
Statistiek, validatie en meetonzekerheid

voor het laboratorium

Dr. J.W.A. Klaessens

Vijfde druk

Syntax Media – Utrecht

©2021, Uitgeverij Syntax Media, Utrecht

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever. Voor zover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van Artikel 16h Auteurswet 1912 dient men de daarvoor verschuldigde vergoedingen te voldoen aan Stichting Reprorecht (www.reprorecht.nl). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) kan men zich wenden tot Stichting UvO (www.stichting-uvo.nl).

ISBN 978 94 91764 50 9

www.syntaxmedia.nl

Ontwerp omslag: Lapis Vivus grafisch ontwerp, Oosterbeek

Omslag foto: © Sebastian Czapnik - Dreamstime.com

Opmaak: AlphaZet prepress, Bodegraven

Tekstredactie: Redactie & zo, ir. Caroline van der Meulen

Eventuele opmerkingen over deze uitgave kunt u richten aan:

Uitgeverij Syntax Media

Hooghiemstraplein 124

3514 AZ Utrecht

e-mail: info@syntaxmedia.nl

Voorwoord

Statistiek heeft een vaste plaats ingenomen in het laboratorium, vaak ten dienste van het kwaliteitssysteem. Het doel van dit boek is om de laboratoriummedewerker de hulpmiddelen in handen te geven voor een juiste toepassing van de statistische technieken die nodig zijn in het laboratoriumproces. Het maakt naar onze mening weinig verschil of deze medewerker werkt in een analytisch-chemisch of een medisch laboratorium, hoewel wij ons realiseren dat het verschillende gemeenschappen zijn. De focus ligt op het routinelaboratorium, maar het boek is zeker ook toepasbaar in een onderzoekslaboratorium. Met zijn benadering vanuit de laboratoriumpraktijk is het boek zeer geschikt voor toepassing in het onderwijs. Daarbij denken wij zowel aan het Hoger Laboratorium Onderwijs, als aan de universitaire opleiding Analytische Chemie.

Het boek bevat alle statistische basistechnieken voor het laboratorium. Er is veel plaats ingeruimd voor praktijkvoorbeelden, waarbij de verschillende laboratoriumgemeenschappen aan bod komen. De uitwerking van de voorbeelden biedt een leidraad voor toepassing in de eigen praktijk van de lezer. Vanwege het praktische karakter zijn afleidingen van de formules tot een minimum beperkt.

Voldoende begrip van de theoretische achtergronden is echter cruciaal. Statistiek is niet simpelweg het volgen van een kookboekrecept. In bijna alle situaties kunnen namelijk verschillende technieken worden toegepast, maar alleen de toepassing van een geschikte techniek leidt tot een zinnig antwoord. Een ongeschikte techniek leidt ook tot een antwoord en aan dit antwoord zelf kun je niet zien dat het geen betekenis heeft. Achtergrondkennis is dus onontbeerlijk. Daarom wordt in het boek steeds aandacht besteed aan begripsvorming en onderbouwing.

De computer is onmisbaar voor de berekeningen. Gebruik van een specifiek statistisch pakket is een mogelijkheid, maar evengoed kan men ervoor kiezen om Excel te gebruiken. Dit programma wordt in bijna elk laboratorium gebruikt en bevat een groot aantal statistische functies. We kunnen stellen dat alle statistische technieken die in dit boek aan bod komen, met Excel kunnen worden berekend.

De belangrijkste wijziging in deze vijfde druk van het boek is de integratie van de statistische functies van Excel. Aanvankelijk waren deze uitgewerkt in een apart boek, maar de ervaring heeft geleerd dat het gemakkelijker en efficiënter is om ze samen te voegen. In de tekst zijn de Excel-onderdelen duidelijk gemarkeerd zodat zij gemakkelijk zijn terug te vinden.

Het boek kan ruwweg in twee delen worden onderverdeeld. In hoofdstuk 2 tot en met 6 worden de statistische basistechnieken behandeld: spreiding, statistische verdelingen, toetsen, regressierekening, variantieanalyse en bemonsteringstechnieken. De Excel-functies komen vooral in deze hoofdstukken aan de orde. In hoofdstuk 7 tot en met 11 worden de statistische technieken op het laboratoriumproces toegepast: kwaliteitscontrole, methodevalidatie, robuustheidsonderzoek, meetonzekerheid en methodenvergelijking.

Dr. Jo Klaessens

Den Dolder
augustus 2021

Inhoudsopgave

Voorwoord	5
1 Inleiding	13
1.1 Het laboratorium	13
1.2 Analyseresultaten en onzekerheid	13
1.3 Analysemethode en onzekerheid	14
2 Spreiding in analyseresultaten	17
2.1 Inleiding	17
2.2 De ware waarde en de benaderingen	18
De gemiddelde waarde	18
De mediaan	18
2.3 Grafische weergave van de spreiding van meetwaarden	20
2.4 Maat voor het uitdrukken van spreiding	22
De variatiecoëfficiënt CV	24
2.5 Onafhankelijke waarnemingen	25
2.6 De normale verdeling en de toepassing ervan	26
Standaardafwijking van een gemiddelde waarde	27
Berekenen van kansen op basis van de normale verdeling	28
Berekenen van kansen met Excel	28
2.7 Variabelen naast elkaar	31
Covariantie en correlatie	31
Propagatie van standaardafwijking in functies	33
Basisregels voor de propagatie van een standaardafwijking	33
2.8 Belangrijke verdelingen	35
De T-verdeling	35
De χ^2 -of chi-kwadraatverdeling	39
De F-verdeling	41
2.9 Gebruik van standaardafwijking	42
Afronden van meetwaarden	43
3 Hulpmiddelen voor het nemen van beslissingen	45
3.1 Beslissen	46
3.2 De methode van hypothesetoetsing	48
3.3 Z-toets	51
3.4 T-toets voor vergelijking van een meting met een groep	53
T-toets voor een enkele meting	53
T-toets voor de vergelijking van het gemiddelde van een groep met een theoretische waarde	54

3.5	T-toets voor vergelijking van groepen metingen	55
3.6	Evaluatie van spreiding	59
	Vergelijking van de spreidingen van twee groepen meetwaarden	59
	Vergelijking van een geschatte spreiding met een theoretische waarde	61
3.7	Toetsen van de verdeling van een groep waarnemingen	63
	χ^2 -toets	63
	Kolmogorov-Smirnov-toets	68
3.8	Toetsen op uitbijters	70
	Dixons Q-toets	70
	Cochran-toets	72
3.9	Toetsen wanneer de normale verdeling niet opgaat	73
	De log-normale verdeling	74
	Rangtekentoets van Wilcoxon voor gepaarde waarnemingen	75
	Steekproventoets van Wilcoxon / Mann-Whitney U-toets	77
4	Kalibreren met regressierekening	79
4.1	Inleiding	79
4.2	Lineaire regressie: het fitten van een rechte lijn	81
4.3	Onzekerheid in de regressielijn	84
	Schatting van s^2 uit regressielijn	85
	Onzekerheid in de richtingscoëfficiënt b	85
	Onzekerheid in asafsnede a	85
	Onzekerheid in de regressielijn	86
	Onzekerheid in een met regressie bepaalde concentratie	87
4.4	Opzet van de kalibratie	88
	Aantal standaarden	88
	Keuze standaarden	88
4.5	Controle op de correctheid van het regressiemodel	89
4.6	Interferentie en matrixeffecten	90
4.7	De standaardadditiemethode	92
4.8	Gewogen lineaire regressie	95
4.9	Niet-lineaire modellen	98
4.10	Kalibratie voor de routinematige analysemethode	102
	Klein werkgebied ruim boven de aantoonbaarheids-grens	103
	Klein werkgebied nabij de aantoonbaarheidsgrens	104
	Groot werkgebied: precieze resultaten in het lage én het hoge gebied	104
	Groot werkgebied, maar het lineaire gebied is beperkt	105
	Relatie tussen respons en concentratie is gebogen	106
	Bijzondere gevallen	106

5	Toepassing van variantieanalyse	107
5.1	Bronnen van spreiding	107
5.2	Uitvoering van variantieanalyse	108
5.3	Uitvoering met Excel	111
5.4	Variantieanalyse en vervolgonderzoek	113
5.5	Multifactoriële variantieanalyse	117
	Tweevoudige variantieanalyse zonder herhaling	117
	Tweevoudige variantieanalyse met herhaling	120
6	Bemonsteren	125
6.1	Inleiding	125
6.2	Bemonsteringsplan	127
	Planningsfase	128
	Optimale invulling van het ontwerp	129
6.3	Verschillende soorten monsters	130
	Representatief monster	130
	Random monstername	130
	Systematische monstername	131
	Gestratificeerde monstername ('stratified sampling')	131
	Samengesteld monster ('composite sample')	131
	Oordeelkundig bemonsteren ('authoritative/ judgemental sampling')	132
6.4	Bepalen van het aantal benodigde monsters	132
	Beschrijven van een object	132
	Bepalen van het gehalte van een object	133
6.5	Evaluatie van de mogelijke foute beslissingen	134
6.6	Het bemonsteren van industriële bulkpartijen	137
	Formule van Gy	137
6.7	Bepaling van de bemonsteringsfrequentie	138
7	Kwaliteitscontrole	141
7.1	Het laboratorium als kwaliteitssysteem	141
7.2	Controleren en beheersen van de kwaliteit	143
7.3	Eerstelijnscontrole met de controlekaart	144
	Controleregels en hun werking	145
	Soorten controlekaarten	147
	Opstart en onderhoud	147
	Uitvoering	148
	Controlemateriaal	149
	Frequentie van controlemetingen	149
	Periodieke evaluatie	150
	Microbiologie	151
	EWMA	151
7.4	Tweedelijnscontrole: interne laboratoriumcontrole	153
7.5	Derdelijnscontrole: ringonderzoek	153
	Z-scores	154
	De Youden-plot	155
	Deelname aan ringonderzoeken	157

8	Validatie van analysemethoden	159
8.1	Inleiding	159
	Validatie of verificatie	160
8.2	Validatie aan de hand van prestatiekenmerken	161
8.3	Bepalen van de prestatiekenmerken	162
	Meetbereik	163
	Meetonzekerheid	163
	Aantoonbaarheidsgrens	164
	Bepalingsgrens of kwantificeringsgrens	165
	Herhaalbaarheid	165
	Reproduceerbaarheid	167
	Juistheid	168
	Bias	168
	Terugvinding	169
	Modelafwijking	170
	Selectiviteit	171
	Geheugeneffect	172
8.4	Praktische uitvoering van de validatie	173
8.5	Eisen aan prestatiekenmerken	179
8.6	Kwalitatieve methoden	184
	<i>LOD</i> 50 als aantoonbaarheidsgrens	184
	Spreiding	186
	Kenmerken van juistheid	187
8.7	Validatie en verificatie in de microbiologie	189
8.8	Validatie en verificatie van medische analysemethoden	190
	Prestatiekenmerken	190
	Evaluatie van de prestatie	192
8.9	Fysische meetmethoden	194
9	Proefopzetten voor methodenontwikkeling en robuustheidsonderzoek	197
9.1	Inleiding	197
9.2	Begrippen	198
9.3	Proefopzet met twee factoren	199
9.4	Proefopzet met drie of meer factoren	201
	Grotere proefopzetten	202
9.5	Het bepalen van de significantie van een effect	203
9.6	Fractionele factoriële proefopzetten	206
9.7	Toepassing in methodeontwikkeling en robuustheids- onderzoek	207
10	Meetonzekerheid	211
10.1	Inleiding	211
10.2	Top-down bepaling van de meetonzekerheid	214
	Statistisch model	215
	Gebruik van interlab of intralab gegevens	216

10.3	Algemene bepalingwijze met interne laboratorium-gegevens	216
	Intralab reproduceerbaarheid	216
	Plaats van juistheid in de meetonzekerheid	218
	Het concept <i>RMS</i>	219
	Meetonzekerheid	220
10.4	Top-down meetonzekerheid voor chemische analysemethoden	222
	Vormen van juistheid	222
	Weergave van meetonzekerheid: relatief of absoluut	223
	Onderdelen van meetonzekerheid en soorten monsters	224
	Toepassing terugvinding voor bepaling meetonzekerheid	225
	Bepaling meetonzekerheid volgens NEN 7779	226
	Intralab reproduceerbaarheid als uitgangspunt	230
	Niet vertegenwoordigde bronnen van onzekerheid	231
	Benadering van VITO voor meetonzekerheid	236
10.5	Meetonzekerheid voor medische analysemethoden	237
	Total error en meetonzekerheid	241
10.6	Meetonzekerheid in de microbiologie	243
	Technische meetonzekerheid	244
	Gecombineerde meetonzekerheid	246
	Rapportage en weergave van de meetonzekerheid	248
10.7	Meetonzekerheid op basis van interlaboratorium-onderzoek	249
10.8	Bottom-up methode voor de bepaling van meetonzekerheid	253
	Het statistisch model	253
	De uitvoering	253
	Stap 1: specificatie	254
	Stap 2: identificatie van de bronnen van meetonzekerheid	254
	Stap 3: kwantificering van de onzekerheidsbronnen	254
	Stap 4: berekening gecombineerde meetonzekerheid	257
10.9	Meetonzekerheid voor fysische methoden	259
11 Hulpmiddelen voor vergelijking en ontwikkeling van analysemethoden		265
11.1	Inleiding	265
	Uitvoering van methodenvergelijking	267
11.2	Regressieanalyse met Excel	267
	Regressie van Gegevensanalyse	268
	Gegevens voor de regressie	268
	Variantieanalyse	269
	Modelparameters (de Excel-uitvoer heeft hier geen kopje boven staan)	270
	Functie LIJNSCH	270
11.3	Methodenvergelijking met regressieanalyse	272
	Geavanceerde technieken van regressierekening	276

11.4	Methodenvergelijking met de T-toets	278
11.5	Controle op de correctheid van het kalibratiemodel	281
	Visuele inspectie residuen	281
	Toets op lineariteit	281
	Toepassing van correlatiecoëfficiënt	284

Tabellen**291****Registers****311**

Nederlands – Engels		
	Statistische functies Excel	312
	Gegevensanalyse	313
Register Engels – Nederlands		314
	Statistical functions Excel	314
	Data analysis	315
Trefwoordenregister		317

Hoofdstuk 1

Inleiding

1.1 Het laboratorium

Een analytisch of klinisch laboratorium is altijd onderdeel van een groter geheel. Het ontvangt monsters van een opdrachtgever en analyseert deze, met als doel informatie te verzamelen in de vorm van analyseresultaten. Op zijn beurt gebruikt de opdrachtgever deze informatie voor zijn eigen doeleinden zoals productkeuring, procesbewaking, trouble shooting, diagnose stellen of productontwikkeling.

Het product van een laboratorium bestaat uit analyseresultaten. De omgeving van het laboratorium, dat wil zeggen de opdrachtgever, stelt eisen aan dit product. Immers, niet alleen de feitelijk gevonden analyseresultaten zijn belangrijk, ook andere aspecten spelen een belangrijke rol: de mate van onzekerheid in het resultaat (zie voorbeeld 1.1), de snelheid van analyse en rapportage en ten slotte de kosten. Een medische analyse die pas wordt gerapporteerd als de patiënt al overleden is, heeft geen waarde meer.

Voorbeeld 1.1 Onzekerheid in analyseresultaten

Een fabrikant levert een product waarvan wettelijk is bepaald dat een component een maximale concentratie mag hebben van 20 ppm. Van een batch van het product wordt een monster genomen dat wordt geanalyseerd in het laboratorium. Dit rapporteert een concentratie van 19 ppm. Het analyseresultaat heeft een onzekerheidsgebied van ± 5 ppm.

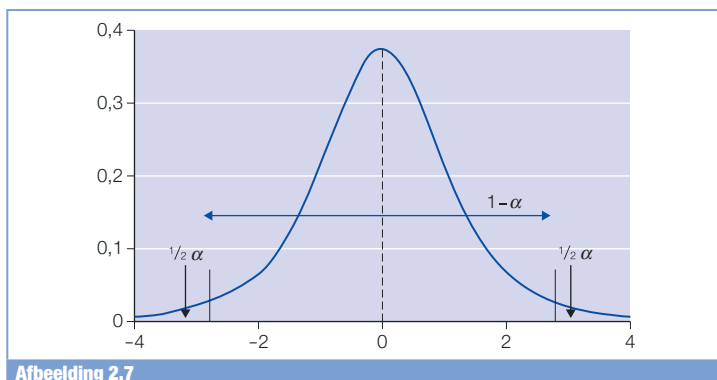
De gevonden waarde ligt weliswaar onder het wettelijke maximum, maar een behoorlijk groot gedeelte van het onzekerheidsgebied ligt erboven. Dit betekent dat er een behoorlijk grote kans is dat de concentratie in werkelijkheid boven het wettelijke maximum ligt.

Het onzekerheidsgebied heeft tot gevolg dat in deze situatie het gevonden analyseresultaat niet geschikt is om vast te stellen of de partij aan de wettelijke eis voldoet.

1.2 Analyseresultaten en onzekerheid

Voor de opdrachtgever geeft het analyseresultaat informatie over het bemonsterde object. Deze informatie is echter nooit volledig. Er is altijd een mate van onzekerheid in de wijze waarop het analyseresultaat iets zegt over het object. Deze onzekerheid is behoorlijk complex

$[-T(\alpha, v), T(\alpha, v)]$ een kansgebied vertegenwoordigt van $1 - \alpha$ (zie afbeelding 2.7). Dus als we kiezen $\alpha = 0,05$, vertegenwoordigt $[-T(0,05, v), T(0,05, v)]$ het 95%-kansgebied; dit wordt meestal het 95%-betrouwbaarheidsinterval genoemd.



Afbeelding 2.7

De T-verdeling voor 4 vrijheidsgraden met daarin aangegeven de overschrijdingskans van $\alpha = 0,05$, dat wil zeggen er is een overschrijdingskans van $1/2 \alpha$ aan de onderkant en van $1/2 \alpha$ aan de bovenkant. Ook het kansgebied $1 - \alpha$ is aangegeven dat hoort bij $-T(\alpha, v)$ en $T(\alpha, v)$; dit is het 95%-betrouwbaarheidsinterval dat ligt tussen de T-waarden $-2,776$ en $2,776$ (tabel II).



Omschrijving	Bepaling		Excel-functie
Overschrijdingskans P bij een waarde T	Bereken bij een T-verdeling met v vrijheidsgraden de dubbelzijdige overschrijdingskans P van de waarden $-T$ en T		T.VERD.2T ($T; v$)
	Bereken bij een T-verdeling met v vrijheidsgraden de linkszijdige overschrijdingskans P van de waarde $-T$		T.DIST ($T; v; \text{cumulatief}$)
	Bereken bij een T-verdeling met v vrijheidsgraden de rechtszijdige overschrijdingskans P van de waarde T		T.VERD.RECHTS ($T; v$)
Interval met overschrijdingskans P	Bereken bij een T-verdeling met v vrijheidsgraden de waarde van T zodanig dat de som van de linker- en rechteroverschrijdingskans van het interval $[-T, T]$ gelijk is aan P		T.INV.2T ($P; v$)
	Bereken bij een T-verdeling met v vrijheidsgraden de waarde van T zodanig dat de linkeroverschrijdingskans P bedraagt		T.INV($P; v$)

Voorbeeld 2.9**Berekening betrouwbaarheidsinterval op basis van kleine steekproef**

Het laboratorium is bezig met de ingangscntrole van een product. Volgens de specificaties mag dit product maximaal 9 ppm van een bepaalde storende component bevatten. Drie monsters worden geanalyseerd waarbij de volgende resultaten worden verkregen: 7,7, 8,6 en 8,8 ppm. Het is bekend dat de component homogeen verdeeld is. Als er geen verdere informatie is over de spreiding, binnen welk interval mogen we met 95% kans een volgend analyseresultaat verwachten? En binnen welk interval zal met 95% kans de ware waarde van de concentratie liggen?



Voor de berekening van het interval passen we T.INV.2T toe. Deze levert niet meteen de gezochte uitkomst, maar een genormaliseerde waarde T voor het interval $[-T, T]$. Deze moeten we omrekenen naar de concentraties in ppm. De bewerking gebeurt in de volgende stappen:

1. We zijn op zoek naar het 95%-betrouwbaarheidsinterval, dus er is een overschrijdingskans van 5% (links en rechts samen). Er zijn drie metingen, dus $\nu = 2$. Bereken dan $T = T.INV.2T(5\%;2)$. Dit levert de waarde 4,303.

	A	B	C	D	E
1	Metingen				
2	7,7		1) T =	4,303	
3	8,6		2) gem =	8,37	
4	8,8		3) s =	0,59	
5					
6			4) volgende meting (95 %)		
7			ondergrens	5,8	
8			bovengrens	10,9	
9					

Afbeelding 2.8

Excel-berekeningen voor de bepaling van het 95%-betrouwbaarheidsinterval.

Omrekenen naar concentratie gebeurt met formule 2.17. Hiervoor hebben we het gemiddelde en de standaardafwijking nodig.

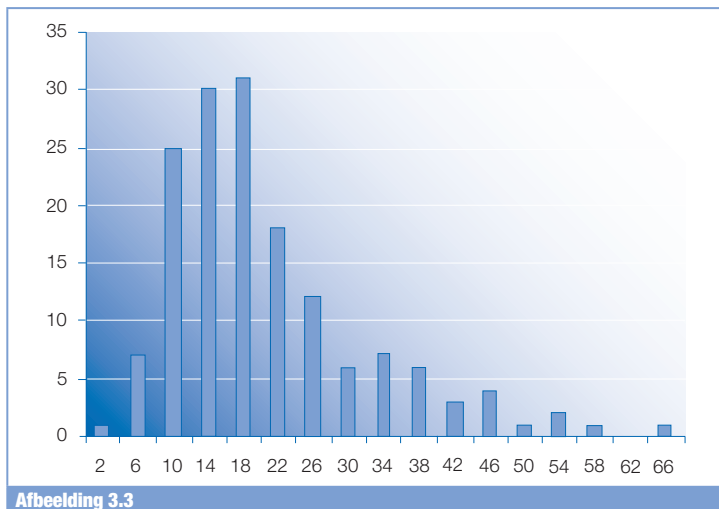
2. Bereken gemiddelde waarde $\bar{x} = \text{GEMIDDELDE}(A2:A4) = 8,37$.
3. Bereken standaardafwijking $s = \text{STDEV.S}(A2:A4) = 0,59$.
4. Bereken met formule 2.17 vervolgens:
 ondergrens: $\bar{x} - T \cdot s = D3 - D2 * D4 = 5,8$.
 bovengrens: $\bar{x} + T \cdot s = D3 + D2 * D4 = 10,9$.

Wanneer we een vierde meting zouden uitvoeren, mogen we deze met 95%-kans verwachten tussen 5,8 en 10,9 ppm. Deze kan dus ook nog boven de specificatie van 9 ppm liggen.

Voor de bepaling van het 95%-betrouwbaarheidsinterval van de ware waarde zijn de stappen 1 – 3 gelijk. De beste schatting van de ware waarde is het gemiddelde van de drie metingen. Dus dat is het centrum van het interval.

Voorbeeld 3.10 Verdeling voor normaalwaarden

Een klinisch laboratorium wil de normaalwaarden bepalen voor de bepaling van Alanine Aminotransferase (ALAT) in serum. Monsters van 155 gezonde personen zijn geanalyseerd. Uit de meetwaarden is berekend: $\bar{x} = 20,76$ en $s = 11,58$ U/L. In afbeelding 3.3 is een histogram met de meetgegevens weergegeven.



Afbeelding 3.3

Histogram van 155 metingen van ALAT in serum bij gezonde personen; in U/L.



Bij een normale verdeling kan als normaalwaarden worden gekozen $\bar{x} \pm 2 \cdot s$. De vraagstelling is of ALAT in serum normaal verdeeld is. We onderzoeken dit met de χ^2 -toets. Omdat \bar{x} en s bepaald zijn uit de meetwaarden, moeten we de werkwijze met CHIKW.INV.RECHTS toepassen.

1. H_0 : de verkregen frequentieverdeling volgt een verdeling met $\mu = 20,76$ en $\sigma = 11,58$.
2. Kies $\alpha = 0,05$.
3. H_1 : de verkregen frequentieverdeling wijkt af van deze normale verdeling.
4. Zoals te zien in afbeelding 3.3 zijn er intervallen met minder dan vijf waarnemingen. Aan de uiteinden worden twee, respectievelijk zes intervallen samengevoegd. Tabel 3.9 geeft de intervallen en hun frequenties aan in kolommen A en B.
5. Bereken op basis van een normale verdeling en $\bar{x} = 20,76$ en $s = 11,58$ de theoretische frequenties van de intervallen. Voer de volgende berekeningen uit:
 - de cumulatieve kansen tot de bovengrenzen van de intervallen (kolom C);
 - de kansen van de intervallen (kolom D);
 - de frequenties in aantallen (kolom E);
 - de toetswaarde met de formule van tabel 3.7 (kolom F);
 - de kritieke waarde $\chi^2(\alpha, \nu) = 14,07$ (cel F18).
6. Bepaal $\chi^2(\alpha, \nu) = 43,64$ (cel F17). Omdat deze waarde groter is dan de kritieke waarde verwerpen we H_0 .

a)

	A	B	C	D	E	F
1						
2		waarnemingen =	155			
3		gem =	20,76			
4		st.afw. =	11,58			
5						
6	grens interval	f_i	cumulatieve kans	kans interval	e_i	$(f_i - e_i)^2 / e_i$
7	<8	8	=NORM.VERD.N(8;C\$3;C\$4;1)	=C7	=D7*C\$2	=(B7-E7)^2/E7
8	12	25	=NORM.VERD.N(A8;C\$3;C\$4;1)	=C8-C7	=D8*C\$2	=(B8-E8)^2/E8
9	16	30				
10	20	31				
11	24	18				
12	28	12				
13	32	6				
14	36	7				
15	40	6				
16	>40	12				
17					som =	=SOM(F7:F16)
18					krit.waarde =	=CHIKW.INV.RECHTS(5%;7)
19						

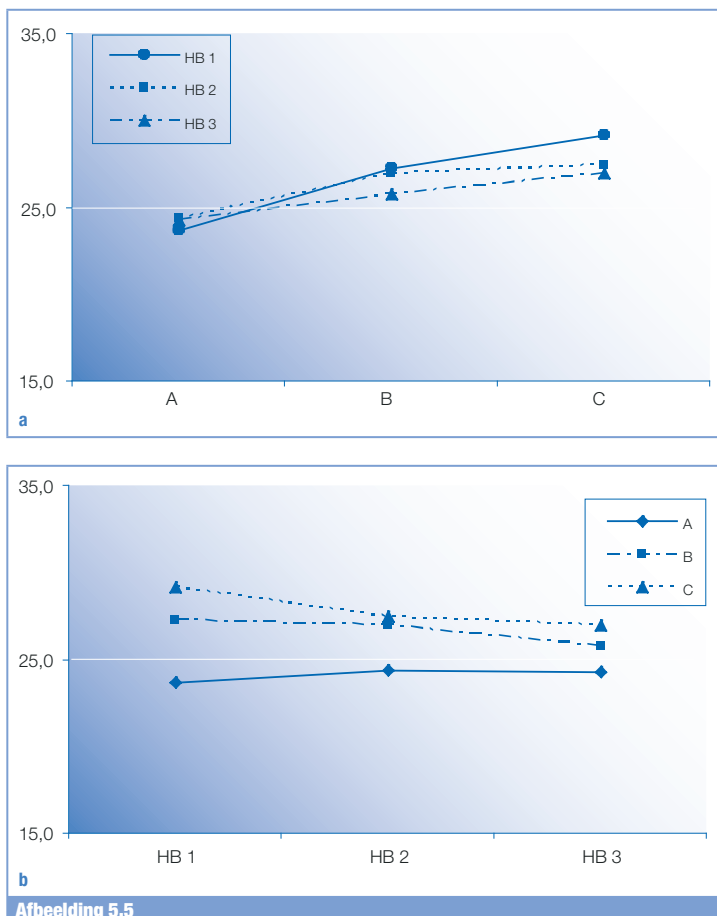
b)

	A	B	C	D	E	F
1						
2		waarnemingen =	155			
3		gem =	20,76			
4		st.afw. =	11,58			
5						
6	grens interval	f_i	cumulatieve kans	kans interval	e_i	$(f_i - e_i)^2 / e_i$
7	<8	8	14%	13,5%	21,0	8,02
8	12	25	22%	8,9%	13,9	8,95
9	16	30	34%	11,6%	18,0	8,09
10	20	31	47%	13,3%	20,7	5,17
11	24	18	61%	13,6%	21,1	0,46
12	28	12	73%	12,4%	19,2	2,70
13	32	6	83%	10,0%	15,5	5,83
14	36	7	91%	7,2%	11,1	1,53
15	40	6	95%	4,6%	7,1	0,17
16	>40	12		4,8%	7,5	2,71
17					som =	43,64
18					krit.waarde =	14,07
19						

Tabel 3.9

Uitvoering van de χ^2 -toets op normale verdeling met Excel. a) De formules: rij 9 t/m 15 bevatten formules die analoog zijn aan die van rij 8. b) De getallen en de resultaten van de bewerkingen.

We concluderen dat ALAT in serum niet normaal verdeeld is. De voorgestelde werkwijze voor de bepaling van de normaalwaarden is dus niet geschikt. We zullen een andere werkwijze moeten volgen, bijvoorbeeld direct op de kansverdeling zonder aanname van normaliteit.



Afbeelding 5.5

Grafische weergave van de gemiddelde waarden per behandeling van de gegevens. a) Legeringen als x-as. b) Hittebehandelingen als x-as.

Uit de variantieanalyse blijkt dat de effecten van de hittebehandeling en de interactie tussen beide factoren niet significant zijn. Afbeelding 5.5a lijkt wel enig effect van de hittebehandeling te tonen, in combinatie met een interactie met de legeringen. Dit blijkt uit de kruisende lijnen. Maar de beschikbare gegevens geven geen uitsluitel of deze effecten echt zijn of slechts toevallig. Mocht dit cruciaal zijn voor het onderzoek, dan kunnen we overwegen om meer duidelijkheid te krijgen door extra metingen.

Noten

- 1 ISO 13528:2015 Statistical methods for use in proficiency testing by interlaboratory comparison.
- 2 L. Davies, Efficiency in research, development and production: the statistical design and analysis of chemical experiments, Royal Society of Chemistry, 1993, pag. 51.

Hoofdstuk 6

Bemonsteren

Samenvatting

De onzekerheid waarmee een analyseresultaat iets zegt over het onderzochte object of de partij wordt mede bepaald door de wijze van bemonsteren. Deze wordt vastgelegd in het bemonsteringsplan.

Het bemonsteringsplan bestaat uit een ontwerp waarin met statistische hulpmiddelen bepaald wordt hoeveel monsters genomen worden, hoe en waar zij genomen worden in het object of de partij. Daarnaast is er een implementatiefase waarin ook de fysieke hulpmiddelen en eventuele wijze van monster conservering worden vastgelegd.

Er bestaan verschillende soorten monsters: representatief monster, random monster, systematisch monster, gestratificeerd monster, samengesteld monster, oordeelkundig monster.

Er zijn geen eenvoudige, algemeen toepasbare regels om het aantal benodigde monsters te bepalen. Vaak kunnen we wel heel specifieke regels opstellen door gebruik te maken van karakteristieken van het toepasingsgebied.

6.1 Inleiding

Voor het bepalen van de concentratie van een component willen we over het algemeen niet het hele object of de hele partij analyseren. We zullen gaan bemonsteren: een klein deel wordt geïsoleerd en naar het laboratorium gezonden ter analyse. Bijvoorbeeld een monster van een hoeveelheid te verplaatsen grond, een bloedmonster van een patiënt, een hoeveelheid schraapsel van een metalen object. Bij de meeste analysetechnieken analyseren we ook niet het hele monster, maar slechts een gedeelte. Er vindt dus ook nog een sub-bemonstering plaats. Vaak wordt de algemene term populatie gebruikt in plaats van het te onderzoeken object of de partij.

In het theoretische geval dat het object homogeen is, is bemonsteren statistisch gezien zeer eenvoudig. Homogeen betekent dat de te onderzoeken component gelijkmatig verdeeld is over het hele object. Elk deel heeft dus precies dezelfde concentratie. Wanneer het object niet homogeen is, wordt het heterogeen genoemd. Bij een homogeen materiaal is het namelijk voldoende om ergens uit het object een hoeveelheid materiaal te verzamelen en te analyseren. Het maakt niet uit

Tabel II

b) Waarden van de T-verdeling als functie van het aantal vrijheidsgraden v voor $\alpha = 0,05$ en $\alpha = 0,01$. Enkelzijdig toetsen.

Enkelzijdig					
v	$\alpha = 0,05$	$\alpha = 0,01$	v	$\alpha = 0,05$	$\alpha = 0,01$
1	6,314	31,821	28	1,701	2,467
2	2,920	6,965	29	1,699	2,462
3	2,353	4,541	30	1,697	2,457
4	2,132	3,747	31	1,696	2,453
5	2,015	3,365	32	1,694	2,449
6	1,943	3,143	33	1,692	2,445
7	1,895	2,998	34	1,691	2,441
8	1,860	2,896	35	1,690	2,438
9	1,833	2,821	36	1,688	2,434
10	1,812	2,764	37	1,687	2,431
11	1,796	2,718	38	1,686	2,429
12	1,782	2,681	39	1,685	2,426
13	1,771	2,650	40	1,684	2,423
14	1,761	2,624	45	1,679	2,412
15	1,753	2,602	50	1,676	2,403
16	1,746	2,583	55	1,673	2,396
17	1,740	2,567	60	1,671	2,390
18	1,734	2,552	65	1,669	2,385
19	1,729	2,539	70	1,667	2,381
20	1,725	2,528	75	1,665	2,377
21	1,721	2,518	80	1,664	2,374
22	1,717	2,508	85	1,663	2,371
23	1,714	2,500	90	1,662	2,368
24	1,711	2,492	95	1,661	2,366
25	1,708	2,485	100	1,660	2,364
26	1,706	2,479	∞	1,645	2,327
27	1,703	2,473			

Tabel III a)

Waarden van de F -verdeling – dubbelzijdig. $\alpha = 0,05$; ν_1 en ν_2 zijn de aantallen vrijheidsgraden van de grootste respectievelijk kleinste variantie.

<i>Dubbelzijdig</i>										
ν_1	ν_2 1	2	3	4	5	6	7	8	9	10
1	647,79	38,51	17,44	12,22	10,01	8,81	8,07	7,57	7,21	6,94
2	799,48	39,00	16,04	10,65	8,43	7,26	6,54	6,06	5,71	5,46
3	864,15	39,17	15,44	9,98	7,76	6,60	5,89	5,42	5,08	4,83
4	899,60	39,25	15,10	9,60	7,39	6,23	5,52	5,05	4,72	4,47
5	921,83	39,30	14,88	9,36	7,15	5,99	5,29	4,82	4,48	4,24
6	937,11	39,33	14,73	9,20	6,98	5,82	5,12	4,65	4,32	4,07
7	948,20	39,36	14,62	9,07	6,85	5,70	4,99	4,53	4,20	3,95
8	956,64	39,37	14,54	8,98	6,76	5,60	4,90	4,43	4,10	3,85
9	963,28	39,39	14,47	8,90	6,68	5,52	4,82	4,36	4,03	3,78
10	968,63	39,40	14,42	8,84	6,62	5,46	4,76	4,30	3,96	3,72
11	973,03	39,41	14,37	8,79	6,57	5,41	4,71	4,24	3,91	3,66
12	976,72	39,41	14,34	8,75	6,52	5,37	4,67	4,20	3,87	3,62
13	979,84	39,42	14,30	8,72	6,49	5,33	4,63	4,16	3,83	3,58
14	982,55	39,43	14,28	8,68	6,46	5,30	4,60	4,13	3,80	3,55
15	984,87	39,43	14,25	8,66	6,43	5,27	4,57	4,10	3,77	3,52
16	986,91	39,44	14,23	8,63	6,40	5,24	4,54	4,08	3,74	3,50
17	988,72	39,44	14,21	8,61	6,38	5,22	4,52	4,05	3,72	3,47
18	990,35	39,44	14,20	8,59	6,36	5,20	4,50	4,03	3,70	3,45
19	991,80	39,45	14,18	8,58	6,34	5,18	4,48	4,02	3,68	3,44
20	993,08	39,45	14,17	8,56	6,33	5,17	4,47	4,00	3,67	3,42
21	994,30	39,45	14,16	8,55	6,31	5,15	4,45	3,98	3,65	3,40
22	995,35	39,45	14,14	8,53	6,30	5,14	4,44	3,97	3,64	3,39
23	996,34	39,45	14,13	8,52	6,29	5,13	4,43	3,96	3,63	3,38
24	997,27	39,46	14,12	8,51	6,28	5,12	4,41	3,95	3,61	3,37
25	998,09	39,46	14,12	8,50	6,27	5,11	4,40	3,94	3,60	3,35
26	998,84	39,46	14,11	8,49	6,26	5,10	4,39	3,93	3,59	3,34
27	999,54	39,46	14,10	8,48	6,25	5,09	4,39	3,92	3,58	3,34
28	1000,2	39,46	14,09	8,48	6,24	5,08	4,38	3,91	3,58	3,33
29	1000,8	39,46	14,09	8,47	6,23	5,07	4,37	3,90	3,57	3,32
30	1001,4	39,46	14,08	8,46	6,23	5,07	4,36	3,89	3,56	3,31
35	1003,8	39,47	14,06	8,43	6,20	5,04	4,33	3,86	3,53	3,28
40	1005,6	39,47	14,04	8,41	6,18	5,01	4,31	3,84	3,51	3,26
45	1007,0	39,48	14,02	8,39	6,16	4,99	4,29	3,82	3,49	3,24
50	1008,1	39,48	14,01	8,38	6,14	4,98	4,28	3,81	3,47	3,22
60	1009,8	39,48	13,99	8,36	6,12	4,96	4,25	3,78	3,45	3,20
70	1011,0	39,48	13,98	8,35	6,11	4,94	4,24	3,77	3,43	3,18
80	1011,9	39,49	13,97	8,33	6,10	4,93	4,23	3,76	3,42	3,17
90	1012,6	39,49	13,96	8,33	6,09	4,92	4,22	3,75	3,41	3,16
100	1013,2	39,49	13,96	8,32	6,08	4,92	4,21	3,74	3,40	3,15
∞	1018,3	39,50	13,90	8,26	6,02	4,85	4,14	3,67	3,33	3,08