



Statistiek in onderzoek

Arno Theune

Een andere statistiekmodule

29 februari 2024

Voorstellen

- Arno Theune

a.theune@paxchristicollege.nl

Werkzaam aan Pax Christi College Druten, 30^e jaar, docent wiskunde

Docent NLT sinds 15 jaar



NLT voor VWO in de regio Nijmegen

- **Uniek (?) samenwerkingsverband**

Radboud Universiteit Nijmegen

6 middelbare scholen: Pax Christi College, Druten; Stedelijk Gymnasium Nijmegen; Stedelijke Scholengemeenschap Nijmegen; Kandinsky College, Nijmegen; NSG Groenewoud, Nijmegen; Dominicus College, Nijmegen (laatste jaar)

- **Voor klas 5 en 6 VWO**

9 modules op de Radboud Universiteit, 2 modules, in klas 4, op school

- **Onderwijs**

Colleges door medewerkers en hoogleraren universiteit, ook eigen docenten

Gebruik maken van faciliteiten van de universiteit (practica, onderzoeksgroepen)



Deze samenwerking

- **Voordelen**

Kennis en faciliteiten van de universiteit: meer mogelijkheden dan op school

Collega's van verschillende scholen, vanuit verschillende vakken, met verschillende ideeën en invalshoeken: inspiratiebron

Als een docent ziek is, kunnen anderen het overnemen (als de module al een tijdje loopt, en de betrokken collega's de module goed kennen)

- **Nadelen**

Je moet het wel eens zijn over het programma: werkgroep

Verschillen in PTA's zorgen soms voor verschillen in toepassen slaag-zakregeling

Vervoer, als de school niet dicht in de buurt zit



Modules

- **We geven (onder andere) de volgende modules:**
 - Atmosfeer en klimaatverandering (aardrijkskunde, natuurkunde)
 - Statistiek in onderzoek (wiskunde, natuurkunde, scheikunde (?))
 - Artificial intelligence (natuurkunde, informatica, wiskunde)
 - Hersenen en leren (biologie, natuurkunde)
 - De toekomst van de landbouw (biologie, scheikunde)
- **Modulekeuze**
 - Binnen de werkgroep (elke school vertegenwoordigd, vertegenwoordiging vanuit de universiteit)
 - Modules gaan en komen, nu een accent op duurzaamheid bij de keuze voor nieuwe modules



Samenwerking met de Radboud Universiteit

- **Gastsprekers**

Hoogleraren, UD's

- **Faciliteiten**

Ondersteuning door studenten, practica die op school niet mogelijk zijn, bezoek aan onderzoeksgroepen op de universiteit (Donders-instituut)



Statistiek in onderzoek

- **Statistiek is relevant!**

 - Waarom veel metingen bij onderzoek?

 - Meetonzekerheid en standaardafwijking

 - Normale verdeling (veel leerlingen zonder wiskunde A (of D))

 - Hypothesetoetsen (veel leerlingen zonder wiskunde A (of D))

- **Relevant voor verschillende richtingen:**

 - Bij de 'harde' β -wetenschappen

 - Bij medische vakken (geneeskunde, gezondheids- en bewegingswetenschappen)

 - Psychologie, sociologie, ...



De module Statistiek in onderzoek in het kort (1)

- **Hoofdstuk 1: Beschrijvende statistiek**

Meetonzekerheden, aan de hand van het rolgootexperiment

Gemiddelde en standaardafwijking bij één meting

Gemiddelde en standaardafwijking bij meerdere metingen: steekproevenverdeling

Samenvoegen van verschillende metingen

- **Hoofdstuk 2: Hypothesetoetsen**

De t -toets voor gekoppelde paren en voor onafhankelijke varianties

Interpretatie en kanttekeningen bij significantietoetsen



De module Statistiek in onderzoek in het kort (2)

- **Hoofdstuk 3: Onzekerheidsvoortplanting en regressieanalyse**
 - Onzekerheidsvoortplanting: één of meer variabelen bij de slingerproef
(Dus twee metingen: slingertijd en slingerlengte)
 - Regressieanalyse: fitten van functies
 - Programma SciDAViS
(Scientific Data Analysis and Visualisation)
- **Toetsing**
 - 2 verslagen: rolgootpracticum en slingerproef
 - 1 schriftelijke toets over de t -toets
(Althans, zo doen wij het)



Vergelijking met Meten en interpreteren (1)

Onderwerp	Meten en interpreteren	Statistiek in onderzoek
Meetschalen (nominaal-ratio)	Ja	Nee
Toevallige en systematische fout	Ja	Ja
Meetonzekerheden	Ja	Nee
Doorwerking	Basaal	Grondig(er)
Verschil tussen SD en SE	Basaal	Grondig(er)
Filteren meetgegevens	Ja	Ja
Combineren gegevens	Nee	Ja
Kansaspecten bij meetonzekerheden	Nee	Ja
Werken met grote dataset	Ja	Nee

Vergelijking met Meten en interpreteren (2)

Onderwerp	Meten en interpreteren	Statistiek in onderzoek
Lineaire regressie	Ja	Nee
Kleinste kwadratenmethode	Ja	Nee
Fitten van grafieken	Alleen lineair	Ja
Correlatiecoëfficiënten	Ja: pmcc, rang-	Nee
Hypothesetoetsen	Nee	Ja: <i>t</i> -toets
Significantie	Nee	Ja
Fouten van 1 ^e en 2 ^e soort	Nee	Ja

Samengevat

- **Meten en interpreteren**

 - Grotere variatie aan onderwerpen

 - Daardoor niet de ruimte om bij alles een grondige behandeling te geven

 - Pluspunt: correlatiecoëfficiënten

- **Statistiek in onderzoek**

 - Minder onderwerpen

 - Grondigere uitdieping bij een aantal onderwerpen

 - Pluspunt: hypothesetoetsen

- **Statistiek in onderzoek vs Meten en interpreteren**

 - Geen kwalitatief verschil, wel wat te kiezen



Statistiek in onderzoek, hoofdstuk 1

- Beschrijvende statistiek

Alles met metingen aan het
rolgootexperiment

Metten en significantie

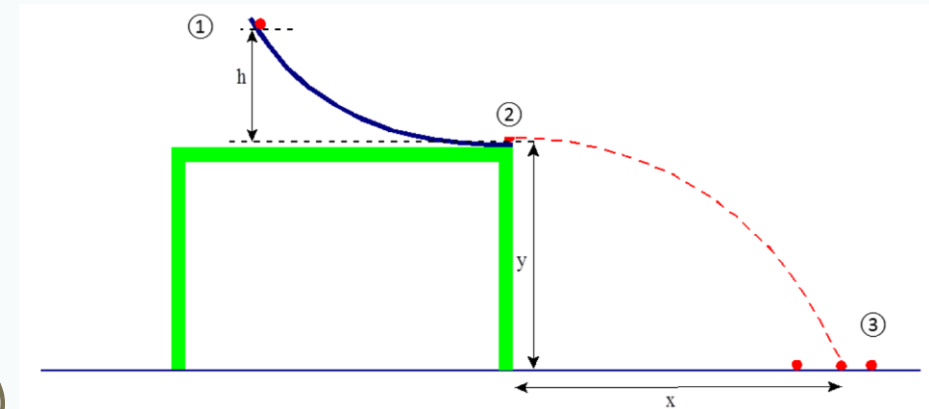
Eén reeks metingen = één steekproef:
gemiddelde en standaardafwijking (SD)

De normale verdeling

Meerdere meetreeksen: de steekproevenverdeling (SE)

Soorten 'fouten': toevallig en systematisch

Samenvoegen van verschillende meetreeksen: hoe en onder welke voorwaarden



Statistiek in onderzoek, hoofdstuk 2

- Hypothesetoetsen

- t -toets voor gekoppelde paren

- t -toets voor onafhankelijke varianties

- t -waarden, p -waarden, significantieniveau

- Door kennis van de normale verdeling begrip van de t -verdeling (WisB-leerlingen!)

- Trekken van conclusies, zuiver formuleren

- Kanttekeningen



Statistiek in onderzoek, hoofdstuk 3

- Onzekerheidsvoortplanting en regressieanalyse

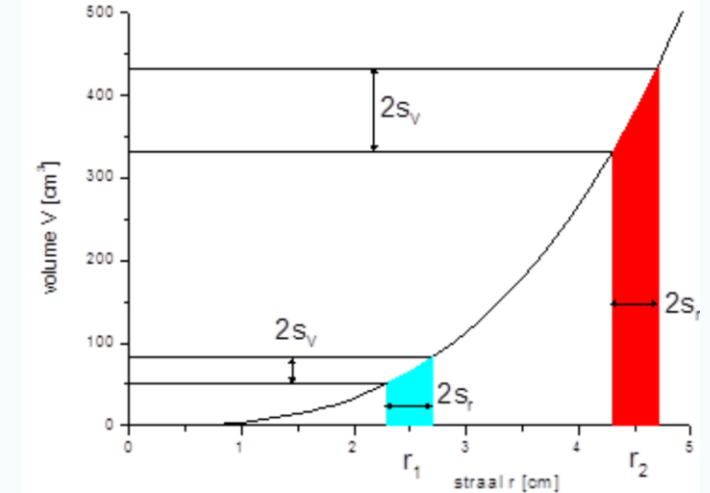
Alles met metingen aan de slingerproef

- Eerst een aantal metingen bij dezelfde lengte
- Daarna metingen bij verschillende lengtes
- We zoeken met de formule $T = 2\pi \sqrt{\frac{l}{g}}$ de waarde van g

Onzekerheidsvoortplanting: basis, en verband met partiële afgeleiden

Fitten van functies (met daaronder natuurlijk stiekem de kleinste kwadratenmethode)

Programma: SciDAVis



SciDAVis?

- Scientific Data Analysis and Visualisation

Geen Excel?

Met Excel:

Mogelijkheid voor een machtsfunctie als fitfunctie, wordt niet per se een wortelfunctie

Dat is te ondervangen: herschalen: T uitzetten tegen \sqrt{l}

Nadeel: weer terugrekenen

Voordeel van SciDAVis: je vindt direct de waarde van g met foutmarge

Mogelijkheden van SciDAVis

- Gedempte trilling

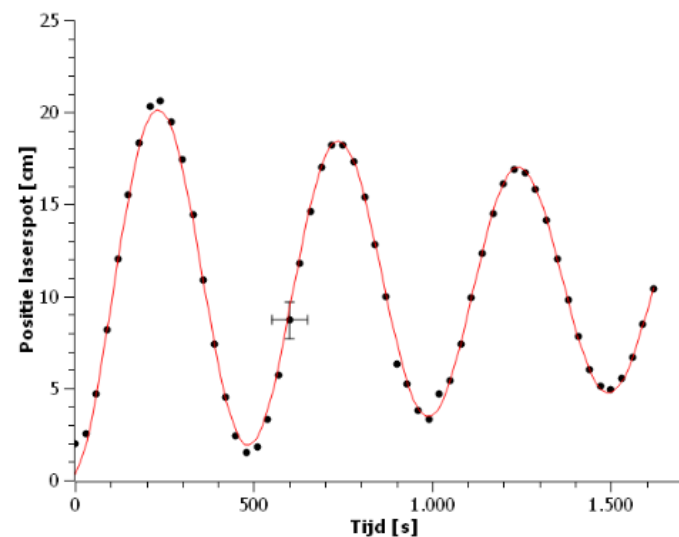
Meetwaarden en gefitte functie:

Volgens de theorie moet de formule er zo uitzien:

$$x(t) = x_0 + A \cdot e^{-\gamma t} \cdot \cos(\omega t + \varphi)$$

We zoeken de parameters x_0 , A , γ , ω en φ

Die kunnen we allemaal vinden met SciDAVis!



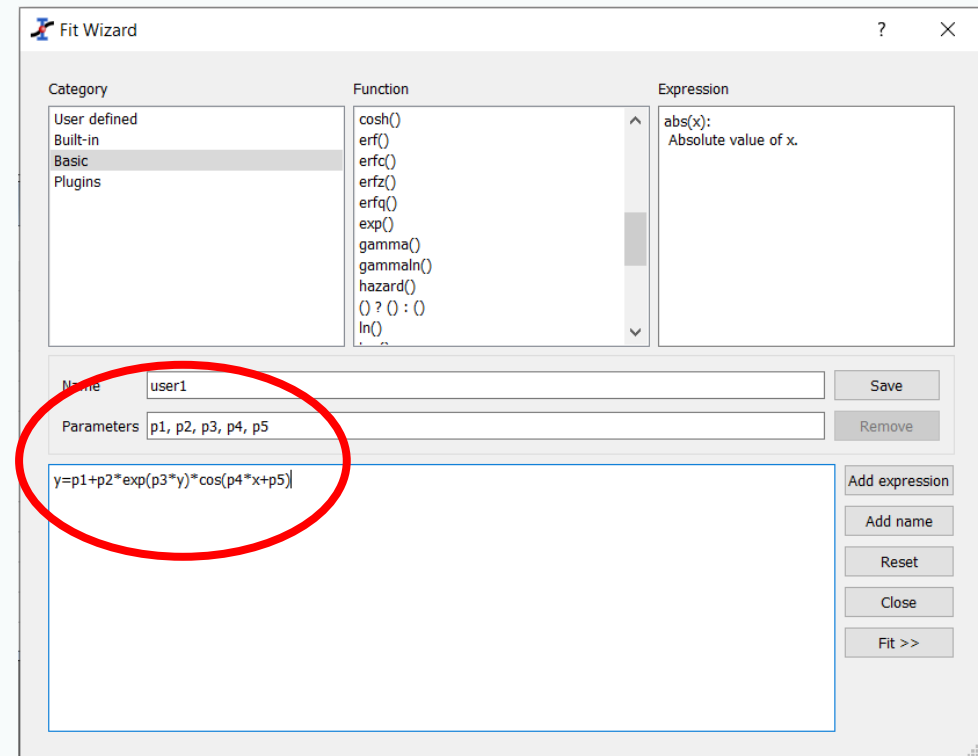
Figuur 3.7: De reflectie van een laser beschrijft een gedempte harmonische trilling. De metingen worden weergegeven door de zwarte punten. De rode lijn geeft de fit weer. Bij één meetpunt zijn foutbalkjes weergegeven, die voor alle punten even groot zijn. De fit past dus binnen één standaarddeviatie door alle meetpunten.

Hoe ziet dat er uit?

- Dit voer je in (naast de meetgegevens):
- Uitvoer na fitten:

```
Results Log
-----
[12-2-2014 14:07:11      Plot: "Graph3"]
Non-linear fit of dataset: Table1_2, using function: p1+p2*exp(p3*x)*cos(p4*x+p5)
Y standard errors: Unknown
Scaled Levenberg-Marquardt algorithm with tolerance = 0,0001
From x = 0 to x = 1.620
p1 = 10,5626940070008 +/- 0,0557069207047163
p2 = -10,523925760058 +/- 0,178672415660577
p3 = -0,000393884027059323 +/- 2,18906693983857e-05
p4 = 0,0124606417723876 +/- 2,16622010770939e-05
p5 = 0,210636316987065 +/- 0,017176732197885
-----
Chi^2/dof = 0,164727480654383
R^2 = 0,995215375716113
-----
Iterations = 4
Status = success
-----
```

- (voorbeeld uit de reader)
- Het programma is dus tamelijk krachtig



Bij de slingerproef is het eenvoudiger

- We zoeken de valversnelling g (op aarde)

Volgens de theorie geldt het volgende verband: $T = 2\pi \sqrt{\frac{l}{g}}$

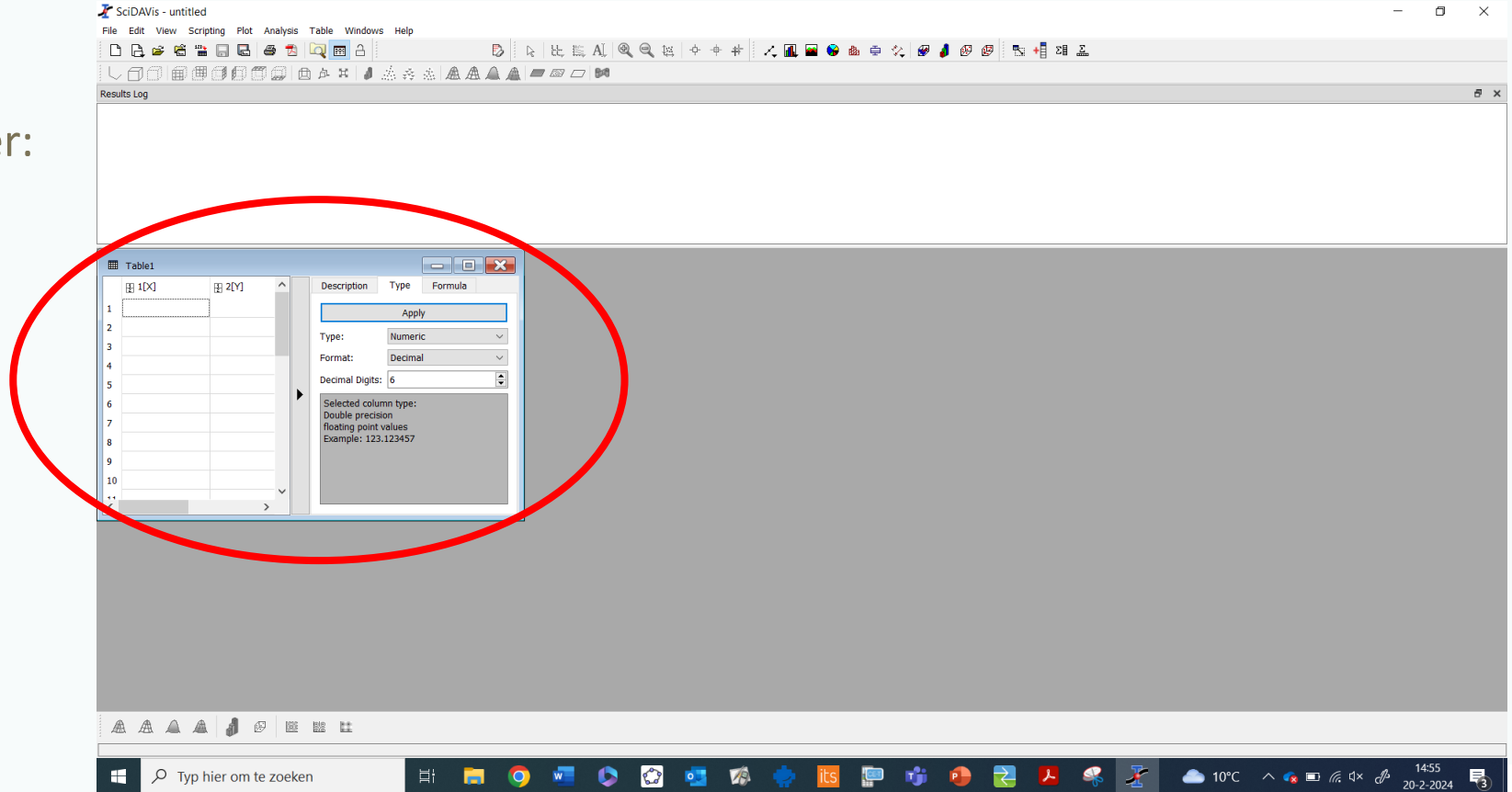
Leerlingen meten bij verschillende waarden van l de slingertijd T

Daarbij horen meetonzekerheden, die willen we meewegen



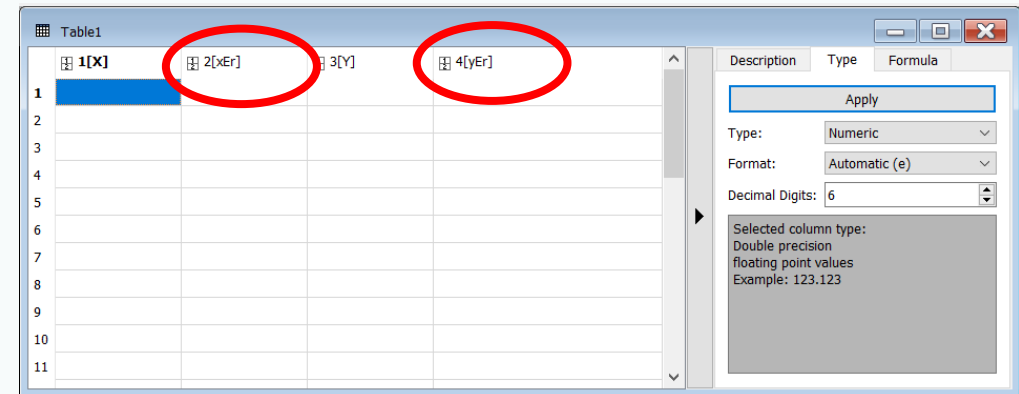
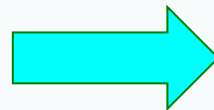
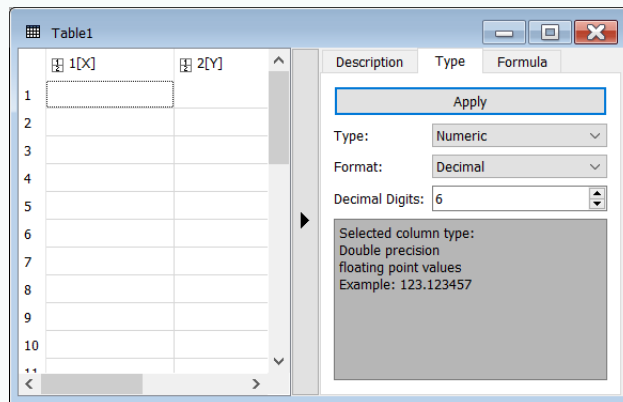
Werken met SciDAVis

- Startscherm:
We beginnen hier:



Meetgegevens en –onzekerheden invoeren

- Door ‘fout’ te klikken kun je kolommen toevoegen en kolommen instellen als kolommen met meetonzekerheden



Nu kunnen de gegevens worden ingevoerd: gebruik een komma, geen decimale punt

Naar een scatterplot, en verder

- Na het invoeren

- Selecteer alle kolommen

- 'Fout' klikken

- Kies 'Plot', en dan 'Scatter'

- Als het goed is staan de punten en foutmarges in de scatterplot

- Je kunt de titel, de astitels en de schaal aanpassen door 'fout' te klikken

- Nu fitten

- Klik op de scatterplot (dit scherm is nu actief)

- 'Fout' klikken

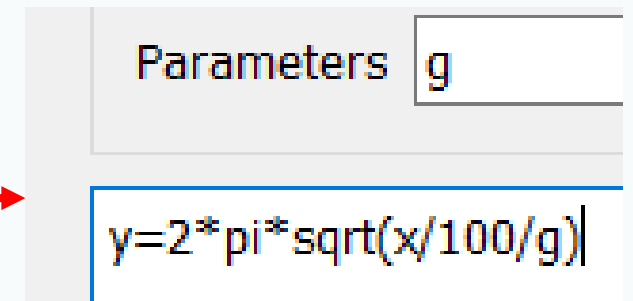
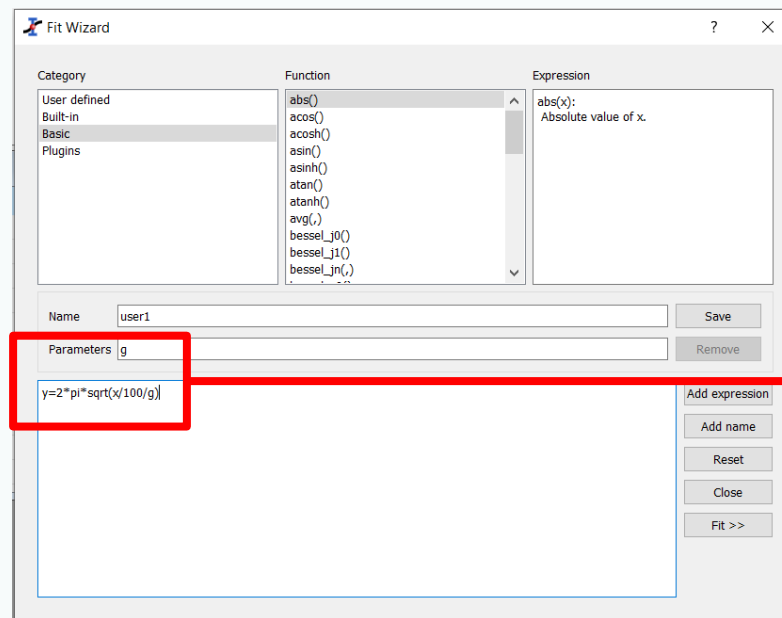
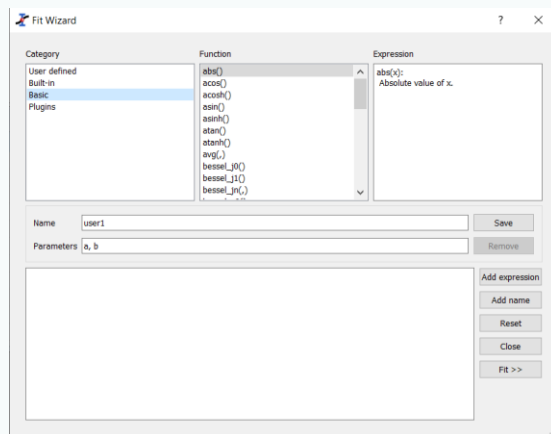
- Kies 'Analyze', en dan 'Fit Wizard'



De Fit Wizard

- Ziet er oorspronkelijk zo uit:

En wordt dan dit:



Resultaten

- Wat leveren de metingen op?

Ideeën?

- Twee natuurkundige contexten (rolgoot en slingerproef)

Zijn andere contexten mogelijk?

Ter vervanging van het rolgootexperiment: bij één hoogte alleen de horizontale valafstand gemeten

Ter vervanging van de slingerproef: bij verschillende lengtes de slingertijd gemeten

Zijn er bijvoorbeeld vanuit de chemie of biologie experimenten te bedenken die het rolgootexperiment of de slingerproef kunnen vervangen?



Belangstelling?

- **Belangstelling voor deze module?**

Kan met docenten middelbaar onderwijs gedraaid worden, geen expertise vanuit de universiteit nodig

De module kan gedeeld worden

Er is materiaal beschikbaar voor het stroomlijnen van de verslagen die voor de practica gemaakt moeten worden

Er zijn toetsen beschikbaar voor hoofdstuk 2: de *t*-toets

We willen naar certificering, maar dat is nog niet gedaan, dus: nog geen handleiding, nog geen heel nette formats en nakijkmodellen voor de verslagen



Vragen, opmerkingen?

- En als de vragen pas later komen: a.theune@paxchristicollege.nl