

Statistiek B – C1 - RESPONS

Oefening 1

Is er een verschil in interesse voor techniek ('Interest.voor') tussen leerlingen met een vader die wel of niet een diploma hoger onderwijs behaalde ('Dipvader')?

- a) Hoe groot is het effect van die hogere studies?
- b) Geef het effect visueel weer.

OEENING 1 a

```
> leveneTest(Techniek$Interest.voor, Techniek$Dipvader)
```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group   1  0.0614 0.8043
      2172
```

→ p is > 0.05 , dus WEL gelijke binnengroepvariantie

```
> t.test(Techniek$Interest.voor~Techniek$Dipvader, var.equal=TRUE)
```

Two Sample t-test

```
data:  Techniek$Interest.voor by Techniek$Dipvader
t = 3.9534, df = 2172, p-value = 7.95e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.08344689 0.24771611
sample estimates:
mean in group 0 mean in group 1
 3.022044      2.856463
```

→ $p < 0.05$: kans dat H_0 opgaat is kleiner dan 5%

⇒ We verwachten in de populatie WEL een verschil in interesse in techniek afhankelijk van het diploma van de vader.

→ gemiddelde interesse groep 0 (vader geen diploma HO) = 3.03

gemiddelde interesse groep 1 (vader wel diploma HO) = 2.86

```
> d(Techniek$Interest.voor, Techniek$Dipvader)
```

```
[1] 0.1701453
```

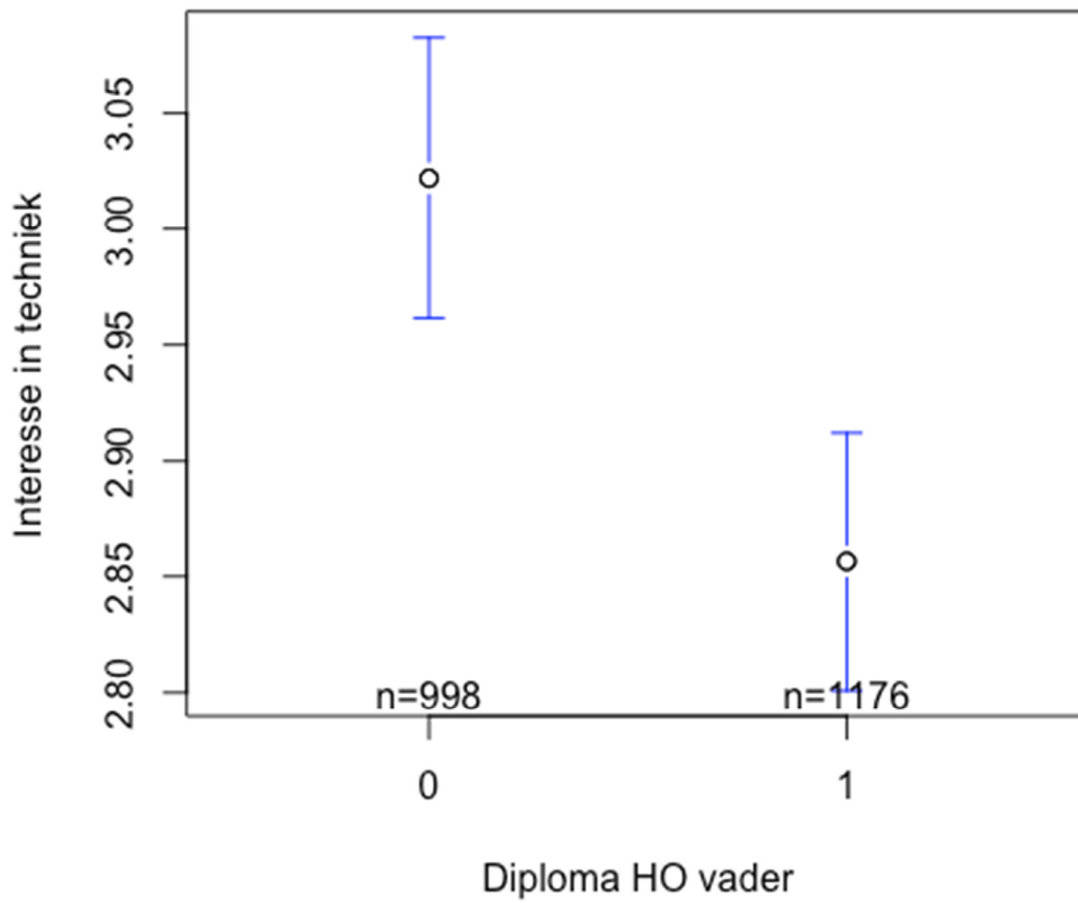
→ $d < 0.2$: het gaat om een verwaarloosbaar effect

CONCLUSIE:

Uit de steekproef blijkt dat leerlingen wiens vader geen diploma hoger onderwijs behaalde ($\mu = 3.02$) geïnteresseerder zijn in techniek dan hun peers met een "gediplomeerde vader" ($\mu = 2.86$). Dit verschil is bovendien statistisch significant ($p < 0.05$) en mogen we dus doortrekken naar de populatie. Op basis van de Cohen's d kunnen we echter stellen dat het om een verwaarloosbaar effect gaat ($d = 0.17$). Leerlingen wiens vader geen diploma heeft, scoren slechts 0.17 SD hoger op interesse in techniek dan leerlingen wiens vader wel een diploma hoger onderwijs behaalde. Het gaat dus om een verwaarloosbaar effect.

OEFFENING 1 b

```
> errorbar(Techniek$Interest.voor~Techniek$Dipvader,xlab="Diploma HO vader",  
+         ylab="Interesse in techniek")
```



Oefening 2

Is er een verschil in interesse voor techniek ('Interesse.voor') tussen eerste- en tweedejaars leerlingen ('Studiejaar')?

- Hoe groot is het effect van studiejaar?
- Geef het effect visueel weer.

OEfening 2 a

```
> leveneTest (Techniek$Interesse.voor, Techniek$Studiejaar)
```

Levene's Test for Homogeneity of Variance (center = median)

```
      Df F value Pr(>F)
group  1      5.5 0.0191 *
      2288
```

Signif. codes: 0 '***' 1e-03 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

→ $p < 0.05$, dus GEEN gelijke binnengroepvariantie

```
> t.test (Techniek$Interesse.voor~Techniek$Studiejaar, var.equal=FALSE)
```

Welch Two Sample t-test

data: Techniek\$Interesse.voor by Techniek\$Studiejaar

t = 13.277, df = 2260.6, p-value < 2.2e-16

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

0.4433075 0.5969520

sample estimates:

mean in group 1 mean in group 2

3.211843 2.691714

→ $p < 0.05$: kans dat H_0 opgaat is kleiner dan 5%

⇒ We verwachten in de populatie WEL een verschil in interesse in techniek tussen beide studiejaaren.

→ gemiddelde interesse leerjaar 1 = 3.21

gemiddelde interesse leerjaar 2 = 2.69

```
> d(Techniek$Interesse.voor, Techniek$Studiejaar)
```

```
[1] 0.5557404
```

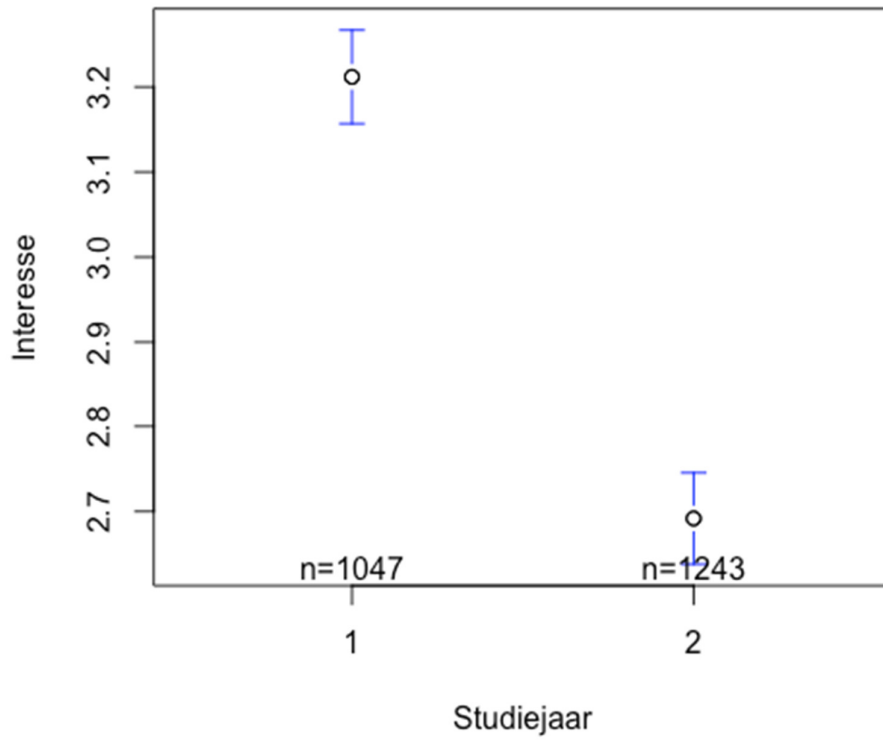
→ d tussen 0.5 en 0.7: het gaat om een medium effect

CONCLUSIE:

Eerstejaars (gemiddelde = 3.21) zijn in onze steekproef geïnteresseerder in techniek dan tweedejaars (gemiddelde = 2.69). We kunnen dit effect bovendien doortrekken naar de populatie ($p < 0.05$). Het gaat hierbij om een medium effect ($d = 0.55$). Eerstejaars scoren gemiddeld 0.55 SD hoger op interesse in techniek dan tweedejaars.

OEFENING 2 b

```
> errorbar(Techniek$Interest.voor~Techniek$Studiejaar,xlab="Studiejaar",  
+          ylab="Interesse in techniek")
```



Oefening 3

Is de interesse van leerlingen verschillend in het begin ('Interest.voor') en het einde van een trimester ('Interest.na')?

- a) Hoe groot is het effect van een trimester technieklessen?
- b) Geef het effect visueel weer.

OEFFENING 3 a

```
> t.test(Techniek$Interest.voor, Techniek$Interest.na, paired=TRUE)
```

Paired t-test

```
data: Techniek$Interest.voor and Techniek$Interest.na
t = 1.9696, df = 2289, p-value = 0.049
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.000149209 0.067973062
sample estimates:
mean of the differences
      0.03406114
```

→ $p < 0.05$: kans dat H_0 opgaat is kleiner dan 5%

⇒ We verwachten in de populatie WEL een verschil in interesse in techniek tussen het begin en het einde van het trimester.

→ gemiddeld verschil in interesse tussen begin en einde trimester = 0.03

```
> dpaired(1.9696, 2290)
[1] 0.04115857
```

→ $d < 0.2$: het gaat om een verwaarloosbaar effect

CONCLUSIE:

Hoewel de p -waarde wijst op een statistisch significant effect ($p < 0.05$), blijkt uit de effectgrootte ($d = 0.04$) dat we overheen het trimester geen verschil in interesse voor techniek kunnen vaststellen bij leerlingen.

OEFFENING 3 b

```
> errorbar2(Techniek$Interest.voor,Techniek$Interest.na,  
+          xlab="Meting voor en na",ylab="Interesse")
```

