanneer die verkillende moontlike veranderlikes en meetinstrumente bestudeer word en wanneer daar onduidelikheid oor die konstrakte onderliggend aan die data waarmee voorspel gaan word, bestaan. Twee statistiese metodes is basies hiervoor geskik, te wete faktoranalise en hoofkomponentanalise. Uit 'n vergelyking blyk dit dat laasgenoemde meer gebruik word.

'n Tweede kwessie met betrekking tot die voorspelling van akademiese prestasie, hou verband met die groepering van leerlinge in subgroepe. Dit is as noodsaaklik beskryf omdat voorspelling daardeur met 'n groter mate van geldigheid kan

plaasvind. Hiervoor is die gebruik van moderatorveranderlikes geskik, alhoewel daar nie duidelikheid in die literatuur oor die suksesvolle gebruik van sodanige veranderlikes is nie.

Die derde kwessie, naamlik die vraag na die wyse waarop voorspel word, is in hierdie hoofstuk spesifiek beantwoord in terme van metodes van meervoudige regressieanalise. Twee besondere metodes, naamlik die alle-moontlike-deelversamelingmetode en die stapsgewyse regressiemetode is beskryf en teen mekaar opgeweeg. Daar is geargumenteer dat eersgenoemde meer geskik is. Ten slotte is aangetoon dat verskeie faktore die akkuraatheid van voorspelling kan beïnvloed.