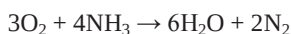


De ontdekking van argon en het belang van een nauwkeurige data-analyse

John William Strutt, derde baron van Rayleigh (1842-1919), is één van de meest indrukwekkende figuren in de geschiedenis van de natuurwetenschappen. Hij is het meest bekend door zijn verklaring van de blauwe kleur in de lucht veroorzaakt door *Rayleighverstrooiing*. Minder bekend is dat Rayleigh één van de ontdekkers was van argon. In dit artikel gaan we in op de manier waarop Rayleigh dit gas ontdekte, en wordt een visuele en een statistische analyse van zijn meetgegevens gepresenteerd.

In de tijd van Rayleigh was bekend dat lucht in de atmosfeer bestond uit ongeveer één-vijfde zuurstofgas en vier-vijfde stikstofgas. Het isoleren van stikstofgas uit lucht was grotendeels gebaseerd op het verwijderen van zuurstofgas uit deze lucht (Giunta, 2012). Een manier om dit te doen was door een luchtstaal door een hete koperen buis te leiden, zodat koper(II)oxide kon worden gevormd. Vervolgens kon Rayleigh de dichtheid van het overblijvende gas bepalen door van een vast volume, bij een constante temperatuur en druk, de massa te meten. Een andere manier was door lucht door vloeibare ammoniak te borrelen en vervolgens door een hete koperen buis te leiden. De reactie die optreedt in vloeibare ammoniak luidt:



Het water, gevormd tijdens de reactie, werd verwijderd via droogmiddelen. Het product van de eerste methode bestond voornamelijk uit atmosferisch stikstofgas, terwijl het product van de tweede methode voornamelijk bestond uit atmosferisch stikstofgas en chemisch stikstofgas. Omdat atmosferische lucht argon bevat (wat men niet wist in die tijd), zorgde het toevoegen van het stikstofgas uit ammoniak ervoor dat de verhouding tussen het stikstofgas en argon veranderde. Door het kleine verschil in samenstelling van de producten die gelijke volumes bezaten en die werden bereid door verschillende methoden bij dezelfde temperatuur en druk, mat Rayleigh licht verschillende massa's. Rayleigh

vond dit kleine verschil (ongeveer 1/1000^{ste}) opmerkelijk. Overtuigd van de correctheid van zijn meting vroeg hij in 1892 aan de wetenschappelijke wereld waaraan dit verschil zou kunnen liggen (Rayleigh, 1892). Gesteund door zijn collega scheikundige, Sir William Ramsay, die eveneens verbaasd was door deze opmerkelijke discrepantie, begon Rayleigh aan een reeks experimenten waarin hij atmosferisch en chemisch stikstofgas bereidde. Het chemische stikstofgas werd gevormd uit lachgas, stikstofmonoxide of ammonium-nitriet. De door Rayleigh gemeten massa's van constante volumes (bij constante temperatuur en druk) van de twee typen stikstofgas staan weergegeven in Tabel 1 (Harris, 2007).

Atmosferisch N ₂ (g)	Chemisch N ₂ (g)
2,31017	2,30143
2,30986	2,29890
2,31010	2,29816
2,31001	2,30182
2,31024	2,29869
2,31010	2,29940
2,31028	2,29849
	2,29869

Tabel 1: Massa's van atmosferisch en chemisch stikstofgas

De gemiddelde massa van atmosferisch stikstof (2,31011) is slechts een klein beetje groter dan de gemiddelde massa van het chemisch stikstof (2,29945). Rayleigh besloot dat dit verschil buiten de experimentele fout moest liggen en dat er in het atmosferische stikstof een zwaardere component aanwezig

moest zijn. In 1895 publiceerden Rayleigh en Ramsay hun ontdekking van argon (Rayleigh & Ramsay, 1895).

Belangrijk om op te merken is dat enkel de gemiddelden van de gemeten resultaten werden vergeleken om te besluiten dat er een opmerkelijk verschil bleek te zijn. Met hedendaagse statistische methoden zou deze manier van werken nogal primitief kunnen worden genoemd. Desalniettemin ontvingen Rayleigh en Ramsay voor hun nauwgezette manier van meten in 1904 respectievelijk de Nobelprijs voor de fysica en de Nobelprijs voor de chemie.

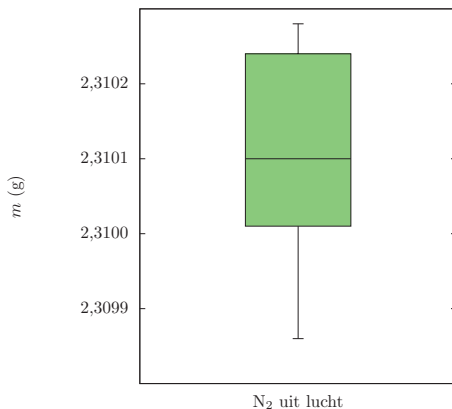
Box-plots

Een beter visueel inzicht in het verschil in massa's kan verkregen worden door middel van box-plots (Larsen, 1990). Een box-plot geeft de middelste positie weer van de data in een doos (box) waarbij de waarde van de mediaan wordt aangegeven door een horizontale streep in de doos. De onderste en de bovenste horizontale lijnen in de doos komen overeen met respectievelijk het eerste en het derde kwartiel (Q_1 en Q_3). De onderste en bovenste lijnstukken behoren in dit voorbeeld tot de uiterste waarden van de datareeks. De waarden kunnen relatief eenvoudig in een spreadsheetprogramma worden berekend en staan weergegeven in Tabel 2.

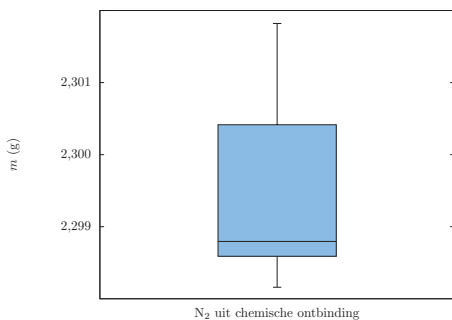
Figuur 1 (a) en Figuur 1 (b) geven de box-plots voor de metingen van Rayleigh weer en in Figuur 2 worden beide box-plots met elkaar vergeleken.

	Atmosferisch N ₂ (g)	Chemisch N ₂ (g)
Mediaan	2,3101	2,298795
Minimumwaarde	2,30986	2,29816
Kwartiel 25%	2,310055	2,29864
Kwartiel 50%	2,3101	2,298795
Kwartiel 75%	2,30986	2,29816
Maximumwaarde	2,31028	2,30182

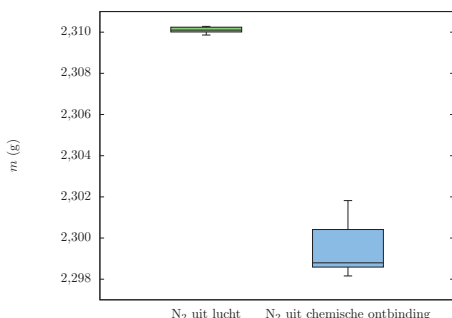
Tabel 2: Gegevens voor de box-plots



Figuur 1a: Box-plot van atmosferisch stikstofgas



Figuur 1b: Box-plot van chemisch stikstofgas



Figuur 2: Vergelijking van de twee boxplots

Het verschil tussen de massa's van atmosferisch stikstofgas en chemisch stikstofgas is nu des te opmerkelijker.

Tweezijdige t-toets

Om twee datareeksen met elkaar te vergelijken, gebruiken chemici tegenwoordig statistische toetsen. Een dergelijk voorbeeld is

de *t*-toets die werd ontwikkeld door William Gosset en die door hem werd gepubliceerd in 1908 (dus na het toekennen van de Nobelprijzen aan Ramsay en Rayleigh) onder het pseudoniem 'Student' (Student, 1908). Een tweezijdige *t*-toets kan worden uitgevoerd met behulp van het *Analysis ToolPak* in Microsoft Excel om de nulhypothese te toetsen dat er geen verschil is tussen de gemiddelde massa's van de twee typen stikstofgas. De alternatieve hypothese is dat er wel degelijk een verschil is tussen de gemiddelde massa's van de twee typen stikstofgas. De resultaten van de in Excel berekende tweezijdige *t*-toets met gelijke varianties op basis van de gegevensreeksen in Tabel 1 staan weergegeven in Tabel 3.

Het komt er nu op aan om de getallen goed te interpreteren zonder dat in dit artikel al te diep kan worden ingegaan op de statistische achtergrond. We kijken eerst naar 'T-statistische gegevens' (t_{berekend}) en 'Kritiek gebied van T' (t_{tabel}). Daar $t_{\text{berekend}} (20.06) > t_{\text{tabel}} (2.16)$, besluiten we dat de twee gemiddelden niet hetzelfde zijn omdat het verschil *significant* is. ' $P(T \leq t)$ tweezijdig' in Tabel 3 is de waarschijnlijkheid om de twee gemiddelden en de standaardafwijkingen waar te nemen als gevolg van toeval de gemiddelde waarden echt hetzelfde zouden zijn. We zien dat $P(T \leq t) = 3.65 \times 10^{-11}$ zodat we kunnen besluiten dat het

verschil *zeer significant* is. Voor elke waarde $P(T \leq t) \leq 0.05$ zal de nulhypothese verworpen worden en zal men besluiten dat de gemiddelden niet hetzelfde zijn.

Besluit

Bestudering van specifieke historische wetenschappelijke ontwikkelingen kan bijdragen aan inzicht van leerlingen in wetenschappelijke methoden en werkwijzen. Via de ontdekking van argon lieten wij het belang van nauwkeurig meten en het belang van statistische analyse van meetgegevens zien. Dit voorbeeld is goed bruikbaar in de bovenbouw van de havo en het vwo. Enkel bij het gebruik van de *t*-toets is specifieke statistische kennis vereist. ●

BRONNEN

- Giunta, C. J. (2012). Using History To Teach Scientific Method: The Case of Argon. *Journal of Chemical Education*, 75 (10), 1322–1325.
- Harris, D. C. (2007). *Quantitative Chemical Analysis*. Seventh Edition. New York: W. H. Freeman and Company.
- Larsen, R. D. (1990). Lessons Learned from Lord Rayleigh on the Importance of Data Analysis. *Journal of Chemical Education*, 67 (11), 925–928.
- Rayleigh, L. (1892). Density of nitrogen. *Nature*, 46, 512–513. <http://www.nature.com/physics/looking-back/rayleigh/Rayleigh.pdf> (laatst bezocht in januari 2017).
- Rayleigh, L. & Ramsay, W. (1895). Argon, a New Constituent of the Atmosphere. *Proceedings of the Royal Society of London*, 57, 265–287. <http://www.periodictable.ru/018Ar/history/Ar1.pdf> (laatst bezocht in januari 2017).
- Student (1908). The probable error of a mean. *Biometrika* 6 (1), 1–25. <http://www.york.ac.uk/depts/maths/histstat/student.pdf> (laatst bezocht in januari 2017).

Tabel 3: Resultaten van de t-toets

	Variabele 1	Variabele 2
Gemiddelde	2,310108571	2,2994475
Variantie	$2,03476 \times 10^{-8}$	$1,94045 \times 10^{-6}$
Waarnemingen	7	8
Gepaarde variatie	$1,05425 \times 10^{-6}$	
Schatting van verschil tussen gemiddelden	0	
Vrijheidsgraden	13	
T- statistische gegevens	20,06215856	
$P(T \leq t)$ eenzijdig	$1,82622 \times 10^{-11}$	
Kritiek gebied van T-toets: eenzijdig	1,770933396	
$P(T \leq t)$ tweezijdig	$3,65244 \times 10^{-11}$	
Kritiek gebied van T-toets: tweezijdig	2,160368656	