9. Testen van meetresultaten.

Opgave 9.1 **Testen van het uit de steekproef geschatte gemiddelde t.o.v. μ**

a $\bar{x} = 24{,}5$ kg en $s = 1{,}0$ kg

b

$$t = \left| (\mu - \bar{x}) \right| \cdot \frac{\sqrt{n}}{s} = |25 - 24{,}5| \times \frac{\sqrt{5}}{0{,}9} = 1{,}21$$

c 2,5 %

d $v = n - 1 = 5 - 1 = 4$

tabel: $t_{\text{kritisch}} = 2{,}78$.

e $1{,}21 < 2{,}78$

f de nulhypothese wordt aangenomen

g Het gewicht voldoet aan de specificatie van 25 kg met een betrouwbaarheid van 95 %

Opgave 9.2 **Paracetamol**

a via de website:

One sample *t* test results
P value and statistical significance:
The two-tailed P value equals 0.0045
By conventional criteria, this difference is considered to be very statistically significant.

zelf berekenen

Nulhypothese:
Het gewicht voldoet aan de specificatie, de waarde wijkt niet significant af van 200 g
H_0: μ = 200 g

Alternatieve hypothese
Het gewicht voldoet niet aan de specificatie, de waarde wijkt significant af van 200 g
H_0: μ ≠ 200

$$t = \left| (\mu - \bar{x}) \right| \cdot \frac{\sqrt{n}}{s} = \left| |200 - 192| \times \frac{\sqrt{6}}{4} \right| = 4{,}90$$

$v = n - 1 = 6 - 1 = 5$
tabel: 95%; tweezijdig, → $t_{\text{kritisch}} = 2{,}57$
4,90 > 2,57, dus de nulhypothese wordt verworpen
Het gewicht is significant lager dan 200 g met een betrouwbaarheid van 95 %

b bij een eenzijdige test is de $t_{\text{kritisch}} = 2{,}02$; deze afwijking is nog groter, dus de conclusie is hetzelfde

Opgave 9.3 **Nieuwe machine**

a eenzijdig, je wilt bewijzen dat hij sneller is.

b **Nulhypothese:**

Het aantal van de nieuwe machine verschilt niet van de oude

H_0: μ = 250

Alternatieve hypothese

Het aantal van de nieuwe machine is groter dan van de oude

H_1: μ > 250

$$t = \left|(\mu - \bar{x})\right| \cdot \frac{\sqrt{n}}{s} = \left|250 - 265\right| \times \frac{\sqrt{10}}{6} = 7,91$$

$v = n - 1 = 10 - 1 = 9$

tabel: $t_{\text{kritisch}} = 1,83$

7,91 > 1,83 dus de nulhypothese wordt afgewezen en de alternatieve dus aangenomen; de nieuwe machine werkt significant sneller dan de oude.

Opgave 9.4 **Slootwater**

a plaatje II past het best

b de meetserie van de partner lijkt nauwkeuriger

c **Eigen metingen**

Nulhypothese:

Het "werkelijke" gehalte wijkt niet significant af: H_0: μ = 0,40

Alternatieve hypothese

Het "werkelijke" gehalte wijkt wel significant af: H_0: μ ≠ 0,40

$$t = \left|(\mu - \bar{x})\right| \cdot \frac{\sqrt{n}}{s} = \left|\left|0,40 - 0,38\right| \times \frac{\sqrt{12}}{0,02}\right| = 3,46$$

$v = n - 1 = 12 - 1 = 11$

tabel: $t_{\text{kritisch}} = 2,20$ (geen voorkeur dus tweezijdig testen)

2,20 < 3,46 dus de nulhypothese wordt afgewezen

De gevonden waarde wijkt significant af van de werkelijke waarde.

Metingen partner

Het "werkelijke" gehalte wijkt niet significant af: H_0: μ = 0,40

Alternatieve hypothese

Het "werkelijke" gehalte wijkt wel significant af: H_0: μ ≠ 0,40

$$t = \left|(\mu - \bar{x})\right| \cdot \frac{\sqrt{n}}{s} = \left|\left|0,40 - 0,44\right| \times \frac{\sqrt{8}}{0,01}\right| = 11,3$$

tabel: $t_{\text{kritisch}} = 2,36$ (geen voorkeur dus tweezijdig testen)

2,36 < 11,3 dus de nulhypothese wordt afgewezen

De gevonden waarde wijkt significant af van de werkelijke waarde.

d Beide meetmethoden voldoen niet

Opgave 9.5 **Vergelijken van twee meetseries**

Opgave 9.6 **T-test van gemiddelde uit twee steekproeven**

a Bereken S.

$$S = \sqrt{\frac{v_1 \cdot s_1^{\ 2} + v_2 \cdot s_2^{\ 2}}{v_1 + v_2}} = \sqrt{\frac{74 \times 420^2 + 74 \times 425^2}{74 + 74}} = 422{,}5$$

b

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{S \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{|3100 - 2750|}{422{,}5 \times \sqrt{\frac{1}{75} + \frac{1}{75}}} = 5{,}07$$

c tabel:
schatting $t_{\text{kritisch}} = 2{,}00$
$5{,}07 > 2{,}00$ dus de nulhypothese wordt verworpen

d Er is wel een significant verschil tussen de gemiddelden
Het gemiddelde gewicht van de behandelde groep is dus groter dan die van de controlegroep

Opgave 9.7 **F-test van standaarddeviaties uit twee steekproeven**

Hier wordt verstandig gekozen voor tweezijdig toetsen.

a $F = \frac{s_A^{\ 2}}{s_B^{\ 2}} = \frac{0{,}40^2}{0{,}25^2} = 2{,}56$

b tabel: $F_{\text{kritisch}} = 3{,}58$
$2{,}56 < 3{,}58$ dus de nulhypothese wordt aangenomen

c De meetseries verschillen niet significant in precisie. Je kunt dus niet zeggen dat serie B nauwkeuriger is. De verschillen zijn aan toeval te wijten

Opgave 9.8 **Afvalwateronderzoek**

Gemiddelde

Nulhypothese: Er is geen significant verschil tussen de gemiddelden: H_0: $|\mu_1 - \mu_2| = 0$

Alternatief: H_1: $|\mu_1 - \mu_2| \neq 0$ (geen voorkeur voor een van beide methoden), dus tweezijdig testen.

$$S = \sqrt{\frac{v_1 \cdot s_1^{\ 2} + v_2 \cdot s_2^{\ 2}}{v_1 + v_2}} = \sqrt{\frac{10 \times 1{,}27^2 + 10 \times 1{,}99^2}{10 + 10}} = 1{,}67$$

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{S \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{|4{,}55 - 6{,}37|}{1{,}67 \times \sqrt{\frac{1}{11} + \frac{1}{11}}} = 8{,}55$$

tabel: $t_{\text{kritisch}} = 2{,}23$
$8{,}55 > 2{,}23$ dus de nulhypothese wordt verworpen. Er is een significant (opvallend) verschil in de gevonden gemiddelden

Standaarddeviatie

Nulhypothese: De precisie van methode B is ***niet*** significant beter dan de precisie van methode A: H_0: $\sigma_A = \sigma_B$

Alternatief: de precisie van methode B is significant slechter dan de precisie van methode A H_1: $\sigma_A < \sigma_B$. Dus tweezijdig testen.

$$F = \frac{s_A{}^2}{s_B{}^2} = \frac{1,99^2}{1,27^2} = 2,46$$

tabel: $F_{kritisch} = 3,72$
2,46 < 3,72 dus de nulhypothese wordt aangenomen. De meetseries zijn wel vergelijkbaar wat betreft precisie.

Opgave 9.9 ***T*-test van gemiddelde uit twee steekproeven met gepaarde waarnemingen**

a op het geluidsignaal
b ja
c nul?
d

$$t = \frac{\bar{x}_v \cdot \sqrt{n}}{s_V} = \frac{22 \times \sqrt{10}}{34,4} = 2,02$$

e tabel: $t_{kritisch} = 2,26$
2,02 < 2,26 dus de nulhypothese wordt aangenomen.
f Er is geen significant verschil tussen de gemiddelde reactietijden.

Opgave 9.10 **Hemoglobinegehalte**

Nulhypothese
Er is geen significant verschil tussen de gemiddelde Hb-gehaltes per patiënt
H_0: $\bar{x}_v = 0$
Alternatieve hypothese
Er is wel een significant verschil tussen de gemiddelde Hb-gehaltes per patiënt
H_1: $x_v \neq 0$

Hb-gehalte (g/dL)			
patiënt	**A**	**B**	verschil
1	12,5	13,2	-0,9
2	13,6	14,1	-1,1
3	16,3	16,8	-0,8
4	15,8	15,2	0,6
5	14,6	15,3	0,7
6	11,3	13,8	-2,5
gemiddeld	14,0	14,9	-0,9
			0,9

$$t = \frac{\bar{x}_v \cdot \sqrt{n}}{s_V} = \frac{0,9 \times \sqrt{6}}{0,9} = 2,45$$ **(neem** $\bar{x}_v > 0$)

tabel: $t_{kritisch} = 2,57$

2,45 < 2,57 dus de nulhypothese wordt aangenomen. Er is geen significant verschil tussen beide meetmethoden

Opgave 9.11 **Opstellen van hypotheses**

CASUS 1

a Het gemiddelde gehalte van een steekproef uit de partij kindervoeding.

b **Nulhypothese**
Er is geen significant verschil tussen het gemiddelde gehalte en de maximale waarde van 0,02 kg
H_0: μ = 0,02 mg/kg
Alternatieve hypothese
Het gemiddelde gehalte is significant lager dan de maximale waarde van 0,02 kg
H_1: μ < 0,02 mg/kg

c Wel een voorkeur dus eenzijdig toetsen.

d De t-test voor vergelijking van een gemiddelde van een steekproef met een (on)gewenste waarde μ

CASUS 2

a Steekproeven met methode A en een met methode B worden vergeleken. De standaarddeviaties worden vergeleken.

b **Nulhypothese**
Er is geen significant verschil tussen de standaarddeviaties van methode A en B
H_0: $\sigma_A = \sigma_B$
Alternatief: de precisie van methode B is significant beter dan de precisie van methode A
H_1: $\sigma_B < \sigma_A$.

c Wel een voorkeur dus eenzijdig toetsen.

d De F-test voor vergelijking van de standaarddeviaties van twee steekproeven.

CASUS 3

a Aan begin en eind van de periode van alle patiënten de bloeddruk meten. Het gemiddelde van de verschillen voor en na wordt vergeleken.

b **Nulhypothese**
Er is geen significant verschil tussen het gemiddelde van de verschillen van de bloeddrukwaarden per patiënt.
H_0: $\bar{x}_v = 0$
Alternatieve hypothese
Het gemiddelde verschil van de bloeddrukwaarden per patiënt is significant lager na de behandeling.
H_1: $\bar{x}_v > 0$

c Wel een voorkeur dus eenzijdig toetsen.

d De gepaarde t-test voor vergelijking van de steekproeven

CASUS 4

a De standaarddeviaties van de metingen van de twee analisten worden vergeleken.

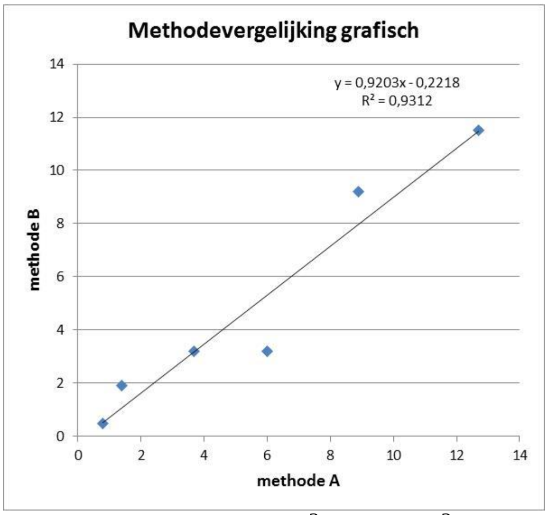
b **Nulhypothese**
Er is geen significant verschil tussen de standaarddeviaties van analist A en analist B
H_0: $\sigma_A = \sigma_B$
Alternatieve hypothese
Er is een significant verschil tussen de standaarddeviaties van analist A en analist B.
H_1: $\sigma_B \neq \sigma_A$.

c Geen voorkeur dus tweezijdig toetsen.

d De F-test voor vergelijking van de standaarddeviaties van twee steekproeven.

Opgave 9.12 Grafische vergelijking van meetmethoden

a



c R = 0,9312 dus $R^2 = 0{,}9312^2 = 0{,}867$
de grenswaarde is 0,811, dus er is aantoonbare ,maar zwakke correlatie

d helling = 1 en asafsnijding = 0

e op het oog lijken deze methoden niet goed vergelijkbaar, de afwijkingen t.o.v. de ideale waarden is redelijk groot

Opgave 9.13 **Grafische vergelijking van meetmethoden - Uitschieters**

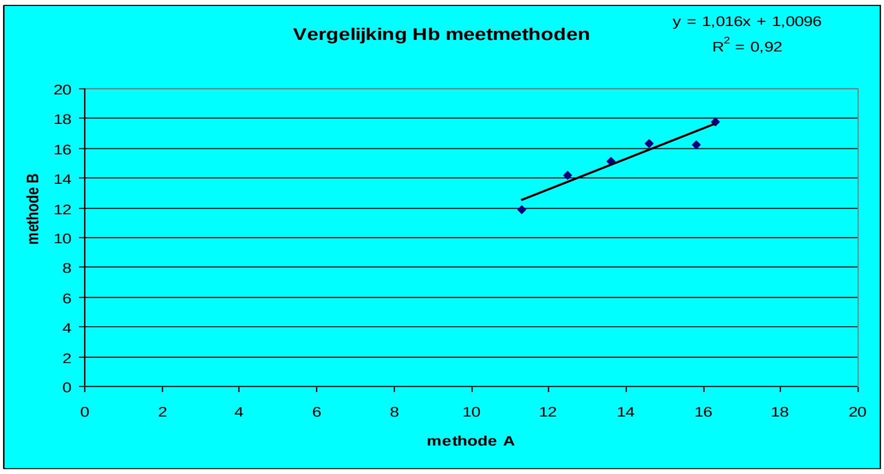
patiënt	methode A	methode B	verschil	verschil abs	test (4x)
1	**0,8**	**0,5**	0,3	0,3	-3,4
2	**1,4**	**1,9**	-0,5	0,5	-3,2
3	**3,7**	**3,2**	0,5	0,5	-3,2
4	**6**	**3,2**	2,8	2,8	-0,9
5	**8,9**	**9,2**	-0,3	0,3	-3,4
6	**12,7**	**11,5**	1,2	1,2	-2,5
		gemiddeld	0,67	0,93	

er zijn geen uitschieters
$y = 0{,}9203x - 0{,}2218$
$R^2 = 0{,}9312$

Opgave 9.14 **Grafische vergelijking van meetmethoden - Valkuilen**

a alle waarden met methode B zijn groter dan dezelfde van A

b



ze komen overeen allen de waarden bij B zijn gemiddeld 1,0 hoger dan die van A

c Een van de methodes vertoont een systematische afwijking. Dat kan zowel A als B zijn

d het zo niet vast te stellen welke methode afwijkt, je zou de kalibratielijnen per methode moeten bekijken

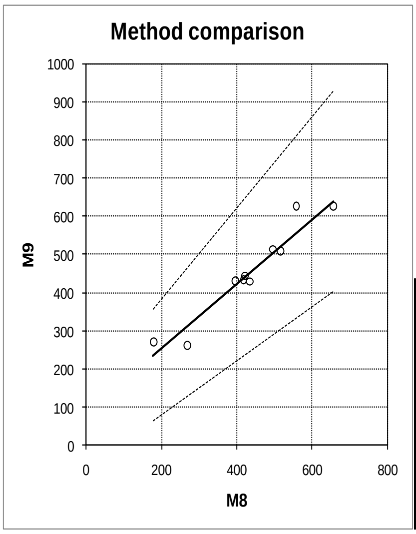
e in het hogere meetgebied wijkt een van de twee methoden af (niet te zeggen welke)

Opgave 9.15 **Vergelijking van meetmethoden volgens Passing en Bablok**

a het verschil zit alleen in de onzekerheid van helling en snijpunt met de y-as, de formules zijn gelijk

b de ideale waarden (1 en 0) liggen binnen de betrouwbaarheidsintervallen, dus de methoden zijn vergelijkbaar

c



Result of PB-Regression	
slope	**0,8429**
lower 95%-CL	0,7037
upper 95%-CL	1,1883
Intercept	**84,3429**
lower 95%-CL	-60,4026
upper 95%-CL	145,8148

de methodes zijn vergelijkbaar

Opgave 9.16 **Vergelijking van meetmethoden volgens Deming**

deming		normaal	
helling	0,97476	helling	1,033429
snijpunt	0,383629	snijpunt	-0,46153
correlatie	0,984639	correlatie	0,993386

wat opvalt is dat de Demingregressie een kleinere correlatie geeft en een duidelijk afwijkend snijpunt met de *y*-as

Opgave 9.17 **De analyse volgens Bland en Altman**

a onderste grens = –4,4 – 2×24,1 = – 52,6
bovenste grens = –4,4 + 2×24,1 = 43,8

b gemiddeld verschil = –4,4

c –4,4 L/min

d Bij de ene serie 4,4 optellen of bij de andere 4,4 eraf halen

e nee

f $n = 10 \rightarrow \nu = 9 \rightarrow t = 2,26$
grens betrouwbaarheidinterval = $t \cdot \frac{s}{\sqrt{n}} = 2,26 \times \frac{24,1}{\sqrt{10}} = 17,2$
ondergrens –4,4 – 17,2 = – 21,6
bovengrens –4,4 + 17,2 = 12,8
dus – 21,6 < afwijking < 12,8

g $SE = \sqrt{\frac{3\sigma^2}{n}} = \sqrt{\frac{3 \times 24,1^2}{10}} = 13,2$

-95,8 < onderste grens < 39,4
30,6 < bovenste grens < 57,0

h De afwijkingen lijken toch behoorlijk groot