

## HOOFSTUK 4

### Gestandaardiseerde Verskille

As groepe vergelyk wil word ten opsigte van 'n kontinue karakteristiek (bv. akademiese prestasie, IK, cholesterolvlak, ens.) is die voor die handliggende metode om die rekenkundige gemiddeldes van die groepe te vergelyk. Beskou die volgende voorbeeld :

#### Voorbeeld 4.1:

Twee ewekansige steekproewe met groottes 100 en 200, uit populasies A en B se gemiddelde IK's is 110 en 107 met standaardafwykings, 10 en 12. Die gewone z-toets (t-toets vir groot steekproewe) lewer:

$$z = \frac{|110 - 107|}{\sqrt{\frac{10^2}{100} + \frac{12^2}{200}}} = 2,29 (p < 0,05).$$

Dit beteken dat die nulhipotese  $H_0 : \mu_A - \mu_B = 0$  verwerp word op 'n 5% betekenispeil, waar  $\mu_A$  en  $\mu_B$  die twee populasiegemiddeldes is. As ons dus sê dat  $\mu_A$  en  $\mu_B$  verskil, is die waarskynlikheid dat hierdie uitspraak foutief is, kleiner as 0,05. Die verskil in gemiddeldes van 3 eenhede is dus betekenisvol in die sin dat slegs 5% van herhaaldelike steekproewe van dieselfde groottes vanuit populasie A en B die bewering (van  $H_0$  se verwerping) nie gaan ondersteun nie. Dit dui egter nie noodwendig op 'n belangrike verskil nie. Omdat sielkundiges die IK-skaal ken, kan hulle oordeel dat 3 eenhede te klein is ten einde *praktiese betekenisvolheid* af te lei. Die verskil in gemiddeldes gee dus vir ons 'n maatstaf om praktiese betekenisvolheid of belangrikheid van die verskil tussen die populasies te karakteriseer. □

In 'n verdere voorbeeld verkry ons die gemiddelde houding van 'n eerstejaarsklas oor 'n aspek van hulle vak as 6,8, gemeet op 'n stanegeskaal (d.i. 'n genormeerde skaal met skaalwaardes 1-9, gemiddeld 5 en standaardafwyking omtrent 2). Omdat 6,8 amper 2 standaardafwykings bokant die gemiddelde skaalwaarde lê, sou mens voel om dit as prakties betekenisvol verskillend van 5 weer te gee, terwyl 'n gemiddelde houding van 'n ander groep van sê 4,6 duidelik te naby 5 is om as 'n verskil beskou te kan word. By beide bogenoemde voorbeelde het ons die skale "geken" in die sin dat ons geweet het wat die standaardafwykings was. Die 3 eenhede wat die twee gemiddeldes verskil het in konteks met standaardafwykings van 10 en 12, is dus klein in vergelyking met die 1,8 eenhede wat die eerstejaars se gemiddeld bokant 5 lê, want in lg. geval is die standaardafwyking 2. Dit lyk dus sinvol om die verskil in gemiddeldes te deel deur 'n standaardafwyking, sodat

$$\delta = \frac{\mu_1 - \mu_2}{\sigma^*} \quad (4.1)$$

'n effekgrootte-indeks lewer wat nie *skaal-afhanklik* is nie. Hier is  $\mu_1$  en  $\mu_2$  die twee populasies se gemiddeldes en  $\sigma^*$  een van vier moontlikhede as  $\sigma_1, \sigma_2$  die twee populasies se SA's is:

(a)  $\sigma^* = \sigma$  die gemeenskaplike SA van die twee populasies (d.i. waar aanvaar word dat beide populasies se SA  $\sigma$  het);

(b)  $\sigma^* = \sigma_1$  of  $\sigma_2$ , afhangende die uitgangspunt (bv. as populasie 1 die kontrole- of standaardgroep behandeling is, neem  $\sigma^* = \sigma_1$ );

(c)  $\sigma^* = \text{maks}(\sigma_1, \sigma_2)$  die maksimum van die twee SA's;

(d)  $\sigma^* = \sqrt{w_1\sigma_1^2 + w_2\sigma_2^2}$ , 'n geweegde SA, waar  $w_1$  en  $w_2$  die gewigte só is dat

$$w_1 + w_2 = 1.$$

In die volgende seksies gee ons aandag aan elkeen van bostaande moontlikhede.

#### 4.1 Cohen se $d$

Cohen (1969, 1977, 1988) neem *homogeniteit van populasievariansies* aan en neem die gemeenskaplike SA  $\sigma$  as noemer:

$$d \equiv \delta = \frac{\mu_1 - \mu_2}{\sigma} \quad (4.2)$$

Opmerking:

Ons sal vir wat volg probeer om effekgrootte-indekse vir *populasies* met Griekse letters (soos bv.  $\delta$ ) aan te dui, daarom verwys ons liefs na die simbool  $\delta$  hier i.p.v.  $d$  soos Cohen.

##### 4.1.1 Beraming van $\delta$ :

Wanneer ewekansige steekproewe uit die twee populasies getrek word, kan  $\delta$  nie direk bereken word nie, maar moet beraam word uit die steekproewe.

#### Voorbeeld 4.2:

Beskou soos in Voorbeeld 1 weer IK's maar van twee volledige populasies.

Gestel  $\mu_1 = 111$  en  $\mu_2 = 105$  met  $\sigma_1 = \sigma_2 = 10$ , dan is

$$\delta = \frac{111 - 105}{10} = 0,6.$$

Dit sê dus dat die verskil tussen die twee populasie gemiddeldes 0,6 standaardafwyking eenhede is. Let op dat indien ons bv. gemiddelde diastoliese bloeddrukke van 75 en 66 met gemeenskaplike SA van 15 verkry, dan is

$$\delta = \frac{75 - 66}{15} = 0,6,$$

wat ons op dieselfde manier kan interpreteer al was die eenhede nou in mmHg.

□

Die voor die handliggende beramer (Hedges se  $g$ ) is

$$g \equiv \hat{\delta} = \frac{\bar{x}_1 - \bar{x}_2}{s_p}, \quad (4.3)$$

waar

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 1}}, \quad (4.4)$$

die gepoelde SA en

$\bar{x}_1, \bar{x}_2$ : die steekproefgemiddeldes

$n_1, n_2$ : die steekproefgroottes

$s_1, s_2$ : die steekproefstandaardafwykings.

Opmerking:

Ons sal deurgaans 'n beramer van 'n populasie-indeks wat op steekproewe gebaseer is met 'n kappie aandui, bv.  $\hat{\delta}$  as 'n beramer van  $\delta$ .

'n Alternatiewe formule in terme van die Student t-statistiek (of z-statistiek as  $n_1$  en  $n_2$  groot is) is:

$$\hat{\delta} = t \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (4.5)$$

Die beramer  $\hat{\delta}$  is egter *positief sydig*, wat beteken dat  $\hat{\delta}$  gemiddeld  $\delta$  oorberaam. Om dit reg te stel, veral by kleiner  $n_1$  en  $n_2$ , kan die volgende gekorrigeerde effekgrootte-indeks gebruik word (kyk Hedges en Olkin, 1985:81):

$$\hat{\delta}_a = \left(1 - \frac{3}{4n - 9}\right) \hat{\delta}, \quad (4.6)$$

waar  $n = n_1 + n_2$ .

**Voorbeeld 4.3:**

(Kline, 2004a: 104-105)

Gestel die steekproefgemiddeldes van aanlegtellings van mans en dames is 13,0 en 11,0 met variansies van 7,5 en 5,0. Tabel 4.1 gee die resultate van die t-toets en effekgrootte-indekse vir verskillende steekproefgroottes ( $n_1 = n_2$ ).

Vir  $n_1 = n_2 = 5$  word  $\hat{\delta}$  en  $\hat{\delta}_a$  aanpas as volg bereken:

$$s_p = \sqrt{\frac{(5-1)7,5 + (5-1)5}{5+5-2}} = \sqrt{\frac{30+20}{8}} = \sqrt{6,25} = 2,5$$

$$\hat{\delta} = \frac{13-11}{2,5} = 0,8$$

$$\hat{\delta}_a = \left(1 - \frac{3}{4 \times 10 - 9}\right) \hat{\delta}$$

$$= \left(1 - \frac{3}{31}\right) 0,8$$

$$= 0,903 \times 0,8 = 0,722.$$

**Tabel 4.1**

**Steekproefgroottes**

<b>Statistiek</b>	<b>5</b>	<b>15</b>	<b>30</b>
<b>t-toets:</b>			
t	1,26	2,19	3,1
vg*	8	28	58
p	0,243	0,037	0,003
<b>Gestandaardiseerde verskil:</b>			
$\hat{\delta}$	0,80	0,80	0,80
$\hat{\delta}_a$	0,72	0,78	0,79

---

vg\* : vryheidsgrade  $n_1 + n_2 - 2$

Uit die tabel is dit duidelik dat die t-waardes groter en p-waardes kleiner word met toename in steekproefgroottes. Die beraamde effekgrootte-indeks  $\hat{\delta}$  bly egter dieselfde omdat dit nie van  $n_1$  en  $n_2$  afhang nie. Verder gee die aangepaste beramer  $\hat{\delta}_a$  kleiner waardes wat al hoe nader aan  $\hat{\delta}$  kom soos  $n_1$  en  $n_2$  groter word. □

#### 4.1.2 Vertrouensintervalle vir $\delta$

As die aanname van normaalpopulasies met gelyke variansies gemaak word, het

$$t_{v, nsp} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{1/n_1 + 1/n_2}} \quad (4.7)$$

'n *nie-sentrale t-verdeling* met  $v = n_1 + n_2 - 2$  vryheidsgrade (vg) en *nie-sentraliteitsparameter*

$$nsp = \frac{(\mu_1 - \mu_2)}{\sigma \sqrt{1/n_1 + 1/n_2}} = \delta \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \quad (4.8)$$

Cumming & Finch (2001) gee 'n volledige bespreking van die nie-sentrale  $t$ -verdeling en die gebruik van hulle ESCI-sagteware en veral die NonCentral tNET-program (kan afgelaai word van handleiding se webblad) help verder om die gebruik daarvan te verstaan.

Uit (4.7) is dit moontlik om 'n  $100(1 - \alpha)\%$  - vertrouensinterval (VI) vir  $nsp$  te bepaal deur van 'n rekenaarprogram gebruik te maak (kyk bv. Kline, 2004a: table 4.6). As inset tot die program is slegs  $t$ ,  $vg$  en  $\alpha$  nodig. Hieruit kan m.b.v. (4.8) vir  $\delta$  'n VI bepaal word. Die SAS-program **VI\_delta** (op die webblad van die handleiding beskikbaar) gebruik hierdie metode. Zou (2007) gee ook hierdie program.

Meer besonderhede van die metode om die VI te bepaal word in bylae A gegee.

'n Benaderde VI vir  $\delta$  word deur Hedges & Olkin (1985:86) gegee:

$$\text{Ondergrens: } \delta_o = \hat{\delta}_a - z_{\alpha/2} \hat{\sigma}_\delta$$

$$\text{Bogrens: } \delta_B = \hat{\delta}_a + z_{\alpha/2} \hat{\sigma}_\delta, \quad (4.9)$$

waar  $z_{\alpha/2}$  die  $100(1 - \alpha/2)$ -de persentiel van 'n standaardnormaalverdeling is en waar

$$\hat{\sigma}_\delta^2 = \frac{n_1 + n_2}{n_1 n_2} + \frac{\hat{\delta}_a^2}{2(n_1 + n_2)}, \quad (4.10)$$

die beraamde asimptotiese variansie van  $\hat{\delta}_a$  is. Uit 'n simulasiestudie van Hedges & Olkin(1985: 86-88), blyk dit dat die korrekte oordekkingswaarskynlikheid verkry word as gelyke steekproefgroottes van 10 en groter beskou word.

Verdere benaderde VI's vir  $\delta$  word gegee deur Wu et. al. (2006). Volgens simulasiestudies deur Wu et al. uitgevoer, gee die volgende metode 'n VI met korrekte oordekkingswaarskynlikheid by steekproewe so klein soos 5:

Laat  $a = \sqrt{4 + 2n_1/n_2 + 2n_2/n_1}$ , dan is 'n variansie-stabiliseringstransformasie van

d:  $h(d) = \sqrt{2} \ln \left[ d/a + \sqrt{d^2/a^2 + 1} \right]$ , sodat  $\sqrt{n}(h(d) - h(\delta))$  benaderd

standaardnormaal verdeeld is. Die  $100(1 - \alpha)\%$  VI vir  $h(\delta)$  is dus

$$(O_{h(\delta)}, B_{h(\delta)}) = h(d) \pm z_{\alpha/2} / \sqrt{n}.$$

Die inverse transformasie is:

$\delta = \frac{a(e^{\sqrt{2}h(\delta)} - 1)}{2e^{h(\delta)/\sqrt{2}}}$ , wat meebring dat die  $100(1 - \alpha)\%$  VI vir  $\delta$  is:

$$\left[ \frac{a(e^{\sqrt{2}O_{h(\delta)}} - 1)}{2e^{O_{h(\delta)}/\sqrt{2}}}, \frac{a(e^{\sqrt{2}B_{h(\delta)}} - 1)}{2e^{B_{h(\delta)}/\sqrt{2}}} \right].$$

#### Voorbeeld 4.4:

In voorbeeld 4.1, as aanvaar word dat die populasievariensies gelyk is, word beramer  $\hat{\delta}$  bereken as:

$$s = \sqrt{\frac{(100-1)10^2 + (200-1)12^2}{100 + 200 - 2}} = \sqrt{\frac{38556}{298}} = 11,37$$

$$\hat{\delta} = (110 - 107) / 11,37 = 0,26.$$

Uit (4.5) volg  $t = \hat{\delta} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = 0,26 \times \sqrt{\frac{100 \times 200}{100 + 200}} = 2,12$ .

Let op dat die  $t$ -waarde kleiner is as  $z=2,29$  (in Voorbeeld 4.1) waar geen aanname van homogeniteit van variensies gemaak word nie. A.g.v. die groot  $n_1$ - en  $n_2$ -waardes, is  $\hat{\delta}_a$  ook 0,26. Kies ons  $\alpha=0,05$  en let op dat die vryheidsgrade  $\nu = 100 + 200 - 2 = 298$  is, dan is die presiese 95% VI vir  $\delta$  wat met die rekenaar bereken is (kyk Bylae A): (0,018 ; 0,500). Dit beteken dat die populasie effekgrootte-indeks so klein soos 0,018 en so groot soos 0,5 kan wees



met 'n waarskynlikheid van 0,95. Anders gestel, indien herhaaldelik ewekansige steekproewe van dieselfde groottes uit populasies A en B getrek sou word, behoort 95% van hulle  $VI$ 's die onbekende  $\delta$  - waarde in te sluit.

Die benaderde 95%  $VI$  vir  $\delta$  kan ook bereken word. Hier is  $z_{0,025} = 1,96$  en

$$\hat{\sigma}_{\delta}^2 = \frac{100+200}{100 \times 200} + \frac{0,26^2}{2(100+200)} = 0,01500 + 0,000113$$

$$= 0,015113$$

$$\delta_O = 0,26 - 1,96\sqrt{0,0151} = 0,26 - 0,241 = 0,019$$

$$\delta_B = 0,26 + 0,241 = 0,501.$$

Die interval (0,019 ; 0,501) is baie naby die presiese interval hierbo bereken.  $\square$

Opmerking:

Die benaderde interval sal al hoe meer verskil van die presiese interval soos  $n_1$  en  $n_2$  kleiner word. Tabel 4.2 gee beide intervale vir die drie stelle steekproefgroottes van Tabel 4.1:

**Tabel 4.2**

<b><math>VI</math></b>	<b><math>n_1 = n_2 = 5</math></b>	<b><math>n_1 = n_2 = 15</math></b>	<b><math>n_1 = n_2 = 30</math></b>
Presies	(-0,523 ; 2,072)	(0,048 ; 1,539)	(0,271 ; 1,324)
Benaderd	(-0,559 ; 1,999)	(0,039 ; 1,521)	(0,264 ; 1,316)

Die benadering word duidelik beter soos  $n_1$  en  $n_2$  vergroot. In die geval van Voorbeeld 4.3 met  $n_1 = n_2 = 30$  gee Kline (2004a: Tabel 4.6) die 95%  $VI$  vir  $\delta$  as (0,271; 1,324).

Die EXCEL – sigblad EffectSizeCalculator. xls maak die bererekening van  $\hat{\delta}, \hat{\delta}_a$  en die benaderde  $VI$  moontlik deur slegs die gemiddeldes,  $SA$ 's en groepgroottes te spesifiseer. Dis beskikbaar by hierdie handleiding se webblad. 'n Ander handige EXCEL – sigblad, wat ook op die handleiding se webblad

besikbaar is, is Effect\_Sizes\_Spreadsheet.xls en dit bereken Cohen se  $d$  met verskillende kombinasies van soos bv.  $t$ ,  $n_1$  en  $n_2$ . Die 90% ( $\alpha=0,10$ ) en 99% ( $\alpha=0,01$ )  $VI$ 's is respektiewelik (0,355 ; 1,239) en (0,016 ; 1,489). Let op dat die interval nouer word as  $\alpha$  groter word, maar wyer word vir kleiner  $\alpha$ -waardes.

#### 4.1.3 Teen-nul effekgroottes

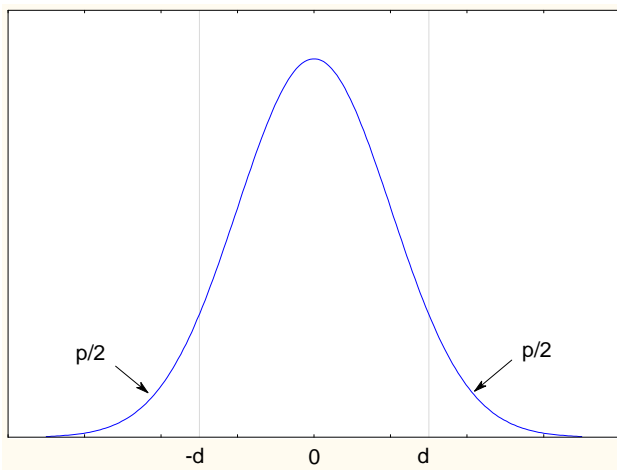
Gestel dat by die effekgrootte  $\delta$  in (4.2)  $\sigma$  bekend is en dit daarom beraam word met

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sigma}.$$

Onder  $H_o : \delta = 0$ , geld dan dat onder  $H_o$ :

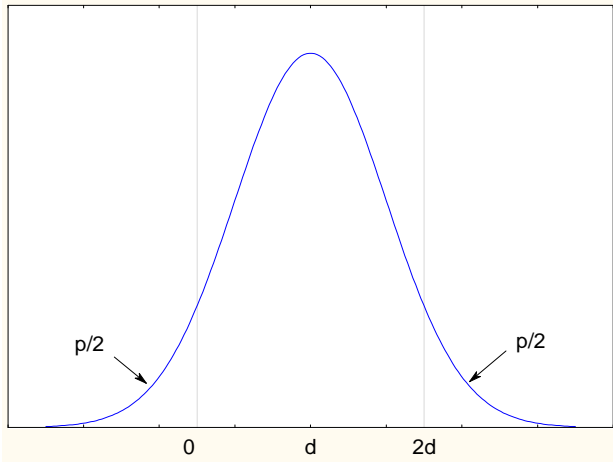
$P\left(\left|\frac{\bar{X}_1 - \bar{X}_2}{\sigma}\right| > d\right) = p$ , die tweekantige p-waarde van bv. 'n z-toets (onder die

aanname van normaalverdeelde populasies of groot steekproewe uit enige populasie). Die volgende figuur illustreer dit i.t.v. die steekproefverdeling van  $(\bar{X}_1 - \bar{X}_2) / \sigma$ :



Indien die hipotese  $H_o : \delta = d$  getoets word, dan geld onder  $H_o$  :

$$P_{\delta=d} \left( \left| \frac{\bar{X}_1 - \bar{X}_2}{\sigma} \right| > d \right) = P_{\delta=d} \left( 0 < \frac{\bar{X}_1 - \bar{X}_2}{\sigma} < 2d \right) = p.$$



In bostaande figuur het die verdeling van  $\frac{\bar{X}_1 - \bar{X}_2}{\sigma}$  net met  $d$  eenhede regs geskuif sonder verandering van vorm. Dit gebeur net as die verdeling se standaardafwyking nie van  $\delta$  afhang nie en ons sê dan dat die effekgrootte  $d$  translasië-invariant is.

Dit beteken dus dat die waarskynlikheid dat die effekgrootte so klein (en kleiner) as 0 is, gelyk is die waarskynlikheid om so groot (en groter) as  $2d$  te wees, indien  $\delta = d$ . Hierdie interval  $(0, 2d)$  word die **nul-teen-nul interval** vir  $\delta$  genoem terwyl  $2d$  die **teen-nul waarde** van die effekgrootte is.

Die  $(1-\alpha)$  100% vertrouensinterval (VI) vir  $\delta$  is:

$$d \pm z_{\alpha/2} \sqrt{\frac{n_1 + n_2}{n_1 n_2}},$$

wat met  $(1-\alpha)100\%$  waarskynlikheid  $\delta$  oordek. Dis duidelik dat daar geen direkte verband is tussen die VI en die nul-teen-nul interval nie, behalwe dat hulle ooreenstem as

$$d = z_{\alpha/2} \sqrt{\frac{n_1 + n_2}{n_1 n_2}} .$$

Wanneer  $H_0: \delta=0$  verwerp word, op 'n  $\alpha$ -peil, is  $d$  gewoonlik groot, is  $(0,2d)$  wyd en kan dit as 'n VI met vertrouenskoëffisiënt groter as  $1 - \alpha$  beskou word.

**Voorbeeld 4.4 (vervolg):**

Die nul-teen-nul interval is  $(0, 2d) = (0; 0,52)$ , wat beteken dat die waarskynlikheid dat die effekgrootte  $\delta$  nul is netso groot is as dat dit 0,52 is. Let op dat dit amper ooreenstem met die 95%VI van  $(0,018; 0,500)$  is wat tevore verkry is.

Opmerkings:

- (a) Die interval  $(0,2d)$  volg uit die simmetrie van die steekproefverdeling van  $d$  en omdat  $d$  translasië-invariant is.
- (b) Indien  $\sigma$  onbekend is en uit die steekproewe beraam moet word, verkry ons 'n beramer soos bv.  $\hat{\delta}$  (in 4.3), wat 'n nie-sentrale t-verdeling het onder die aanname van normaliteit van die populasies. Hierdie beramer vir  $\delta$  is nie meer translasië-invariant nie en is positief sydig. Dit bring mee dat net benaderd geld dat onder  $\delta=d$  die waarskynlikhede dat die effekgroottes 0 (of kleiner) en  $2d$  (of groter) sal wees, gelyk is. Hierdie benadering word beter met groter steekproewe en kleiner waardes van  $d$ .

Grissom & Kim (2005: 65-67) bespreek die teen-nul effekgrootte in groter diepte. Ons sal in latere hoofstukke weer terugkom hierop.

#### 4.1.4 Interpretasie van die teen-nul effekgrootte

Rosenthal et.al (2000: 14) gee die volgende twee voorbeelde om die nut van die teen-nul effekgrootte te illustreer:

1. Gestel  $\hat{\delta} = 0,3$  en die toets  $H_0 : \mu_1 - \mu_2 = 0$  het  $p = 0,15$  opgelewer. Die benaderde teen-nul waarde is dus  $2\hat{\delta} = 0,6$  wat mens daaraan herinner dat  $\delta$  netsowel so groot soos 0,6 kon wees as so klein soos 0, niesteenstaande die gematigde p-waarde van 0,15. Ons word dus gewaarsku om nie noodwendig 'n statisties nie-betekenisvolle resultaat as 'n nulwaarde van die effek (d.i.  $\delta = 0$ ) te interpreteer nie. Dit gee die probleem in die linker-onderste sel van Tabel 1 van Hoofstuk 1.
2. In 'n baie groot kliniese proef word 'n nuwe duur medisyne om liggaamstemperatuur te verlaag, in vergelyking met aspirien, getoets.

'n Duidelike statistiese betekenisvolle effek ten gunste van die nuwe middel word verkry deurdat  $p = 0,013$  (eenkantig) hoewel  $\hat{\delta} = 0,03$ . Dit gee die nul-teen-nul interval van  $(0,00; 0,06)$ , wat 'n  $1-2(0,013) = 0,974$ , d.i. 97,4% vertrouensinterval gee. Omdat  $\delta$  met hoë waarskynlikheid maar so hoog soos 0,06 kan wees, is dit duidelik nie van praktiese waarde om die duur medikasie te regverdig nie. Dit wys dat 'n statisties-betekenisvolle nie-nul effekgrootte nie noodwendig dui op 'n wetenskaplike belangrike effek nie – die probleem in die regter-boonste sel van Tabel 1 van Hoofstuk 1.

Die nul-teen-nul interval gee dus soortgelyk aan 'n vertrouensinterval vir die effekgrootte, meer inligting om die werklike effekgrootte te beoordeel.

## 4.2 Glass se $\Delta$

Indien daar 'n eksperimentele- en 'n kontrole populasie is, stel Glass (1976) die effekgrootte-indeks

$$\Delta = \frac{(\mu_E - \mu_K)}{\sigma_K} \quad (4.11)$$

voor, waar  $\mu_E$  en  $\mu_K$  die eksperimentele- en kontrole populasiegemiddeldes en  $\sigma_K$  die kontrole populasie standaardafwyking is.

### Voorbeeld 4.5:

By voorbeeld 4.2, gestel  $\mu_E=111$  met SA van 10 en  $\mu_K=105$  met  $\sigma_K=15$ , dan is

$$\Delta = (111 - 105) / 15 = 0,4.$$

Hierdie indeks waarde is kleiner as die waarde van 0,6 in Voorbeeld 4.2, omdat deur 'n ander SA gedeel is, nl.  $\sigma_K$  in plaas van  $\sigma$ , die gemeenskaplike SA.  $\square$

### 4.2.1 Beraming en vertrouensinterval vir $\Delta$ .

In die geval van ewekansige steekproewe uit die populasies kan die *beramer*

$$\hat{\Delta} = \frac{\bar{X}_E - \bar{X}_K}{s_K} \quad (4.12)$$

gebruik word, met die steekproefgemiddeldes en die kontrolegroep se standaardafwyking in die plek van die populasieparameters.

Volgens Kline (2004a: 108) is die asimptotiese standaardfout van  $\hat{\Delta}$ :

$$\hat{\sigma}_{\Delta} = \sqrt{\frac{n_E + n_K}{n_E n_K} + \frac{\hat{\Delta}^2}{2(n_K - 1)}}, \quad (4.13)$$

sodat 'n benaderde  $100(1-\alpha)\%$  VI vir  $\Delta$  se grense is:

$$\begin{aligned}\Delta_O &= \hat{\Delta} - Z_{\alpha/2} \hat{\sigma}_{\Delta} \\ \Delta_B &= \hat{\Delta} + Z_{\alpha/2} \hat{\sigma}_{\Delta}\end{aligned}\tag{4.14}$$

**Voorbeeld 4.6:**

In Voorbeeld 4.3, neem die dames as die kontrolegroep met groepgrootte 20 en die mans se groepgrootte as 30.

$$\begin{aligned}\hat{\sigma}_{\Delta} &= \sqrt{\frac{30+20}{30 \times 20} + \frac{0,894^2}{2(20-1)}} = \sqrt{0,0833 + 0,0210} \\ &= 0,323\end{aligned}$$

‘n 90% VI se grense:

$$\begin{aligned}\Delta_O &= 0,894 - 1,645 \times 0,323 = 0,894 - 0,531 \\ &= 0,363\end{aligned}$$

$$\Delta_B = 0,894 + 0,531 = 1,425$$

Dit beteken dat met ‘n 90% waarskynlikheid kan  $\Delta$  so laag soos 0,363 wees, maar so hoog soos 1,425. □

By hierdie indeks word nie noodwendig homogeniteit van variansies veronderstel nie en word die kontrole populasie as uitgangspunt geneem. Soms veroorsaak ‘n behandeling of ingreep dat ‘n eksperimentele groep se gemiddelde verhoog (of verlaag) en daarmee saam ook die variasie. In so ‘n geval sou Glass se  $\Delta$  die aangewese indeks wees om te gebruik.

### 4.3 Effekgrootte- indekse by heterogene variansies

Gestel dit kan *nie* veronderstel word dat  $\sigma_1 = \sigma_2$  *nie*, dan is daar verskeie opsies om die noemer van die indeks te kies.

#### 4.3.1 Keuse van enige populasie SA:

$$\hat{\Delta}_1 = (\bar{x}_1 - \bar{x}_2) / s_1,\tag{4.15}$$

as beramer vir

$$\Delta_1 = (\mu_1 - \mu_2) / \sigma_1 . \quad (4.16)$$

Netso dien

$$\hat{\Delta}_2 = (\bar{x}_1 - \bar{x}_2) / s_2 \quad (4.17)$$

As beramer vir

$$\Delta_2 = (\mu_1 - \mu_2) / \sigma_2 . \quad (4.18)$$

Hierdie effekgrootte-indekse se waardes kan grootliks verskil as  $\sigma_1$  en  $\sigma_2$  se waardes baie verskil.

Die VI's vir  $\Delta_1$  en  $\Delta_2$  kan netsoos by  $\Delta$  bepaal word met

$$\hat{\sigma}_{\Delta_1} = \sqrt{\frac{n_1 + n_2}{n_1 n_2} + \frac{\hat{\Delta}_1^2}{2(n_1 - 1)}} .$$

#### Voorbeeld 4.7

(a) Gestel  $\mu_1 - \mu_2 = 5$  en  $\sigma_1 = 20$ ,  $\sigma_2 = 5$ , dan is  $\Delta_1 = 5/20 = 0,25$  en

$\Delta_2 = 5/5 = 1,0$  wat beteken dat  $\Delta_1$  vier maal kleiner is as  $\Delta_2$ .

(b) By Voorbeeld B, Hoofstuk 3 se J/P voorkeurtelling sou

$\Delta_1 = (91,08 - 70,07)/28,6 = 0,73$ .

Die waarde van 0,73 is effens kleiner as die waarde van  $\Delta_2 = 0,81$  wat verkry is as die dosente as uitgangspunt geneem word.

□

Daar is nie 'n korrekte of verkeerde keuse tussen  $\Delta_1$  en  $\Delta_2$  nie, want dit hang net af watter populasie se variasie die uitgangspunt is. Dit sou dus die beste wees om beide te rapporteer.

#### 4.3.2 Keuse van grootste SA

Om meer konsekwent te wees, kan besluit word om altyd te deel deur die grootste variansie (kyk ook Steyn, 1999, 2000 en Ellis & Steyn, 2003), sodat



$$\Delta_m = (\mu_1 - \mu_2) / \sigma_{maks} = \min(\Delta_1, \Delta_2) \text{ waar } \sigma_{maks} = \text{maksimum}(\sigma_1, \sigma_2). \quad (4.19)$$

As beramer sou geld:

$$\hat{\Delta}_m = (\bar{x}_1 - \bar{x}_2) / s_{maks} \quad (4.20)$$

In Voorbeeld 4.7(a) is  $\Delta_m = \Delta_1 = 0,25$ , want  $\sigma_1$  is die grootste van die twee SA's, terwyl in Voorbeeld 4.7(b) is  $\Delta_m = \Delta_1 = 0,73$  omdat  $\sigma_1 = 28,6$  hier groter was as  $\sigma_2 = 25,93$ .

### 4.3.3 Geweegde SA

Wanneer die populasiestandaardafwykings merkbaar verskil, kan die vierkantswortel van die geweegde gemiddelde van die twee variansies geneem word, sodat

$$\delta_g = \frac{\mu_1 - \mu_2}{\sqrt{W_1\sigma_1^2 + W_2\sigma_2^2}}, \quad (4.21)$$

met  $W_1 + W_2 = 1$ .

Die populasiegroottes kan gebruik word as gewigte, d.i.  $W_1 = N_1 / (N_1 + N_2)$  en  $W_2 = 1 - W_1$ , waar  $N_1$  en  $N_2$  die populasiegroottes is. Indien die gewigte as gelyk ( $W_1 = W_2 = 1/2$ ) geneem word, stel Cohen (1977: 44) 'n spesiale geval van (4.21) voor as indeks, nl.

$$\delta_c = \frac{\mu_1 - \mu_2}{\sqrt{(\sigma_1^2 + \sigma_2^2) / 2}}. \quad (4.22)$$

#### **Voorbeeld 4.8:**

By Voorbeeld B, Hoofstuk 3, is uit Tabel B.1,  $W_1 = 254 / (254 + 28) = 0,9$  en  $W_2 = 0,1$ . Nou word vir J/P die effekgrootte-indeks:

$$\delta_g = \frac{91,08 - 70,07}{\sqrt{0,9 \times 28,6^2 + 0,1 \times 25,93^2}} = \frac{21,01}{\sqrt{736,16 + 67,24}} = 21,01 / 28,34 = 0,74. \text{ Dit gee}$$

'n effens groter waarde as  $\Delta_m = \Delta_1$ .

Sou ons Cohen se gelyke gewigte neem, word 
$$\delta_c = \frac{91,08 - 70,07}{\sqrt{(28,6^2 + 25,93^2) / 2}}$$

= 21,01 / 27,3 = 0,77, wat nader aan die waarde van  $\Delta_2 = 0,81$  is, waar die dosente as uitgangspunt geneem is. □

Wanneer  $\delta_g$  beraam moet word uit steekproewe, word

$$\hat{\delta}_g = (\bar{x}_1 - \bar{x}_2) / \sqrt{W_1 s_1^2 + W_2 s_2^2}. \quad (4.23)$$

Indien populasiegroottes onbekend is, of dis moeilik om die gewigte te kies, kan die beramer

$$\hat{\delta}_c = (\bar{x}_1 - \bar{x}_2) / \sqrt{\frac{1}{2}(s_1^2 + s_2^2)} \quad (4.24)$$

gebruik word of ook  $\hat{\Delta}_m$  in (4.20). Beide hierdie beramers is sydig vir  $\delta_g$ . Steyn (1999) het in simulاسies gevind dat  $\hat{\Delta}_m$  vir  $\delta_g$  onderberaam, maar nie meer as 0,08 as  $\delta_g = 0,5$  is en  $n_2 > 10, n_1 \leq 1,5n_2$  en  $\sigma_1 < 2\sigma_2$ . Met dieselfde beperkings op  $n_1, n_2, \sigma_1$  en  $\sigma_2$ , het  $\hat{\Delta}_m$  vir  $\delta_g$  nie meer as 0,13 onderberaam as  $\delta_g = 0,8$ .

#### Voorbeeld 4.9:

Gestel in Voorbeeld 4.1 was die steekproefgroottes proporsioneel aan die populasiegroottes, sodat  $W_1 = 100 / 300 = \frac{1}{3}$  en  $W_2 = \frac{2}{3}$ . As aanvaar word dat  $\sigma_1 \neq \sigma_2$ , kan die indeks

$$\begin{aligned} \hat{\delta}_g &= (111 - 105) / \sqrt{\frac{1}{3} \times 10^2 + \frac{2}{3} \times 12^2} \\ &= 6 / \sqrt{33,33 + 96} = 0,527 \end{aligned}$$

'n beraming gee vir  $\delta_g$ . Met geen kennis oor  $W_1$  en  $W_2$  sou

$\hat{\delta}_c = 6 / \sqrt{\frac{1}{2}(10^2 + 12^2)} = 6 / \sqrt{122} = 0,543$  of  $\hat{\Delta}_m = 6/12 = 0,5$  as 'n konserwatiewe beramings kon dien.

□

#### 4.4 Effekgrootte-indeks vir afhanklike groepe

Gestel soortgelyke metings word op dieselfde mense op verskillende geleenthede gedoen. Hierdie groepe metings is dan afhanklik en word gewoonlik verkry as een of ander ingreep of behandeling gedoen word tussen die metingsgeleenthede. Dis dan belangrik om die effek van die ingreep te kan meet.

##### Voorbeeld 4.10:

In Voorbeeld A, Hoofstuk 3 gee die volgende tabel die beskrywende statistiek van POMS-depressie voor en na terapie aan die  $n=25$  hartpasiënte in die eksperimentele groep:

Voor		Na		Verskil			
$\bar{x}_1$	$s_1$	$\bar{x}_2$	$s_2$	$\bar{x}_D$	$s_D$	$t$	$p$
18,00	12,26	8,08	9,84	9,92	13,38	3,71	0,001

□

Steekproefdata soos in Voorbeeld 4.10 se gemiddelde van die verskille van na- en voortoetstellings  $x_1 - x_2$ , nl.  $\bar{x}_D$  behoort die effek van die terapie uit te wys, maar soos tevore deel ons dit deur 'n standaardafwyking om 'n effekgrootte-indeks te verkry.

Die populasie-indeks is dus

$$\delta_D = \mu_D / \sigma^*, \quad (4.25)$$

met  $\mu_D$  die gemiddeld van die populasieverdeling van verskille en  $\sigma^*$  weer soos aan die begin van die hoofstuk. Die keuse  $\sigma^* = \sigma_1 = \sigma_2$ , maak die aanname van gelyke standaardafwykings by die twee meetgeleenthede. Gewoonlik met 'n basislyn of voormeting en dan 'n nameting nadat 'n ingreep gedoen is, is hierdie aanname nie noodwendig realisties nie. 'n Beter keuse sou  $\sigma_1$  wees omdat dit 'n soort verwysingsvariasie verteenwoordig (die basislyn, voortoets) (kyk Kline, 2004a: 105). Dit lewer dan

$$\Delta_D = \mu_D / \sigma_1, \quad (4.26)$$

met beramer

$$\hat{\Delta}_D = \bar{x}_D / s_1, \quad (4.27)$$

ook "Becker se g" genoem.

'n Tweede moontlikheid is om  $\sigma^* = \sigma_D$ , die SA van die populasieverdeling van verskille, te kies. Gewoonlik is  $\sigma_D$  heelwat kleiner as  $\sigma_1$  of  $\sigma_2$ , veral as daar 'n hoë korrelasie tussen die twee metings is. Dit lewer die effekgrootte-indeks

$$\delta_D = \mu_D / \sigma_D, \quad (4.28)$$

met beramer

$$\hat{\delta}_D = \bar{x}_D / s_D. \quad (4.29)$$

Hierdie beramer is soos by die onafhanklike-groepe geval met  $\hat{\delta}$ , sydig vir  $\delta_D$ .

Volgens Hedges & Olkin (1985: 79) is

$$\hat{\delta}_{D,a} = \left(1 - \frac{3}{4n-5}\right) \hat{\delta}_D, \quad (4.30)$$

'n aangepaste beramer, wat veral vir n klein, gebruik moet word.

Die probleem is egter dat  $\sigma_D$  (en  $s_D$ ) nie meer die oorspronklike skaal waarin metings by elke geleentheid gedoen is, se variasie weergee nie, maar die

verskille. Cumming & Finch, (2001: 569 – 570) bespreek die volgende voorbeeld, waar  $\delta_D$  vergelyk word met die indeks

$$\delta'_D = \mu_D / \sigma, \quad (4.31)$$

(waar  $\sigma = \sigma_1 = \sigma_2$  aanvaar word), met beramer

$$\hat{\delta}'_D = \bar{x}_D / s_P, \quad (4.32)$$

waar  $s_P = \sqrt{\frac{1}{2}(s_1^2 + s_2^2)}$  (4.33)

as beramer vir  $\sigma$  gebruik word.

**Voorbeeld 4.11** (Cohen, 1988: 49):

By 'n verbale vermoë-toets is dit bekend dat  $\mu = 100$ ,  $\sigma = 15$ . Gestel die effek van 'n gesonde ontbyt word bepaal op verbale vermoë deur elkeen van 'n groep kinders met en sonder gesonde ontbyt te toets. Die gemiddelde verskil van 4,1 in die guns van die toets na gesonde ontbyt, terwyl die SA van die verskiltelling 7,7 was. Die beraamde effekgroottes is:

$$\hat{\delta}_D = 4,1 / 7,7 = 0,53, \text{ terwyl } \delta'_D = 4,1 / 15 = 0,27.$$

(waar ons  $\sigma = 15$  gebruik omdat dit bekend is).

In terme van die oorspronklike skaal waarop verbale vermoë gemeet word, is die indeks omtrent die helfte van wat dit is in terme van die variasie van

verskilmettings. Uit Voorbeeld 4.10 is  $\hat{\delta}_D = 9,92 / 13,38 = 0,74$  en

$$\hat{\delta}_{D,a} = \left(1 - \frac{3}{4 \times 25 - 5}\right) 0,74 = 0,72. \text{ Verder is } \hat{\delta}'_D = \frac{9,92}{\sqrt{(12,26^2 + 9,84^2) / 2}} = 9,92 / 11,12$$

= 0,89. Hier is die teenoorgestelde waar as in (a) hierbo: die variasie van verskille tussen voor- en nametings was groter as die variasie op die oorspronklike skaal, vandaar  $\hat{\delta}'_D > \hat{\delta}_D$ . Hierdie verskynsel van groter variasie by verskille is te wyte aan 'n swak korrelasie tussen voor- en nametings (in hierdie geval was dit 0,28). Omdat dit egter lyk of die voor- en nametings se SA's verskil, sou  $\Delta_D$  dalk die aangewese indeks gewees het, wat beraam kon word deur

$\hat{\Delta}_D = 9,92 / 12,26 = 0,81$ . Indien die oorspronklike skaal as uitgangspunt geneem word, sou  $\Delta_D$  dus die aangewese indeks gewees het, eerder as  $\delta_D'$ .

Bespreking:

Die volgende word aanbeveel:

- Indien daar aanduidings is dat die variasie groter (of kleiner) word by die twee metings (a.g.v. die ingreep): gebruik  $\Delta_D$  en  $\hat{\Delta}_D$ .
- Indien homogeniteit van variansies aanvaar kan word en die variasie van die oorspronklike metings as uitgangspunt geneem wil word: gebruik  $\delta_D'$  en  $\hat{\delta}_D'$ .
- In die res van die gevalle, gebruik  $\delta_D$  en  $\hat{\delta}_{D,a}$ .

Slegs  $\hat{\delta}_{D,a}$  is *nie 'n sydigse beramer nie*. Veral as  $n$  klein is, moet dit in ag geneem word dat  $\hat{\Delta}_D$  en  $\hat{\delta}_D'$  vir  $\Delta_D$  en  $\delta_D'$  kan oor- of onderberaam.

Opmerking:

Die effekgrootte- indekse  $\Delta_D$  en  $\delta_D'$  word in die afhanklike-groepe geval op dieselfde wyse bereken en beraam as  $\Delta$  en  $\delta$  vir onafhanklike groepe. As  $\hat{\delta}_D'$  egter uit die t-statistiek bereken wil word, is formule (4.5) nie die korrekte een nie, maar (Kline, 2004a: 107):

$$\hat{\delta}_D' = t_D \sqrt{\frac{2s_D^2}{n(s_1^2 + s_2^2)}}, \quad (4.34)$$

waar  $t_D$  die t-statistiek van afhanklike groepe is.

#### 4.4.1 Vertrouensintervalle vir $\delta_D$ en $\delta_D'$

Volgens Johnson et. al. (1995: 513) is die asimptotiese variansie van  $\hat{\delta}_{D,a}$ :

$$\hat{\sigma}_{\delta_{D,a}}^2 = \frac{1}{n} (1 + \frac{1}{2} \hat{\delta}_{D,a}^2), \quad (4.35)$$

sodat die  $100(1-\alpha)\%$  VI vir  $\delta_D$  die volgende benaderde grense lewer:

$$\begin{aligned}\delta_{D,O} &= \hat{\delta}_{D,a} - z_{\alpha/z} \hat{\sigma}_{\delta_{D,a}} \\ \delta_{D,B} &= \hat{\delta}_{D,a} + z_{\alpha/z} \hat{\sigma}_{\delta_{D,a}} .\end{aligned}\tag{4.36}$$

Vir  $\delta'_D$  gee Kline (2004a) die asimptotiese variansie van  $\hat{\delta}'_D$  as :

$$\hat{\sigma}_{\delta'_D}^2 = \frac{2(1-r_{12})}{n} + \frac{\hat{\delta}'_D{}^2}{2(n-1)}\tag{4.37}$$

waar  $r_{12}$  die korrelasiekoëffisiënt tussen  $x_1$  en  $x_2$  is, waaruit die  $100(1-\alpha)\%$  VI vir  $\delta'_D$  se grense benaderd volg as:

$$\begin{aligned}\delta'_{D,O} &= \hat{\delta}'_D - z_{\alpha/2} \hat{\sigma}_{\delta'_D} \\ \hat{\delta}'_{D,B} &= \hat{\delta}'_D + z_{\alpha/2} \hat{\sigma}_{\delta'_D}\end{aligned}\tag{4.38}$$

Let op die VI'e (4.36) en (4.38) net benaderd geld en dat  $n$  redelik groot moet wees (sê  $n > 30$ ).

Soos in paragraaf 4.1.2, bestaan hier ook onder die aanname van 'n normaalverdeelde populasie 'n presiese VI vir  $\delta_D$  uit die feit dat

$$\frac{\bar{x}_D - \mu_D}{s_D / \sqrt{n}}$$

'n nie-sentrale  $t$ -verdeling met  $v=n-1$  vryheidsgrade en nie-

sentraliteitsparameter  $\sqrt{n}\delta_D$  besit. Kyk Bylae A vir meer besonderhede van die metode. Rekenaarprogramme (bv. die CIdeltaNET-program van die ESCI-sagteware of die SAS-program VI\_delta\_D) is op die webbladsy beskikbaar om die berekenings te doen. Algina & Keselman (2003) het met behulp van simulaties aangetoon dat hierdie VI die korrekte oordekkingswaarskynlikheid gee vir  $\delta_D = 0,0$  (0,2) 1,6 en  $\rho = 0,0$  (0,2) 0,8 (waar  $\rho$  die korrelasie tussen die twee afhanklike metings is).

Uit (4.34) is dit duidelik dat  $\hat{\delta}'_D$  ook nog van  $s_1$  en  $s_2$  afhang, sodat die verdeling daarvan nie in terme van 'n nie-sentrale t-verdeling geneem kan word nie, maar baie meer ingewikkeld is. Daarom word hier volstaan met die benaderde VI's vir  $\delta'_D$ .

#### Voorbeeld 4.12:

By Voorbeeld 4.11 is

$$\hat{\sigma}_{\delta_{D,a}} = 0,72,$$

sodat  $\hat{\sigma}_{\delta_{D,a}}^2 = \frac{1}{25}(1 + \frac{1}{2} \times 0,72^2) = 0,0504$ , en

$$\delta_{D,O} = 0,72 - 1,96\sqrt{0,05} = 0,28$$

en

$$\delta_{D,B} = 0,72 + 1,96\sqrt{0,05} = 1,16.$$

Die presiese 95% VI is: (0,292 ; 1,180). Die benaderde VI vir  $\delta_D$  verskil nie veel van hierdie presiese VI nie, selfs met  $n = 25$  nie baie groot nie.

Verder is  $\hat{\sigma}_{\delta'_D}^2 = \frac{2(1-0,28)}{25} + \frac{0,89^2}{2 \times 24} = 0,074$ ,

sodat

$$\delta'_{D,O} = 0,89 - 1,96 \times \sqrt{0,074} = 0,359$$

en  $\delta'_{D,B} = 0,89 + 1,96 \times \sqrt{0,074} = 1,423$ . □

#### 4.5 Teen-nul waardes vir ander effekgroottes om twee gemiddelde te Vergelyk

Soos reeds aangetoon, is  $2\hat{\delta}$  net die benaderde teen-nul waarde vir  $\hat{\delta}$  omdat  $\hat{\delta}$  se standaardfout 'n funksie is van  $\delta$ . Dit is ook waar vir al die ander effekgroottes nl.  $\hat{\delta}_a, \hat{\Delta}, \hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_m, \hat{\sigma}_g, \hat{\sigma}_c, \hat{\Delta}_D, \hat{\sigma}_D, \hat{\sigma}_{D,a}$  en  $\hat{\delta}'_D$ .



Wanneer die steekproewe groot is, is hierdie benadering goed en so ook wanneer  $\delta$  klein is.

#### 4.6 Riglynewaardes vir effekgrootte-indekse gebaseer op gestandaardiseerde verskille

Die onmiddellike vraag wat opkom nadat 'n effekgrootte-indeks bereken is, is wanneer is dit groot genoeg of wanneer is dit klein? Cohen (1969, 1977, 1988) gee die volgende riglynewaardes vir  $\delta$  (sy 'd'):

- $|\delta| = 0,2$  : klein effekgrootte
- $|\delta| = 0,5$  : medium effekgrootte
- $|\delta| = 0,8$  : groot effekgrootte.

Opmerkings:

Let op dat indien  $\delta$  negatief is, die teken geïgnoreer word by die riglynewaardes.

Omdat  $\delta = (\mu_1 - \mu_2) / \sigma$ , word die teken van  $\delta$  negatief indien  $\mu_1 < \mu_2$ , sodat dit die *rigting* van die effek gee. As ons nie in die rigting belangstel nie, kan ons altyd  $\delta$  bereken deur die kleinste gemiddeld van die grootste af te trek.

Die riglynewaardes is van toepassing ook vir die indekse

$\Delta$ ,  $\Delta_1$ ,  $\Delta_2$ ,  $\Delta_m$ ,  $\delta_g$ ,  $\delta_D$ ,  $\Delta_D$  en  $\delta'_D$  met al hulle beramers.

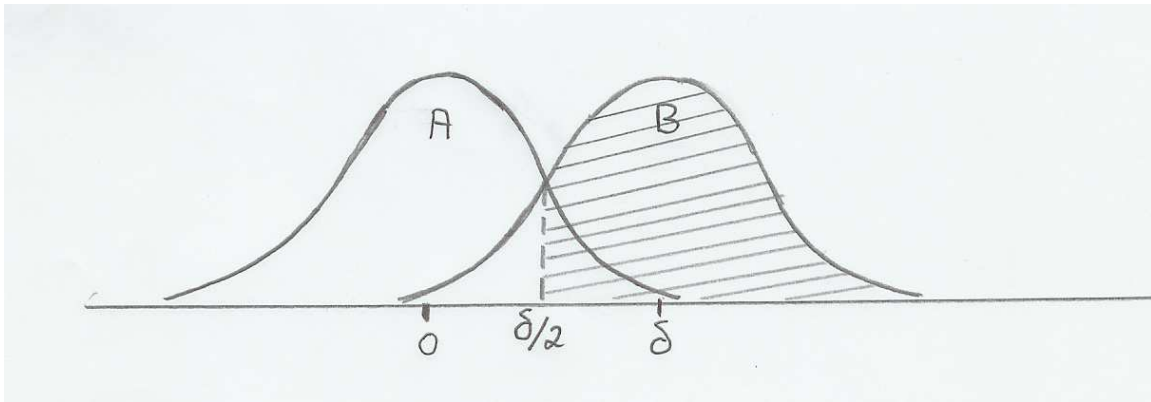
Cohen (1969, 1977, 1988) gee verder die interpretasie van 'n

gestandaardiseerde verskil in terme van twee populasies se oorvleueling.

Beskou die eenvoudige geval van twee normaalpopulasies A en B met verskillende gemiddeldes en gemeenskaplike variansie  $\sigma^2$ . Die effekgrootte om  $\mu_A$  en  $\mu_B$  te vergelyk is

$$\delta = \frac{\mu_B - \mu_A}{\sigma}.$$

Cohen definieer die proporsie oorvleueling  $U_2$  as die proporsie in populasie B wat die waarde halfpad tussen die gemiddeldes van die twee populasies oorskry (wat ekwivalent is met die proporsie van populasie A wat onder hierdie waarde lê). Die volgende figuur illustreer dit mooi, wanneer, sonder verlies aan algemeenheid, die gemiddeldes van populasie A en B as 0 en  $\delta$  geneem word.



In terme van  $\delta$ -eenhede vergelyk ons die verdelings  $N(0;1)$  en  $N(\delta,1)$  en is

$$U_2 = P\left(Z \leq \frac{1}{2}\delta\right).$$

Die waarskynlikheid van 'n misklassifikasie is

$$P_{MK} = 1 - U_2 = P\left(Z > \frac{1}{2}\delta\right) \quad (4.39)$$

Cohen heg die volgende betekenis aan sy riglynwaardes, met voorbeelde:

#### 4.6.1 Klein effek:

Dit kom voor in nuwe navorsingsareas waar metings gedoen word sonder behoorlike eksperimentele kontrole om agtergrondsveranderlikes uit te skakel. Daarom is die verskil in gemiddeldes klein relatief tot die foutvariasie (0,2 en selfs kleiner). Voorbeelde is die verskil in gemiddelde IK van tweeling en nie-tweeling, die verskil in gemiddelde lengte van 15- en 16- jarige dogters

(12,5mm, waar  $\sigma = 53,3$  mm). Hier is die waarskynlikheid van misklassifikasie ( $P_{MK}$ ) 0,46.

#### 4.6.2 Medium effek:

Dit is groot genoeg om deur die blote oog waargeneem te word. Afhangende van die konteks, kan sulke verskille as òf klein òf groot beskou word. Voorbeelde is die verskil in gemiddelde IK van klerikale- en semi-geskoolde werkers, verskil tussen 14- en 18- jarige dogters se gemiddelde lengtes (25mm, met  $\sigma = 51$  mm).  $P_{MK}$  is hier 0,41.

#### 4.6.3 Groot effek:

Dit is 'n belangrike verskil en strook met die algemene opvatting van 'n definitiewe verskil. Voorbeelde is die verskil in gemiddelde IK's van persone met PhD-grade en eerstejaarstudente of tussen 13- en 18- jarige dogters se gemiddelde lengtes. Cohen waarsku self dat die terme "klein", "medium" en "groot" relatief is, nie net tot mekaar nie, maar ook tot die veld van navorsing. Feinstein(1999:2569) – kyk sy motivering verder aan in paragraaf 5.1.2 – en Fleiss(1981) gee die grens 0,6 as 'n praktiese afsnypunt vir die effekgrootte-indeks  $\delta$ , wat ook met 'n empiriese studie van Burnand et.al. (1990) bevestig word. Daar is 'n  $P_{MK} = 0,34$  kans vir misklassifikasie by groot effekte.

Dit is goed om voor die gebruik van die riglyne op die volgende waarskuwings te let.

#### 4.6.4 Waarskuwings (Kline, 2004a: 132)

1. Die riglyne is nie op empiriese wyse verkry nie. Vandaar Cohen se bostaande waarskuwing.
2. Die waardes 0,2 vir "klein", 0,5 vir "medium" en 0,8 vir "groot" moet nie te streng toegepas word nie. Bv. waardes van 0,49 en 0,51 moenie as "klein" en

“medium” geneem word deur 0,5 as ‘n afsnypunt te gebruik nie, maar beide as “medium”, synde beide in die *omgewing* van 0,5 te wees. Dis waarom 0,2, 0,5 en 0,8 as *riglynwaardes* bekend is sodat dit deurgaans so gebruik behoort te word.

3. Soos Cohen hierbo waarsku, geld die definisies van groottes nie noodwendig vir alle navorsingsvelde nie. Wat in een veld miskien as ‘n groot effek beskou word, kan as matig deurgaan in ‘n ander veld, ens.
4. Cohen se riglynwaardes mag meer geskik wees in *nie-eksperimentele studies*, terwyl by studies waar van eksperimentele ontwerpe gebruik gemaak word, die foutvariëansie kleiner behoort te wees, sodat groter riglynwaardes dalk meer relevant gaan wees.
5. In gevestigde navorsingsvelde kan m.b.v. meta-analise, sistematies tussen klein en groot effekgroottes onderskei word en moet hierdie kennis liefers gebruik word om as riglyne te dien. Slegs in nuwe velde waar min gepubliseer is, sou die gebruik van Cohen se riglyne ‘n uitweg wees.
6. ‘n Groter voordeel by die rapportering van effekgroottes is om dit met vorige studies s’n te vergelyk en nie soseer met die arbitrêre riglynwaardes van Cohen nie.

#### **4.7 Praktiese betekenisvolheid**

Die vraag is nou: wanneer is die effekgrootte-indekse se waarde substansieel, betekenisvol of belangrik? Ons vat bg. terme saam deur na *praktiese betekenisvolheid* te verwys sodat dit onderskei kan word van betekenisvolheid in ‘n statistiese sin (wanneer die nulhipotese verwerp word). Hierdie kwessie vra vir kundigheid in die bepaalde navorsingskonteks. Kline (2004a: 134) noem as voorbeeld dat vir die verskil in gemiddelde lengtes van mans en vrouens se  $\delta$ -waarde 2 kan wees sonder dat dit van belang is vir bv. ‘n fisioloog maar van kritiese belang by voertuigveiligheid waar lugsakke in motors nie vir vrouens ‘n

groter risiko moet meebring as vir mans, soos by vroeëre modelle van motors die geval was, nie.

By vroeëre fases van navorsing kan groter waardes van die indeks voorkom as by latere fases – wat ook die praktiese betekenisvolheid bepaal.

Deel van die uitdagings van 'n nuwe navorsingsveld behoort te wees om standarde te stel vir praktiese betekenisvolheid.

Die term *kliniese betekenisvolheid* is ook hier van toepassing. Volgens Kline (2004a: 135) is dit wanneer 'n ingreep 'n sinvolle verskil maak. Op groepsvlak is 'n *kriterium-kontras* 'n manier om hierdie betekenisvolheid te standaardiseer. Dit verteenwoordig 'n bekende verskil in die uitkomsveranderlike waarin belang gestel word, soos bv. tussen pasiënte met dieselfde siekte wat die ergste simptome het, versus dié met die minste simptome. Indien so 'n kriterium-kontras bv. 0,8 is voor enige behandeling toegepas word en die effekgrootte na 'n behandeling 0,4 is, beteken dit dat behandeling se effek is die helfte van die afstand tussen die ergste en minste simptome. Dit kan klinies betekenisvol wees al is dit nie statisties betekenisvol nie, en omgekeerd.

Volgens Kirk (1996) is die evaluering van praktiese betekenisvolheid 'n *kwitatiewe beslissing*, want dit behoort op 'n navorser se kennis van die veld staat te maak, sonder dat dit sy persoonlike waardes of samelewingsbetrokkenheid hoef te weerspieël.

Vervolgens gee ek van die resultate van voorbeelde in hierdie hoofstuk tesame met die interpretasie daarvan n.a.v. riglynwaardes en praktiese betekenisvolheid.

1. In Voorbeeld 4.2 was die waardes van  $\delta$  in beide die gevalle van verskille in gemiddelde IK's en diastoliese bloeddrukke 0,6. Volgens Cohen se riglynwaardes, is dit 'n medium effek. Die sielkundige sou kon redeneer dat

aangesien beide populasiegemiddeldes bokant die gemiddeld van 100 lê en net 6 skaalpunte verskil, dit nie 'n belangrike verskil is nie. Die bloeddrukverlagings-effek was dalk vir die biokinetikus uit sy ervaring genoeg om te dui op 'n suksesvolle oefenprogram.

2. In voorbeeld 4.3 waar  $n_1 = n_2 = 5$ , was  $\hat{\delta}_a = 0,72$  die beraamde effekgrootte. Ons sou kon oordeel dat dit 'n groot effek is, omdat dit amper 0,8 is en daarom dalk prakties betekenisvol. Kyk ons egter na die 95% VI wat (-0,523 ; 2,072) is, is ons versigtiger om dit te sê -  $\delta$  kon wel so hoog soos 2,07 wees maar ook so laag soos -0,52 (met 95% waarskynlikheid). Daar kon dus selfs 'n omgekeerde verskil gewees het met medium effek! Wanneer  $n_1 = n_2 = 30$ , gee die VI (0,271 ; 1,324) steeds nog aanduiding van klein tot groot effekte.

3. In voorbeeld 4.11 was die beraamde effekgroottes  $\hat{\delta}'_D = 0,89$  en  $\hat{\delta}_D = 0,74$ . In terme van die oorspronklike skaal van depressietellings is hier 'n groot effek wat kan dui op 'n prakties betekenisvolle afname na die ingreep. Kyk ons egter na die effekgrootte waar die verskil-skaal as uitgangspunt geneem word, neig dit na 'n medium effek. Die 95% VI van  $\delta'_D$ : (0,359; 1,423) dui op 'n medium tot groot effek, as Cohen se riglyne gebruik word, terwyl die VI vir  $\delta_D$  nl. (0,292 ; 1,180) ons versigtig maak om selfs van 'n medium effek te praat.