HOOFSTUK IX

Die berekening van die gegewens

Die versamelde gegewens is, soos reeds gemeld (Hoofstuk V - VII) op kaartjies neergeskryf ten einde die verwerking daarvan te vergemaklik en akkuraatheid te verseker. Vir elke leeftydsgroep van elke ras is 'n kaartjie met 'n ander kleur gebruik.

Die berekening van die gegewens is met 'n Monroe en Facit rekenmasjien deur skrywer gedoen. Die Rekenkundige gemiddelde en Standaardafwyking vir al die proefpersone van die Blanke ras en vir al die Bantoes is bereken asook vir die leeftydgroepe 19-24 jaar, 25-29 jaar en 30-35 jaar en verder ook nog vir elke leeftydsjaar, Blankes en Bantoes afsonderlik.

Ten einde die prestasies van die rasse aanskoulik voor te stel, is grafieke geteken waarop die prestasie en die aantal proefpersone wat die besondere prestasie bereik het, aangedui word.

Die verskil tussen die prestasies van byvoorbeeld negentien- tot vier-en-twintigjarige Blanke en Bantoemans is bereken en die beduidenheid of nie van die verskil is daarna bepaal deur die t- toets toe te pas, dit wil sê Standaard Fout van die Verskil= $\sqrt{\frac{\sigma \text{ Blanke RG.}}{n} + \frac{\sigma \text{ Bantoe RG}}{n}}$, $t = \frac{\text{S.F van Verskil}}{\text{Werklike Verskil}}$

Die beduidendheid of nie van die verskille is dan op die 1%- of 5%-peil van beduidendheid afgelees. Indien die verskil op die 1%-peil betekenisvol was, is dit as hoogsbeduidend gemerk en as beduidend op die 5%-peil. Indien die werklike verskil so klein was dat die Standaard fout van die verskil gedeel deur die werklike verskil, kleiner as die grense vir die 5%-peil van betekenisvolheid was, is sodanige verskil as onbeduidend beskou.

Nadat die Rekenkundige gemiddeldes en standaardafwykings vir die verskillende toetsnommers bereken was en die beduidendheid van die verskille vasgestel is, het dit geblyk dat die

144/....

Blankes en die Bantoes wat vir hierdie ondersoek gebruik is, geen beduidende leeftydsverskil toon nie. In liggaamslengte en liggaamsgewig was daar egter hoogs beduidende verskille tussen die twee rassegroepe.

Die aanduiding van die verskil tussen byvoorbeeld die standverspringprestasie van die twee rassegroepe as agt duim, maak die interpretasie van die gegewens vaag. Ten einde hierdie probleem op te los, is besluit om van prestasieskale gebruik te maak. Aangesien sodanige skale vir volwassenes vir al die toetsnommers van hierdie ondersoek nie bestaan nie, moes skale vir die doel saamgestel word. Die samestelling van sodanige prestasieskale is reeds in Hoofstuk V (C b) uiteengesit. Die rou punte van die proefpersone kon vinnig in skaalpunte omgesit word. Die skaalpunte is ook op die prestasiekaart van die betrokke proefpersoon ingevul.

Die skaalpunte vir al ses die toetsnommers van die betrokke proefpersoon is bymekaargetel en die som is deur ses (aantal toetse) gedeel ten einde die prestasievermoë in persentasie uit te druk. Hierdie persentasie is as die Prestasie-indeks van die persoon beskou.

Met behulp van die T-skaal is dit moontlik om die verskil in die prestasies van die twee rasse in elke toetsnommer in persentasievorm uit te druk. Die persentasie-eenheid is maklik verstaanbaar en maak dit moontlik dat die verskil tussen die twee rasse in die verskillende toetsnommers ook met mekaar vergelyk kan word.

Ten einde die bydrae van elke toetsnommer tot die Prestasie-indeks te bereken, is elke toetsnommer met die Prestasie-indeks gekorreleer. Alhoewel die interkorrelasies tussen die ses toetsnommers en liggaamslengte en liggaamsgewig reeds by die samestelling van die toetsbattery bereken is, is die interkorrelasies tussen die betrokke toetsnommers vir die twee rasse weer eens afsonderlik bereken.

145/....

Die interkorrelasies het aangedui dat liggaamslengte en liggaamsgewig nie ŉ betekenisvolle verband met die prestasie-indeks by sowel die Blankes as die Bantoes het nie. Liggaamslengte en liggaamsgewig beïnvloed elk van die ses toetsnommers by die Blankes en by die Bantoes tot ŉ sekere hoogte. Op grond hiervan moes daar gepoog word om die invloed van die twee faktore op die toetsnommers uit te skakel aangesien daar ŉ beduidende verskil in liggaamslengte en liggaamsgewig tussen die twee rasse (Blankes en Bantoes) bestaan.

Die voor-die-handliggende oplossing vir die bogemelde probleem, was „random sampling". Dit beteken dat daar op ŉ ewekansige wyse ŉ sekere aantal Blankes en Bantoes uit die twee proefgroepe getrek word. Hierdie proefpersone word dan vergelyk wat hulle leeftyd, liggaamslengte en liggaamsgewig betref. Wanneer daar nog ŉ beduidende verskil in die drie faktore tussen die rasse bestaan, word ŉ volgende ewekansige trekking gedoen. Hierdie prosedure word herhaal totdat twee groepe gevind word wat vergelykbaar is wat leeftyd, lengte en gewig betref. Hierdie twee groepe se prestasies kon dan met mekaar vergelyk word. Die metode openbaar egter sekere tekortkominge, naamlik die langste en swaarste Bantoes word dan met die ligter en korter Blankes vergelyk. Die twee groepe is dus nie verteenwoordigend van die betrokke rasse nie. Die faktore leeftyd, lengte en gewig was gelyk; dog die vraag het ontstaan of die Bantoes oor die algemeen, as hulle so lank en so swaar as die Blankes is, dieselfde prestasies sou behaal.

Ten einde die bogenoemde tekortkominge uit te skakel, is die gemiddeld prestasie van die Bantoes in byvoorbeeld ŉ sekere toetsnommer geneem, die invloed van lengte en gewig op die prestasie is bereken en dan is bepaal hoe hoog die prestasie sou wees indien die Bantoes so swaar en so lank as die Blankes was. Dit is dus ŉ meervoudige regressievergelyking, gedoen volgens die metode van Doolittle (2).

Die prosedure wat vir die berekening van sodanige meervoudige regressievergelykings gevolg is, is die volgende:

1. Bereken die som van die waarnemings vir elke toetsnommer (wat gebruik gaan word) asook die som van die kwadrate van die waarnemings byvoorbeeld rugkrag = X_3, liggaamsgewig = X_1 en liggaamslengte = X_2. Ons kry dan SX_1, SX_2 en SX_3 asook SX_1^2, SX_2^2 en SX_3^2.

2. Bereken die som van die produkte van X_1 en X_2, X_1 en X_3, X_2 en X_3, dit wil sê SX_1X_2, SX_1X_3 en SX_2X_3.

3. Bereken nou verder $SX_1^2 - \frac{(SX_1)^2}{N}$, $SX_2^2 - \frac{(SX_2^2)^2}{N}$ en $SX_3^2 - \frac{(SX_3)^2}{N}$ (N = aantal proefpersone).

4. Bereken die kovariansie tussen X_1 en X_2, X_1 en X_3, X_2 en X_3 volgens die formule:

   $$\text{Kovariansie } X_1X_2 = SX_1X_2 - \frac{(SX_1)(SX_2)}{N}$$

5. Die matrix wat op hierdie wyse bereken is, kan nou neergeskryf word en die Doolittle-metode vir die omdraai van die matrix word gevolg. Duidelikheidshalwe sal die volledige verwerking hier aangetoon word. (Dit is die gegewens van hierdie ondersoek met betrekking tot rugkrag, lengte en gewig)

Die oorspronklike matrix sal as volg daar uitsien:

	X_1	X_2	X_3
X_1	66486.014+	5781.000	83783.000
X_2	5781.000	1617.887+	14208.000
X_3	83783.000	14208.000	871542,321+

+ Dui die getalle aan van $SX_1^2 - \frac{(SX_1)^2}{N}$ ensovoorts.

Ten einde die verwerking van die gegewens te vergemaklik, word al die getalle deur 1000 gedeel, dit wil sê die desimale punt skuif drie plekke na links. Om te verseker dat akkuraat gewerk word,. moet daar ses desimale syfers na die punt gebruik word.

6. Die verwerking van die meervoudige regressie verloop nou soos volg:

I. Skryf die getalle van X_1 en X_2 as volg neer:

	X_1	X_2	
X_1	66.486014	5.781000	A-matrix
X_2	5.781000	1.617887	

II. Trek nog drie kolomme langs die A-matrix. In die eerste twee kolomme word die B-matrix neergeskryf en in die derde kolom word die berekenings gekontroleer. Die twee kolomme van die B-matrix bestaan nou uit 1 en 0. Wanneer die getalle van die A- en die B-matrix bymekaar getel word, word dit in die tjekkolom neergeskryf. Al die berekenings wat op die A- en B-matrix uitgevoer word, moet ook op die tjekkolom gedoen word. Wanneer die getalle van die A- en die B-matrix vir 'n betrokke ry algebraïes opgetel word, moet dit dieselfde antwoord verstrek as wat in die tjekkolom voorkom, byvoorbeeld

	X_1	X_2	BX_1	BX_2	Tjek.
1) X_1	66.86014	5.781000	1.000000	0	73.267014
2) X_2	5.781000	1.617887	0	1.000000	8.398887

III. Deel al die getalle in die horisontale X_1-ry deur die getal in ry 1 X_2 en al die getalle in ry 2 deur die getal in ry 2 X_2. Die volgende antwoorde word dan verkry:

148/....

	X_1	X_2	BX_1	BX_2	Tjek.
3)	11.500780	1.000000	0.172980	0	12.673760
4)	3.573179	1.000000	0	0.618090	5.191269

IV. Trek die getal in reël 4 van die getal direk bokant die getal in reël 3 af. Die volgende antwoorde word verkry:

	X_1	X_2	BX_1	BX_2	Tjek.
5)	7.927601	0	0.172980	-0.618090	7.482491

V. Die C-matrix word nou bereken deur al die getalle in reël 5 te deel deur die getal in reel 5 X_1. Dit sal die eerste reël van die C-matrix gee, byvoorbeeld:

	X_1	X_2	BX_1	BX_2	Tjek.
6)	1.000000	0	0.021820	-0.077967	0.943853

VI. Die tweede reël van die C-matrix word verkry deur al die getalle in reël 6 te vermenigvuldig met die getal in reël 3 X_1 en hierdie produk dan van die getal in reël 3 BX_1 af te trek, byvoorbeeld 0.172980 - (0.021820 x 11.500780) = 0.077967. Dieselfde prosedure word vir die berekening van reël 7 BX_2 gevolg en vir die tjek-kolom:

	X_1	X_2	BX_1	BX_2	Tjek.
7)	0	1.000000	-0.177967	0.996681	1.818714

Die C-matrix is hiermee voltooi.

VII. In reël 8 BX_1 en 8 BX_2 word die kovariansie tussen X_1 en X_3 en X_2 en X_3 respektiewelik neergeskryf, dit is gi. byvoorbeeld:

	X_1	X_2	BX_1	BX_2	Tjek.
8) gi.	-	-	83.783000	14.208000	-
9) bi.			0.720390	6.207695	

VIII. Reël 9 (bi) word bereken deur die getal in reël 6 BX_1 te vermenigvuldig met die getal in reël 8 BX_1 en die produk by te tel by die produk van die getal in reël 7 BX_1 x 7 BX_2, dit is gelyk aan 0.720390. Reël 9 BX_2 word verkry deur (6 BX_2 x 8 BX_1) + (7 BX_2 x 8 BX_2) = 6.207695.

Hierdie getalle dui die invloed van die twee betrokke faktore (in hierdie geval liggaamsgewig en liggaamslengte op die rugkragprestasie van die Bantoemans) aan. In hierdie geval is dit dus 0.720390 liggaamsgewig (lb.) + 6.207695 liggamslengte (duim). Die regressievergelyking kan reeds op hierdie stadium neergeskryf word, naamlik:

rugkrag = 0.720390 (gewig in ponde) + 6.207695 (lengte in duime) + K (ŉ konstante faktor).

K word verkry deur die gemiddelde rugkragprestasie, liggaamslengte en liggaamsgewig van die betrokke groep persone in die geval, die Bantoes, in die bogenoemde vergelyking te vervang:

$$340.440 = 0.720390 \times 139.020 + 6.207695 \times 66.670 + K$$

$$K = 340.440 - 100.094 - 413.887 = -173.541$$

Met behulp van die liggaamslengte en die liggaamsgewig en deur die gebruikmaking van K kan die waarskynlike rugkrag van ŉ Bantoeman bereken word. Die vraag wat nou egter beantwoord moet word, is naamlik hoe betroubaar sal so ŉ voorspelling van die rugkrag wees?

IX. Die beduidendheid of die betroubaarheid van die regressielyn moet derhalwe bereken word. Die berekening geskied soos volg:

S.S.R. = 8 BX_1 x 9 BX_1 + 8 BX_2 x 9 BX_2 (in hierdie geval is dit 83.783000 x 0.720390 + 14.208000 x 6.207695 = 148.555366.

150/....

X. $S.S.E. = SX_3^2 - \frac{(SX_3)^2}{N} - S.S.R.$

Dit is 871.542321 - 148.555366 = 722.986955

XI. $S^2 = \frac{SS.E.}{N-P-1}$ Waar N = aantal proefpersone, P = aantal veranderlikes. In hierdie geval was N = 293 en P = 3, dit wil sê $S^2 = \frac{722.986955}{289} = 2.5016849.$

XII. Sbi. of die beduidendheid van liggaamsgewig en liggaamslengte se invloed op prestasie in die rugkrag word bereken deur die formule:

Sbi. = S^2 x getal in reël 6 BX_1

= 2.5016849 x 0.021820

= 0.545868 (Liggaamsgewig) Beduidend

Sbi. = S^2 x getal in reël 7 BX_2

= 2.5016849 x 0.896681

= 2.432133 Hoogs beduidend

Die beduidendheid van die faktore kan nou op die t-tabelle afgelees word deur die t-waarde vir N x Sbi. te verkry. Indien die t-waarde vir die getal op die 1%-peil beduidend is, is die invloed van die betrokke faktor op die rugkragsprestasie hoogs beduidend en op die 5%-peil is dit beduidend.

XIII. Die betroubaarheid van die regressievergelyking word as volg bereken:

Betroubaarheid van regressievergelyking:

Oorsprong van variasie	Vryheidsgrade	Som van Kwadrate	Gemiddelde kwadrate
Regressie	2	148.555366	74.277683
Fout	289	722.986955	2.5016849

$F = \frac{74.277683}{2.5016849} = 29.691.$

Die betroubaarheid van die regressievergelyking word op die F-tabelle afgelees. In hierdie geval was die betroubaarheid groter as 99%, met ander woorde hoogs beduidend.

L.W. S.S.R. = Som van kwadrate vir regressie

S.S.E. = Som van kwadrate vir fout.

Gemiddelde kwadraat word verkry deur die som van die kwadrate deur die aantal vryheidsgrade te deel.

XIV. Die verwerkings het aangedui dat die regressievergelyking hoogs beduidend is en derhalwe kan met gerustheid van die gegewens gebruik gemaak word.

Die konstante, soos in VIII uiteengesit, is bereken.

Nou kan bereken word wat die rugkragprestasie van die Bantoes sou wees indien hulle net so lank en net so swaar as die gemiddelde Blankes was:

$$\begin{aligned} \text{rugkrag} &= 0.720 \text{ liggaamsgewig} + 6.208 \text{ liggaamslengte} \\ &\quad -173.541 \\ &= 0.720 \times 170.9 + 6.208 \times 70.66 - 173.541 \text{ lb.} \\ &= 123.214 + 438.657 - 173.541 \text{ lb.} \\ &= 388.330 \text{ lb.} \end{aligned}$$

Die betekenis van liggaamslengte en liggaamsgewig in elke toetsnommer is op die voorgaande wyse bereken. Die besondere invloed van die twee faktore op die Prestasie-indeks is op 'n voortgelyke wyse vasgestel.

Op grond van die gegewens wat op die bogenoemde wyse bereken is, kon tot 99% van betroubaarheid bepaal word welke prestasies die Bantoes in die besondere toetsnommers sou behaal indien hulle net so lank en net so swaar as die Blankes sou wees.

Vir die berekenings is slegs met die gemiddeldes gewerk aangesien die gemiddelde prestasie van die Blankes met die gemiddelde prestasie van die Bantoes vergelyk is. Die resultate van die verwerking sal in Hoofstuk X bespreek word.