Chemometrie, ISBN: 9789491764332
Vierde herziene druk
www.syntaxmedia.nl

Uitwerkingen

In de uitgewerkte voorbeelden worden vanwege de leesbaarheid afgeronde tussenresultaten gepresenteerd. De eindresultaten zijn echter altijd berekend zonder tussentijds afronden.

Hoofdstuk 6

In hoofdstuk 4 staat in de tekst op bladzijde 68 dat rechts in vergelijking (4.28) voor *F* de '*tweezijdige F*-waarde uit tabel 5 van bijlage 1' moet worden ingevuld. Dit moet zijn de '*eenzijdige F*-waarde uit tabel 4 van bijlage 1'.

Deze correctie heeft de volgende effecten op de tekst in hoofdstuk 6.

<u>Tekst op bladzijde 154:</u>

De volgende tekst,

> Voor het rechterdeel van de vergelijking geldt, met 2 vrijheidsgraden in de teller en (n – 2 =) 4 vrijheidsgraden in de noemer, de *tweezijdige* $F_{(0,05;2;4)}$ = 10,65 (tabel 5 in bijlage 1):
>
> $$2 \cdot 10{,}65 \cdot \frac{0{,}0144^2}{6} = 7{,}35 \cdot 10^{-4}$$

Moet worden vervangen door,

> Voor het rechterdeel van de vergelijking geldt, met 2 vrijheidsgraden in de teller en (n – 2 =) 4 vrijheidsgraden in de noemer, de *eenzijdige* $F_{(0,05;2;4)}$ = 6,94 (tabel 4 in bijlage 1):
>
> $$2 \cdot 6{,}94 \cdot \frac{0{,}0144^2}{6} = 4{,}79 \cdot 10^{-4}$$

N.B. De tekst die hier op aansluit is correct. Ook Afbeelding 6.2b is correct.

<u>Tekst op bladzijde 160:</u>

De volgende tekst,

> Voor het rechterdeel van de vergelijking geldt, met 2 vrijheidsgraden in de teller en (n – 2 =) 10 vrijheidsgraden in de noemer, met de *tweezijdige* $F_{(0,05;2;10)}$ = 5,46 (tabel 5 in bijlage 1):
>
> $$2 \cdot 5{,}46 \cdot \frac{1{,}4191^2}{12} = 1{,}8314$$

Moet worden vervangen door,

> Voor het rechterdeel van de vergelijking geldt, met 2 vrijheidsgraden in de teller en (n – 2 =) 10 vrijheidsgraden in de noemer, met de *eenzijdige* $F_{(0,05;2;10)}$ = 4,10 (tabel 4 in bijlage 1):
>
> $$2 \cdot 4{,}10 \cdot \frac{1{,}4191^2}{12} = 1{,}3771$$

N.B. De tekst die hier op aansluit is correct. Ook Afbeelding 6.4b is correct.

Antwoord 6.1

n	10
c_ref	0,012
nr.	x
1	0,0137
2	0,0128
3	0,0116
4	0,0146
5	0,0136
6	0,0131
7	0,0123
8	0,0118
9	0,0111
10	0,0127

0,01273	gem
0,00107	sdev
2,262	t_krit
2,156	t_ber
0,0073	bias

Herhaalbaarheid $r = s_r = 0{,}00107$.
Bias, berekend met (6.18) is $d = \bar{x} - c_{\text{ref}} = 0{,}01273 - 0{,}012 = 0{,}0073$.

De t-test op significantie van de bias wordt uitgevoerd met (6.20):

$$|t| = \frac{|\bar{x} - c_{\text{ref}}|}{s/\sqrt{n}} = \frac{|0{,}01273 - 0{,}012|}{0{,}00107/\sqrt{10}} = 2{,}156$$

De kritische t-waarde is $t_{(0,05;9)} = 2{,}262$.
Omdat $|t|_{\text{ber}} < t_{(0,05;9)}$ is de bias niet significant.

Antwoord 6.2

	A	B	C	D	E
1	30,5	26,7	29,1	31,3	30,1
2	29,2	28,2	29,3	30,8	29,2
3	30,1	27,6	29,3	31,1	32,2
4	29,1	27,8	28,8	31,8	31,5

Aantal groepen (laboratoriummedewerkers A-E) $n = 5$; aantal herhalingen $k = 4$.

Unifactoriële variantieanalyse

SAMENVATTING

Groepen	Aantal	Som	Gemiddelde	Variantie
1	4	118,9	29,725	0,4692
2	4	110,3	27,575	0,4025
3	4	116,5	29,125	0,0558
4	4	125	31,25	0,1767
5	4	123	30,75	1,8300

Variantieanalyse

Bron van variatie	Kwadratensom	Vrijheidsgraden	Gemiddelde kwadraten	F	*p*-waarde	Kritische gebied van *F*-toets
Tussen groepen	33,403	4	8,3508	14,23	5,44E-05	3,06
Binnen groepen	8,8025	15	0,5868			
Totaal	42,2055	19				

De P-waarde voor de berekende F-waarde is kleiner dan 0,05. Dat betekent dat er een significant verschil is tussen de tussengroepvariantie en binnengroepvariantie. Er is dus ook een significant verschil tussen de resultaten van de laboratoriummedewerkers.

De variantie onder herhaalbaarheidscondities s_r^2 is volgens (6.12) gelijk aan de binnengroepvariantie: $s_r^2 = s_{\mathrm{BG}}^2 = 0{,}5868$. De variantie onder reproduceerbaarheidscondities s_R^2 kan met (6.17) worden berekend uit de ANOVA-resultaten:

$$s_R^2 = s_{\mathrm{BG}}^2 + (s_{\mathrm{TG}}^2 - s_{\mathrm{BG}}^2)/k = 0{,}5868 + (8{,}3508 - 0{,}5868)/4 = 2{,}5278$$

De herhaalbaarheid is $r = s_r = \sqrt{s_r^2} = \sqrt{0{,}5868} = 0{,}77$.
De reproduceerbaarheid is $R = s_R = \sqrt{s_R^2} = \sqrt{2{,}5278} = 1{,}59$.

Antwoord 6.3

	A	B	C	D
1	0,480	0,382	0,374	0,423
2	0,444	0,407	0,256	0,491
3	0,470	0,448	0,358	0,572
4	0,677	0,417	0,282	0,498
5	0,475	0,416	0,348	0,658

Aantal groepen (laboratoria A-D) n = 4; aantal herhalingen k = 5.

Unifactoriële variantieanalyse
SAMENVATTING

Groepen	Aantal	Som	Gemiddelde	Variantie
A	5	2,546	0,5092	0,008992
B	5	2,070	0,4140	0,000561
C	5	1,618	0,3236	0,002655
D	5	2,642	0,5284	0,008032

Variantieanalyse

Bron van variatie	Kwadratensom	Vrijheidsgraden	Gemiddelde kwadraten	F	*p*-waarde	Kritische gebied van *F*-toets
Tussen groepen	0,1339	3	4,462E-02	8,82	0,001108	3,24
Binnen groepen	0,0810	16	5,060E-03			
Totaal	0,2148	19				

De *P*-waarde voor de berekende *F*-waarde is kleiner dan 0,05. Dat betekent dat er een significant verschil is tussen de tussengroepvariantie en binnengroepvariantie. Er is dus ook een significant verschil tussen de resultaten van de laboratoria.

De variantie onder herhaalbaarheidscondities s_r^2 is volgens (6.12) gelijk aan de binnengroepvariantie: $s_r^2 = s_{\mathrm{BG}}^2 = 5{,}060 \cdot 10^{-3}$. De variantie onder reproduceerbaarheidscondities s_R^2 kan met (6.17) worden berekend uit de ANOVA-resultaten:

$$s_R^2 = s_{\mathrm{BG}}^2 + (s_{\mathrm{TG}}^2 - s_{\mathrm{BG}}^2)/k = 5{,}060 \cdot 10^{-3} + (4{,}462 \cdot 10^{-2} - 5{,}060 \cdot 10^{-3})/5 = 1{,}30 \cdot 10^{-2}$$

De herhaalbaarheid is $r = s_r = \sqrt{s_r^2} = \sqrt{5{,}060 \cdot 10^{-3}} = 7{,}11 \cdot 10^{-2}$.
De reproduceerbaarheid is $R = s_R = \sqrt{s_R^2} = \sqrt{1{,}30 \cdot 10^{-2}} = 1{,}14 \cdot 10^{-1}$.

Antwoord 6.4

nr	$x_{c,i}$	$x_{c+\Delta c,i}$	T_i
1	25,4	50,0	0,984
2	24,6	50,6	1,040
3	25,0	49,8	0,992
4	25,4	50,4	1,000
5	25,2	49,2	0,960
6	24,9	50,5	1,024
7	25,5	49,4	0,956
8	25,2	50,6	1,016
9	25,9	49,6	0,948
10	25,8	49,8	0,960
gem	25,29	49,99	0,9880
sdev	0,398	0,513	0,0319
abs(t)			1,190
t_krit			2,262
halve BI			0,023
BI_BG			1,011
BI_OG			0,965

<u>Berekening van de terugvinding met test op significantie</u>

De terugvinding T is gelijk aan het gemiddelde van T_i: $T = 0{,}988$.
Er kan worden getest of de terugvinding T significant van 1 verschilt met behulp van een tweezijdige t-toets op het 95%-betrouwbaarheidsniveau:

$$|t| = \frac{|T-1|}{s/\sqrt{n}} = \frac{|0{,}988-1|}{0{,}0319/\sqrt{10}} = 1{,}190$$

De kritische tweezijdige t-waarde voor het gemiddelde van 10 metingen met 9 vrijheidsgraden is $t_{(0,05;9)} = 2{,}262$ (tabel 3 in bijlage 1). Omdat $|t| < 2{,}262$ is de terugvinding T niet significant verschillend van 1 en is de methode juist.

Een alternatief is om het 95%-betrouwbaarheidsinterval van T te berekenen, met $t_{(0,05;9)} = 2{,}262$:

$$\mathrm{BI}(T) = T \pm \frac{t_{(0,05;n-1)}s}{\sqrt{n}} = 0{,}988 \pm \frac{2{,}262 \cdot 0{,}0319}{\sqrt{10}} = 0{,}988 \pm 0{,}023$$

Dus $0{,}965 < T < 1{,}011$. Omdat 1 binnen het 95%-betrouwbaarheidsinterval ligt, is T niet significant verschillend van 1.

Antwoord 6.5

Correctie op antwoorden:
b. rechterdeel $3{,}78 \cdot 10^{-3}$ moet zijn $2{,}68 \cdot 10^{-3}$.

nr.	x_i	y_i		
	x1	x2		x_1^2
1	92,0	89,3		8464
2	93,0	90,3		8649
3	94,0	91,3		8836
4	95,0	92,2		9025
5	96,0	93,3		9216
6	97,0	94,2		9409
7	98,0	95,2		9604
8	99,0	96,1		9801
gem	95,50	92,74	som	73004
sdev	2,4495	2,3886		

Berekening van de schijnbare terugvinding met test op significantie

Er kan regressie worden uitgevoerd met de concentraties van de standaarden x_1 op de x-as en de teruggevonden concentraties x_2 op de y-as. De resultaten van lineaire regressie met Excel zijn:

Gegevens voor de regressie	
Meervoudige correlatiecoëfficiënt R	0,9998
R-kwadraat	0,9997
Aangepaste kleinste kwadraat	0,9996
Standaardfout	0,0456
Waarnemingen	8

Variantieanalyse

	Vrijheidsgraden	Kwadratensom	Gemiddelde kwadraten	F	Significantie F
Regressie	1	39,92625	39,92625	19165	9,5818E-12
Storing	6	0,0125	2,083E-03		
Totaal	7	39,93875			

	Coëfficiënten	Standaardfout	T-statistische gegevens	P-waarde	Laagste 95%	Hoogste 95%
Snijpunt	-0,3750	0,6728	-0,5574	0,5974	-2,0213	1,2713
Variabele X 1	0,9750	0,0070	138,4363	0,0000	0,9578	0,9922

Hieruit kan worden afgelezen dat:
b_0 = -0,3750; b_1 = 0,9750; s_r = 0,0456; $s(b_0)$ = 0,6728; $s(b_1)$ = 0,0070.
BI(b_0): $-2{,}02 \leq b_0 \leq 1{,}27$ en BI(b_1): $0{,}96 \leq b_1 \leq 0{,}99$.

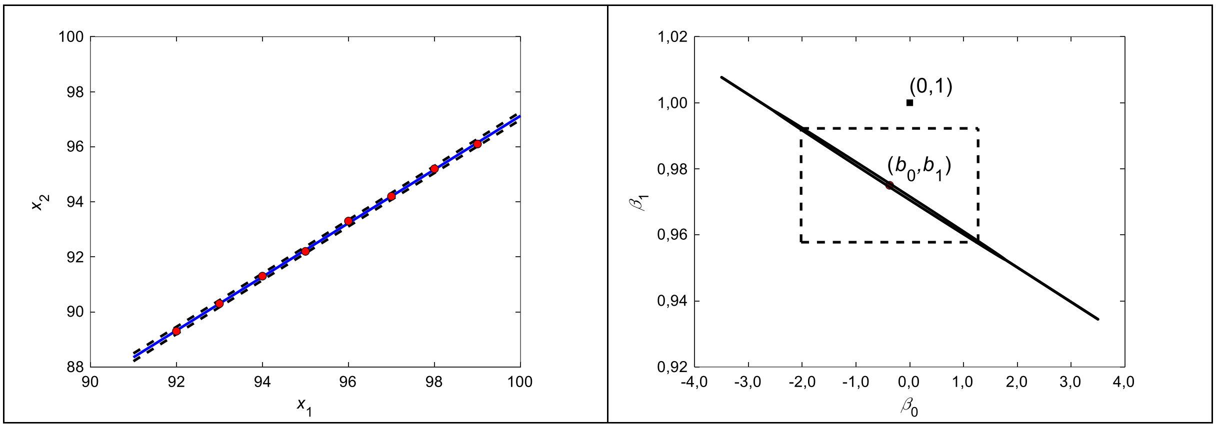
De uiterste grenzen van de individuele 95%-betrouwbaarheidsintervallen van de modelparameters staan in de Excel regressie-uitvoer. Hieruit kan de helft van het betrouwbaarheidsinterval worden berekend door het verschil te nemen van de hoogste waarde van het 95% betrouwbaarheidsinterval en de waarde van de parameter.
De resultaten zijn: $b_0 = -0{,}357 \pm 1{,}646$; $b_1 = 0{,}975 \pm 0{,}017$.

De onzekerheid in b_0 is veel groter dan die in b_1.
De schijnbare terugvinding is gelijk aan de helling: $T' = b_1 = 0{,}975$.
Omdat voor BI($T' = b_1$) geldt $0{,}96 \leq T' = b_1 \leq 0{,}99$, ligt 1 niet in het BI($T'$) en verschilt T' daarom significant van 1.

De bijbehorende regressielijn met 95%-betrouwbaarheidsintervallen is links weergegeven in de volgende afbeelding. De individuele betrouwbaarheidsintervallen en het gecombineerde betrouwbaarheidsinterval van b_0 en b_1 zijn rechts weergegeven. Het punt waarvoor de coördinaten gelijk zijn aan de geschatte b_0 en b_1 ligt in het midden van de rechthoek en de ellips. Merk op dat het gecombineerde betrouwbaarheidsinterval erg smal is.

Test op bias van de individuele 95% betrouwbaarheidsintervallen van b_0 en b_1

Omdat nul binnen het individuele betrouwbaarheidsinterval van b_0 ligt, bestaat er geen significante absolute bias. Er bestaat wel een significante relatieve bias omdat één buiten het individuele betrouwbaarheidsinterval ligt van b_1. De rechter afbeelding laat zien dat de ideale combinatie ($\beta_0 = 0$; $\beta_1 = 1$) buiten de rechthoek valt die gevormd wordt door de individuele betrouwbaarheidsintervallen.



Test op bias van de gecombineerde 95% betrouwbaarheidsintervallen van b_0 en b_1

Met behulp van vergelijking (4.28) kan worden getest of de ideale combinatie ($\beta_0 = 0$; $\beta_1 = 1$) binnen of buiten het gecombineerde betrouwbaarheidsinterval ligt van b_0 en b_1.

$$(\beta_0 - b_0)^2 + 2\bar{x}(\beta_0 - b_0)(\beta_1 - b_1) + \left(\frac{\sum x_i^2}{n}\right)(\beta_1 - b_1)^2 \leq \frac{2F_{(0,05;2;n-2)}s_r^2}{n}$$

$\bar{x} = 95{,}50$; $\sum x_i^2 = 73004$.

Voor het linkerdeel van de vergelijking geldt:

$$\big(0 - (-0{,}3750)\big)^2 + 2 \cdot 95{,}50 \cdot \big(0 - (-0{,}3750)\big)(1 - 0{,}9750) + \left(\frac{73004}{8}\right)(1 - 0{,}9750)^2 = 7{,}63$$

Voor het rechterdeel van de vergelijking geldt, met 2 vrijheidsgraden in de teller en (n – 2 =) 6 vrijheidsgraden in de noemer, de *eenzijdige* $F_{(0,05;2;6)} = 5{,}14$ (tabel 4 in bijlage 1):

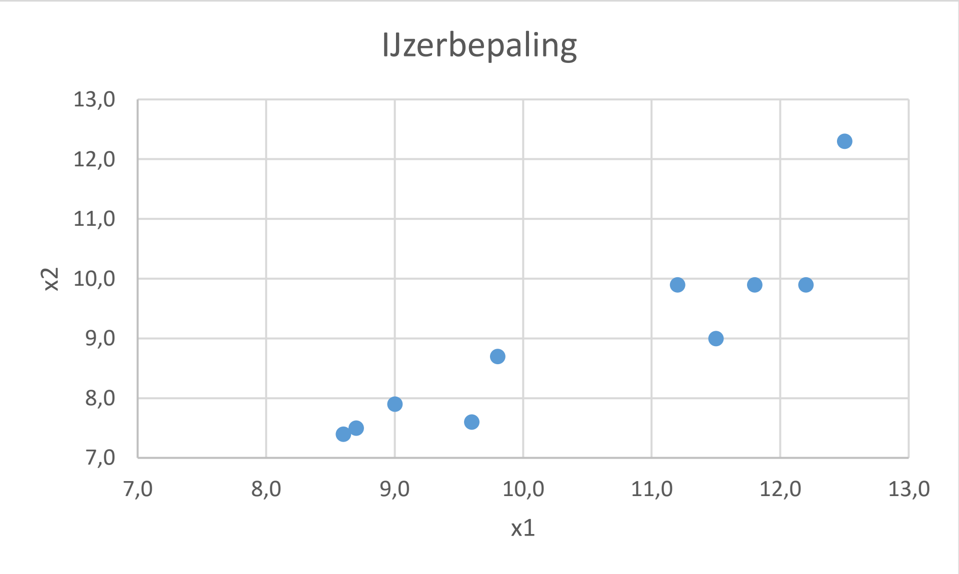
$$\frac{2 \cdot 5{,}14 \cdot 0{,}0456^2}{8} = 2{,}68 \cdot 10^{-3}$$

Omdat het linkerdeel van de vergelijking groter is dan het rechterdeel ligt de ideale combinatie ($\beta_0 = 0$; $\beta_1 = 1$) buiten het zeer smalle ellipsvormige gecombineerde betrouwbaarheidsinterval van b_0 en b_1, zie voorgaande afbeelding rechts.

Op basis van zowel de individuele betrouwbaarheidsintervallen als het gecombineerde betrouwbaarheidsinterval van b_0 en b_1 verschillen daarom de teruggevonden concentraties significant van de bekende concentraties.

Antwoord 6.6

nr.	x1	x2	gem (x1+x2)/2	verschil d=x1-x2	
1	8,6	7,4	8,00	1,2	
2	8,7	7,5	8,10	1,2	
3	9,0	7,9	8,45	1,1	
4	9,6	7,6	8,60	2,0	
5	9,8	8,7	9,25	1,1	
6	11,2	9,9	10,55	1,3	
7	11,5	9,0	10,25	2,5	
8	11,8	9,9	10,85	1,9	
9	12,2	9,9	11,05	2,3	
10	12,5	12,3	12,40	0,2	
gem	10,490	9,010		1,480	gem
sdev	1,508	1,542		0,689	sdev
				2,262	t_krit
				6,790	t_ber
				0,49	halve BI bias
				0,99	ondergrens bias
				1,97	bovengrens bias



<u>Gepaarde t-test</u>

Correctie vergelijking (6.24) op bladzijde 155:
De deelstreep in de noemer ontbreekt. De correcte formule is:

$$|t| = \frac{|\bar{d}|}{s_d/\sqrt{n}}$$

Op bladzijde 168 ontbreekt de deelstreep in de noemer ook. De correcte formule is:

$$|t| = \frac{|\bar{d}|}{s_d/\sqrt{n}} = \frac{|-1{,}71|}{1{,}36/\sqrt{12}} = 4{,}36$$

De beschreven resultaten zijn wel correct.

De gepaarde *t*-test kan ook worden uitgevoerd met Excel in de module ‘Gegevensanalyse’. De uitvoer daarvan is weergegeven in de volgende tabel. Uit deze tabel blijkt dat de *p*-waarde voor de tweezijdige gepaarde *t*-test significant is met $p = 8 \cdot 10^{-5}$.

De berekende $|t| = 6{,}79$. De kritische t is $t_{krit} = t_{(0,05;9)} = 2{,}26$. Omdat $|t| > t_{krit}$ is er een significant verschil tussen de methoden.

T-toets: twee gepaarde steekproeven voor gemiddelden

	x_1	x_2
Gemiddelde	10,490	9,010
Variantie	2,274	2,377
Waarnemingen	10	10
Pearson-correlatie	0,8981	
Schatting van verschil tussen gemiddelden	0	
Vrijheidsgraden	9	
T-statistische gegevens	6,790	
p(T ≤ t) eenzijdig	4E-05	
Kritiek gebied van T-toets: eenzijdig	1,833	
p(T ≤ t) tweezijdig	8E-05	
Kritiek gebied van T-toets: tweezijdig	2,262	

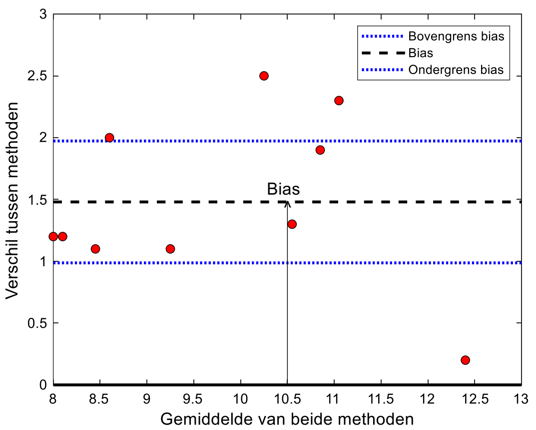
Bland-Altman methode

Het gemiddelde van de verschillen is $\bar{d} = 1{,}480$ ppm. Dit is gelijk aan de bias.
De standaarddeviatie van de verschillen is $s_d = 0{,}689$ ppm.
Het 95%-betrouwbaarheidsinterval van de bias BI$(\bar{d})$ kan met (6.25) voor het verschil van 10 metingen met 9 vrijheidsgraden worden berekend met behulp van de tweezijdige $t_{(0,05;9)} = 2{,}262$ (tabel 3 in bijlage 1):

$$\mathrm{BI}(\bar{d}) = \bar{d} \pm \frac{t_{(0,05;n-1)} s_d}{\sqrt{n}} = 1{,}480 \pm \frac{2{,}262 \cdot 0{,}689}{\sqrt{10}} = 1{,}48 \pm 0{,}49$$

BI$(\bar{d})$: $0{,}99 \leq \bar{d} \leq 1{,}97$.

Nul ligt niet binnen het 95%-betrouwbaarheidsinterval van het gemiddelde verschil $\bar{d}$. Zie ook volgende grafiek. De gemodificeerde methode verschilt dus significant van de oorspronkelijke methode.



Orthogonale regressie

i	x_i	y_i	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
1	8,6	7,4	-1,890	3,5721	-1,6100	2,5921	3,0429
2	8,7	7,5	-1,790	3,2041	-1,5100	2,2801	2,7029
3	9,0	7,9	-1,490	2,2201	-1,1100	1,2321	1,6539
4	9,6	7,6	-0,890	0,7921	-1,4100	1,9881	1,2549
5	9,8	8,7	-0,690	0,4761	-0,3100	0,0961	0,2139
6	11,2	9,9	0,710	0,5041	0,8900	0,7921	0,6319
7	11,5	9,0	1,010	1,0201	-0,0100	0,0001	-0,0101
8	11,8	9,9	1,310	1,7161	0,8900	0,7921	1,1659
9	12,2	9,9	1,710	2,9241	0,8900	0,7921	1,5219
10	12,5	12,3	2,010	4,0401	3,2900	10,8241	6,6129
som	104,9	90,1		20,4690		21,3890	18,7910
gem	10,49	9,01		SSxx		SSyy	SSxy

De helling b_1 en de as-afsnede b_0 van de orthogonale regressielijn kunnen worden berekend met respectievelijk (5.11) en (5.12):

$$b_1 = \frac{SS_{\mathrm{yy}} - SS_{\mathrm{xx}} + \sqrt{\left(SS_{\mathrm{yy}} - SS_{\mathrm{xx}}\right)^2 + 4SS_{\mathrm{xy}}^2}}{2SS_{\mathrm{xy}}}$$

$$b_1 = \frac{21{,}3890 - 20{,}4690 + \sqrt{(21{,}3890 - 20{,}4690)^2 + 4 \cdot 18{,}7910^2}}{2 \cdot 18{,}7910} = 1{,}025$$

$$b_0 = \bar{y} - b_1\bar{x} = 9{,}01 - 1{,}025 \cdot 10{,}49 = -1{,}740$$

Antwoord 6.7

Correctie op antwoorden:
a4 rechterdeel 2,561 moet zijn 1,926.

Gepaarde *t*-test

nr.	x1	x2	gem (x1+x2)/2	verschil d=x1-x2		x_1^2
1	5,2	6,0	5,60	-0,8		27,04
2	5,4	5,5	5,45	-0,1		29,16
3	4,9	8,4	6,65	-3,5		24,01
4	6,8	7,6	7,20	-0,8		46,24
5	19,4	21,6	20,50	-2,2		376,36
6	20,1	25,1	22,60	-5,0		404,01
7	26,7	26,1	26,40	0,6		712,89
8	26,6	27,8	27,20	-1,2		707,56
9	53,5	56,0	54,75	-2,5		2862,25
10	61,2	61,4	61,30	-0,2		3745,44
11	61,4	64,2	62,80	-2,8		3769,96
12	66,2	68,1	67,15	-1,9		4382,44
gem	29,78	31,48	gem	-1,700	som	17087,36
sdev	24,20	24,34	sdev	1,602		
			t_krit	2,201		
			t_ber	3,675		
			halve BI bias	1,018		
			ondergrens bias	-2,718		
			bovengrens bias	-0,682		

De gepaarde t-test kan ook worden uitgevoerd met Excel in de module ‘Gegevensanalyse’. De uitvoer daarvan is weergegeven in de volgende tabel. Uit deze tabel blijkt dat de p-waarde voor de tweezijdige gepaarde t-test significant is met p = 0,0037.

De berekende $|t|$ = 3,675. De kritische t is $t_{\text{krit}} = t_{(0,05;11)}$ = 2,201. Omdat $|t| > t_{\text{krit}}$ is er een significant verschil tussen de methoden.

T-toets: twee gepaarde steekproeven voor gemiddelden

	x_1	x_2
Gemiddelde	29,7833	31,4833
Variantie	585,7088	592,3633
Waarnemingen	12	12
Pearson-correlatie	0,9978	
Schatting van verschil tussen gemiddelden	0	
Vrijheidsgraden	11	
T-statistische gegevens	-3,6754	
p(T ≤ t) eenzijdig	0,0018	
Kritiek gebied van T-toets: eenzijdig	1,7959	
p(T ≤ t) tweezijdig	0,0037	
Kritiek gebied van T-toets: tweezijdig	2,2010	

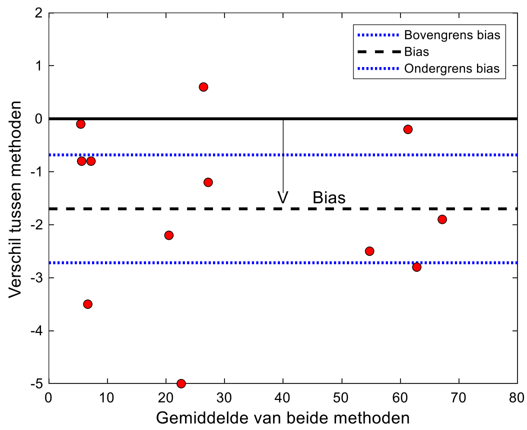
<u>Bland-Altman methode</u>

Het gemiddelde van de verschillen is $\bar{d}$ = -1,700. Dit is gelijk aan de bias.
De standaarddeviatie van de verschillen is s_d = 1,602.
Het 95%-betrouwbaarheidsinterval van de bias BI($\bar{d}$) kan met (6.25) voor het verschil van 12 metingen met 11 vrijheidsgraden worden berekend met behulp van de tweezijdige $t_{(0,05;11)}$ = 2,201 (tabel 3 in bijlage 1):

$$\mathrm{BI}(\bar{d}) = \bar{d} \pm \frac{t_{(0,05;n-1)} s_d}{\sqrt{n}} = -1{,}70 \pm \frac{2{,}201 \cdot 1{,}602}{\sqrt{12}} = -1{,}70 \pm 1{,}02$$

BI($\bar{d}$): $-2{,}72 \le \bar{d} \le -0{,}68$.

Nul ligt niet binnen het 95%-betrouwbaarheidsinterval van het gemiddelde verschil $\bar{d}$. Zie ook de volgende grafiek. De gemodificeerde methode (x_2) verschilt dus significant van de oorspronkelijke methode (x_1).



Test op basis van de individuele 95% betrouwbaarheidsintervallen van b_0 en b_1

Gegevens voor de regressie	
Meervoudige correlatiecoëfficiënt R	0,9978
R-kwadraat	0,9957
Aangepaste kleinste kwadraat	0,9952
Standaardfout	1,6781
Waarnemingen	12

Variantieanalyse

	Vrijheidsgraden	Kwadratensom	Gemiddelde kwadraten	F	Significantie F
Regressie	1	6487,8351	6487,8351	2303,791	3,72E-13
Storing	10	28,1616	2,8162		
Totaal	11	6515,9967			

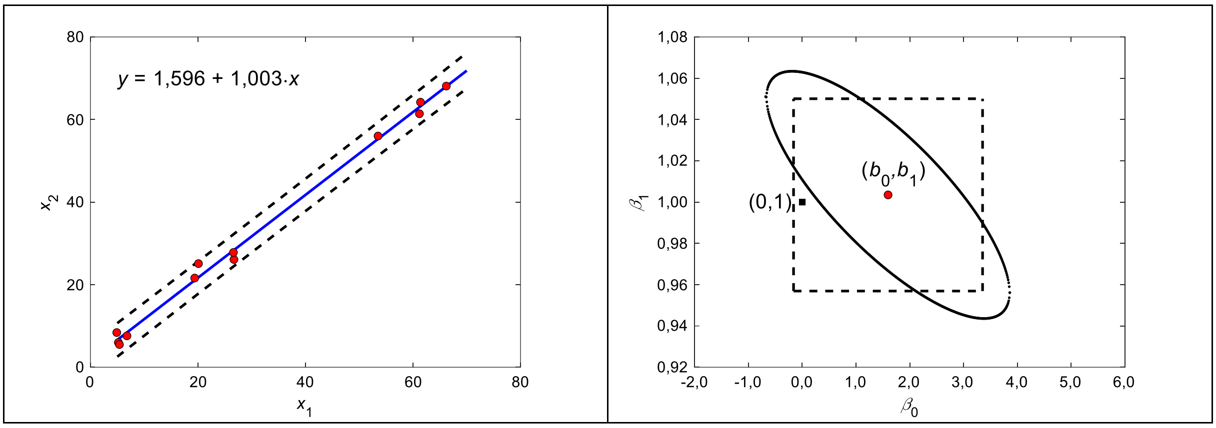
	Coëfficiënten	Standaardfout	T-statistische gegevens	P-waarde	Laagste 95%	Hoogste 95%
Snijpunt	1,5961	0,7889	2,0231	0,0706	-0,1618	3,3539
Variabele X 1	1,0035	0,0209	47,9978	0,0000	0,9569	1,0501

Hieruit kan worden afgelezen dat:
b_0 = 1,5961; b_1 = 1,0035; s_r = 1,6781; $s(b_0)$ = 0,7889; $s(b_1)$ = 0,0209.
BI(b_0): -0,162 ≤ b_0 ≤ 3,354 en BI(b_1): 0,957 ≤ b_1 ≤ 1,050.

De uiterste grenzen van de individuele 95%-betrouwbaarheidsintervallen van de modelparameters staan in de Excel regressie-uitvoer. Hieruit kan de helft van het betrouwbaarheidsinterval worden berekend door het verschil te nemen van de hoogste waarde van het 95% betrouwbaarheidsinterval en de waarde van de parameter.
De resultaten zijn: b_0 = 1,596 ± 1,758; b_1 = 1,003 ± 0,047.

De bijbehorende regressielijn met 95%-betrouwbaarheidsintervallen is links weergegeven in de volgende afbeelding. De individuele betrouwbaarheidsintervallen en het gecombineerde betrouwbaarheidsinterval van b_0 en b_1 zijn rechts weergegeven. Het punt waarvoor de coördinaten gelijk zijn aan de geschatte b_0 en b_1 ligt in het midden van de rechthoek en de ellips.

Omdat nul binnen het individuele betrouwbaarheidsinterval van b_0 ligt, bestaat er geen significante absolute bias. Omdat één binnen het individuele betrouwbaarheidsinterval van b_1 ligt, bestaat er ook geen significante relatieve bias. De rechter afbeelding laat zien dat de ideale combinatie (β_0 = 0; β_1 = 1) binnen de rechthoek valt die gevormd wordt door de individuele betrouwbaarheidsintervallen. Op basis van de individuele 95% betrouwbaarheidsintervallen van b_0 en b_1 bestaat er geen significante bias tussen beide methoden.



<u>Test op basis van de gecombineerde 95% betrouwbaarheidsintervallen van b_0 en b_1</u>

Met behulp van vergelijking (4.28) kan worden getest of de ideale combinatie (β_0 = 0; β_1 = 1) binnen of buiten het gecombineerde betrouwbaarheidsinterval ligt van b_0 en b_1.

$$(\beta_0 - b_0)^2 + 2\bar{x}(\beta_0 - b_0)(\beta_1 - b_1) + \left(\frac{\sum x_i^2}{n}\right)(\beta_1 - b_1)^2 \leq \frac{2F_{(0,05;2;n-2)}s_r^2}{n}$$

$\bar{x} = 29{,}78$; $\sum x_i^2 = 17087{,}36$.

Voor het linkerdeel van de vergelijking geldt:

$$(0 - 1{,}5961)^2 + 2 \cdot 29{,}78 \cdot (0 - 1{,}5961)(1 - 1{,}0035) + \left(\frac{17087{,}36}{12}\right)(1 - 1{,}0035)^2 = 2{,}897$$

Voor het rechterdeel van de vergelijking geldt, met 2 vrijheidsgraden in de teller en (n – 2 =) 10 vrijheidsgraden in de noemer, de *eenzijdige* $F_{(0,05;2;10)}$ = 4,10 (tabel 4 in bijlage 1):

$$\frac{2 \cdot 4{,}10 \cdot 1{,}6781^2}{12} = 1{,}926$$

Omdat het linkerdeel van de vergelijking groter is dan het rechterdeel ligt de ideale combinatie (β_0 = 0; β_1 = 1) buiten het ellipsvormige gecombineerde betrouwbaarheidsinterval van b_0 en b_1, zie voorgaande afbeelding rechts. Op basis van het gecombineerde betrouwbaarheidsinterval van b_0 en b_1 bestaat er een significante bias tussen beide methoden.

<u>Orthogonale regressie</u>

i	x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	5,2	6,0	-24,583	604,3403	-25,4833	649,4003	626,4653
2	5,4	5,5	-24,383	594,5469	-25,9833	675,1336	633,5603
3	4,9	8,4	-24,883	619,1803	-23,0833	532,8403	574,3903
4	6,8	7,6	-22,983	528,2336	-23,8833	570,4136	548,9186
5	19,4	21,6	-10,383	107,8136	-9,8833	97,6803	102,6219
6	20,1	25,1	-9,683	93,7669	-6,3833	40,7469	61,8119
7	26,7	26,1	-3,083	9,5069	-5,3833	28,9803	16,5986
8	26,6	27,8	-3,183	10,1336	-3,6833	13,5669	11,7253
9	53,5	56,0	23,717	562,4803	24,5167	601,0669	581,4536
10	61,2	61,4	31,417	987,0069	29,9167	895,0069	939,8819
11	61,4	64,2	31,617	999,6136	32,7167	1070,3803	1034,3919
12	66,2	68,1	36,417	1326,1736	36,6167	1340,7803	1333,4569
som	357,4	377,8		6442,7967		6515,9967	6465,2767
gem	29,78	31,48		SSxx		SSyy	SSxy

De helling b_1 en de as-afsnede b_0 van de orthogonale regressielijn kunnen worden berekend met respectievelijk (5.11) en (5.12):

$$b_1 = \frac{SS_{\mathrm{yy}} - SS_{\mathrm{xx}} + \sqrt{\left(SS_{\mathrm{yy}} - SS_{\mathrm{xx}}\right)^2 + 4SS_{\mathrm{xy}}^2}}{2SS_{\mathrm{xy}}}$$

$$b_1 = \frac{6515{,}9967 - 6442{,}7967 + \sqrt{(6515{,}9967 - 6442{,}7967)^2 + 4 \cdot 6465{,}2767^2}}{2 \cdot 6465{,}2767} = 1{,}006$$

$$b_0 = \bar{y} - b_1\bar{x} = 31{,}48 - 1{,}006 \cdot 29{,}78 = 1{,}531$$

Antwoord 6.8

Correctie op antwoorden:
$x_{\mathrm{AG}} = 7{,}72 \cdot 10^{-4}$ moet zijn $x_{\mathrm{AG}} = 2{,}12 \cdot 10^{-3}$.
$x_{\mathrm{BG}} = 1{,}22 \cdot 10^{-3}$ moet zijn $x_{\mathrm{BG}} = 2{,}71 \cdot 10^{-3}$.

nr.	y
1	0,048
2	0,031
3	0,066
4	0,002
5	0,016
6	0,029
7	0,010
8	0,055
9	0,034
10	0,025
11	0,022
12	0,022
gem	0,0300
sdev	0,01858

De aantoonbaarheidsgrens *AG* voor de extinctie y, y_{AG}, berekend met (6.28) voor 12 metingen met $n - 1 = 12 - 1 = 11$ vrijheidsgraden en een eenzijdige $t_{(0,05;11)} = 1{,}796$ (tabel 3 bijlage 1, in kolom $\alpha = 0{,}10$):

$$y_{\mathrm{AG}} = \bar{y}_{\mathrm{bl}} + \frac{t_{(0,05;n-1)}s_{\mathrm{bl}}}{\sqrt{n}} = 0{,}0300 + \frac{1{,}796 \cdot 0{,}01858}{\sqrt{12}} = 0{,}0396$$

De aantoonbaarheidsgrens *AG* voor de concentratie x, x_{AG}, berekend met de vergelijking voor de kalibratielijn is:

$$x_{\mathrm{AG}} = \frac{y_{\mathrm{AG}} - b_0}{b_1} = \frac{0{,}0396 - 0{,}0045}{16{,}54} = 2{,}12 \cdot 10^{-3} \text{ mol/L.}$$

De bepalingsgrens *BG* voor de extinctie y, y_{BG}, berekend met (6.29) voor 12 metingen met 11 vrijheidsgraden en een eenzijdige $t_{(0,05;11)} = 1{,}796$ (tabel 3 bijlage 1, in kolom $\alpha = 0{,}10$):

$$y_{\mathrm{BG}} = \bar{y}_{\mathrm{bl}} + \frac{2t_{(0,05;n-1)}s_{\mathrm{bl}}}{\sqrt{n}} = 0{,}0300 + \frac{2 \cdot 1{,}796 \cdot 0{,}01858}{\sqrt{12}} = 0{,}0493$$

De bepalingsgrens *BG* voor de concentratie x, x_{BG}, berekend met de vergelijking voor de kalibratielijn is:

$$x_{BG} = \frac{y_{BG} - b_0}{b_1} = \frac{0{,}0493 - 0{,}0045}{16{,}54} = 2{,}71 \cdot 10^{-3} \text{ mol/L.}$$

Antwoord 6.9

De aantoonbaarheidsgrens AG voor de extinctie y, y_{AG}, berekend met (6.30) voor 7 metingen met (n – 2 =) 5 vrijheidsgraden en de eenzijdige $t_{(0,05;5)}$ = 2,015 (tabel 3 bijlage 1, in kolom α = 0,10) is:

$$y_{AG} = b_0 + t_{(0{,}05;n-2)} s_{b_0} = 0{,}0230 + 2{,}015 \cdot 0{,}025 = 0{,}073$$

De bijbehorende concentratie voor de aantoonbaarheidsgrens, x_{AG}, berekend met (6.32) is:

$$x_{AG} = \frac{t_{(0{,}05;n-2)} s_{b_0}}{b_1} = \frac{2{,}015 \cdot 0{,}025}{21{,}5} = 2{,}34 \cdot 10^{-3}$$

De bepalingsgrens BG voor de extinctie y, y_{BG}, berekend met (6.31) voor 7 metingen met (n – 2 =) 5 vrijheidsgraden en de eenzijdige $t_{(0,05;5)}$ = 2,015 (tabel 3 bijlage 1, in kolom α = 0,10) is:

$$y_{BG} = b_0 + 2t_{(0{,}05;n-2)} s_{b_0} = 0{,}0230 + 2 \cdot 2{,}015 \cdot 0{,}025 = 0{,}124$$

De bijbehorende concentratie voor de bepalingsgrens, x_{BG}, berekend met (6.33) is:

$$x_{BG} = \frac{2t_{(0{,}05;n-2)} s_{b_0}}{b_1} = \frac{2 \cdot 2{,}015 \cdot 0{,}025}{21{,}5} = 4{,}69 \cdot 10^{-3}$$