HP Prime Graphing Calculator

Tabellenboeken en de wet van Newcombe en Benford

Meer over de HP Prime te weten komen:
http://www.hp-prime.nl

Inleiding

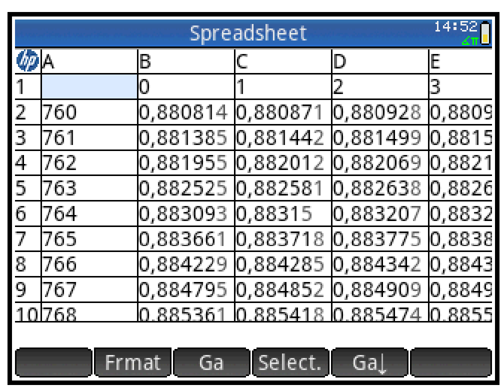
Frank Benford

De wet werd oorspronkelijk ontdekt door Simon Newcombe in 1881, jaren later in 1938 werd de wet 'herontdekt' door Benford. Benford onderzocht vele tabellen met (meet)gegevens waarna hij publiceerde in een wetenschappelijk tijdschrift. Hierdoor bleef vooral zijn naam verbonden aan de wet, die doorgaans als wet van Benford wordt aangeduid.

Waarschijnlijk ontdekten beide mannen deze wet doordat ze in tabellenboeken getallen opzochten en zagen dat de pagina's met getallen die begonnen met een 1 veel vaker werden bekeken dan pagina's met getallen die met een hoger cijfer zoals 8 of 9 begonnen. De eerstgenoemde pagina's waren veel meer 'beduimeld'.

Tabellenboek

Na de introductie van rekenmachines zijn tabellenboeken snel op de achtergrond verdwenen. Voordat er rekenmachines bestonden moesten wetenschappers die snellere berekeningen wilden maken gebruik maken van tabellenboeken. Met een tabellenboek kon men o.a. logaritmen, sinus- en cosinuswaarden opzoeken. Op de pagina's stonden getallen van 4 decimalen, daarvan stonden er drie in de eerste kolom, en het vierde cijfer van 0 tot 9 stond bovenaan de kolommen ernaast.



	A	B	C	D	E
1		0	1	2	3
2	760	0,880814	0,880871	0,880928	0,8809
3	761	0,881385	0,881442	0,881499	0,8815
4	762	0,881955	0,882012	0,882069	0,8821
5	763	0,882525	0,882581	0,882638	0,8826
6	764	0,883093	0,88315	0,883207	0,8832
7	765	0,883661	0,883718	0,883775	0,8838
8	766	0,884229	0,884285	0,884342	0,8843
9	767	0,884795	0,884852	0,884909	0,8849
10	768	0,885361	0,885418	0,885474	0,8855

Hiernaast een voorbeeld van hoe zo'n tabel eruitzag. In dit geval worden de logaritmen weergegeven van de getallen 7,600 tot rechtsonder 7,683. De eerst drie decimalen staan in de kolom onder A, de vierde decimaal staat bovenaan de kolom, hier in beeld ziet u alleen de waarden van 0 tot en met 3, in het boekje liep dit natuurlijk door tot 9. Zo stonden op de hele pagina de logaritmen van de getallen 7,600 tot 7,799, de volgende pagina begon dan weer bij 7,800 enzovoort. De laatste pagina eindigde bij 9,999. De tabel begint op de eerste pagina bij 0,000 en met dit boekje kon men dus de logaritmen opzoeken van de getallen 0,000 tot aan 9,999. Nu denkt u misschien dat dit niet handig is, maar hiermee kan men de logaritmen opzoeken van bijvoorbeeld 0,00000001 tot en met $9{,}999 \cdot 10^{100}$. Laten we eens kijken hoe dat kan.

Wat zijn logaritmen?

Stel dat ik vraag: Hoeveel moet ik bij 13 optellen zodat er 25 uit komt? Iedereen met vijf handen kan snel bedenken dat het antwoord 12 is. En met hoeveel moet ik 7 vermenigvuldigen zodat er 91 uit komt? Even tellen:

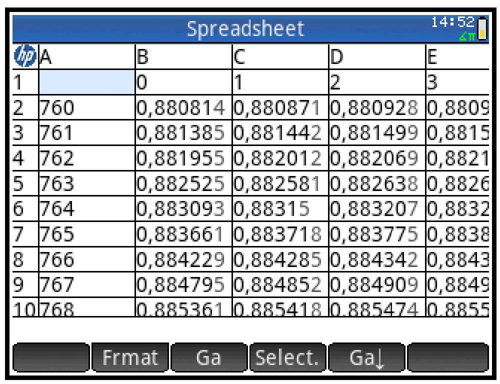
- 10 keer 7 is 70
- Als ik 70 van 91 afhaal houd ik er 21 over
- 21 is 3 keer 7
- Dus moet ik 7 vermenigvuldigen met 13 zodat er 91 uitkomt.

Zo kan ik ook vragen: tot welke macht moet ik 5 verheffen zodat er 625 uit komt?

- 5 keer 5 is 25
- 5 keer 5 keer 5 is: 5 keer 25 is 125
- 5 keer 5 keer 5 keer 5 is: 5 keer 125 is 625

Dus 5 tot de macht 4 (normaal schrijven we: 5^4) is 625. Dit kan (eventueel zelfs uit het hoofd) worden uitgerekend, maar tot welke macht moet ik 10 verheffen zodat er 7,642 uitkomt? Tja, dat rekent u niet een twee drie uit. En daar werden de tabellenboeken voor gebruikt.

Kijk in de tabel; de eerste drie getallen zijn 7, 6 en 4, die vindt u in de 6e rij, het laatste getal is een 2 dus kijk in de D kolom en u ziet het antwoord: 0,883207. Pak uw rekenmachine en bereken $10^{0,883207}$. Mijn HP Prime geeft als antwoord 7,64199.



	A	B	C	D	E
1		0	1	2	3
2	760	0,880814	0,880871	0,880928	0,8809
3	761	0,881385	0,881442	0,881499	0,8815
4	762	0,881955	0,882012	0,882069	0,8821
5	763	0,882525	0,882581	0,882638	0,8826
6	764	0,883093	0,88315	0,883207	0,8832
7	765	0,883661	0,883718	0,883775	0,8838
8	766	0,884229	0,884285	0,884342	0,8843
9	767	0,884795	0,884852	0,884909	0,8849
10	768	0,885361	0,885418	0,885474	0,8855

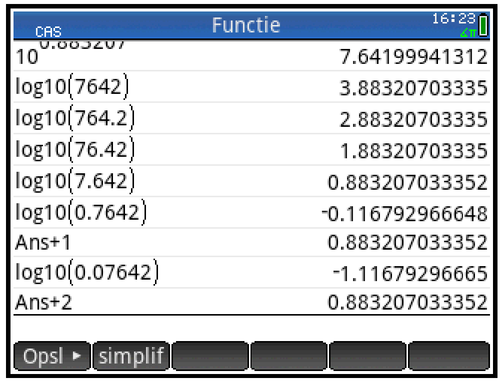
Bereken met uw rekenmachine de volgende logaritmen:

Log(7642)
Log(764,2)
Log(76,42)
Log(7,642)
Log(0,7642). Tel er nu 1 bij op.
Log(0,07642). Tel er nu 2 bij op.

Hoe komt het dat die getallen zo op elkaar lijken?

7642 =
10 * 764,2 =
10 * 10 * 76,42 =
10 * 10 * 10 * 7,642 =
10 * 10 * 10 * 10 * 0,7642 =
10 * 10 * 10 * 10 * 10 * 0,07642

Als u de logaritme van een getal bepaald, wordt de waarde voor de komma alleen bepaald door de grootte orde van het getal; zit het tussen 0 en 10, tussen 10 en 100 of tussen 100 en 1000 en zo verder en zo ook verder omlaag; zit het tussen 1/10 en 1, zit het tussen 1/100 en 1/10. Dit zijn allemaal machten van 10 en de Log knop op de rekenmachine is ook eigenlijk de 10 logaritme; schrijf $\text{Log}_{10}(7642)$. Omdat we alle dagen met het decimale (tientallige) getallenstelsel werken mag de 10 weggelaten worden. Alleen als een ander getal als grondtal wordt gebruikt, bijvoorbeeld 2, schrijft u Log_2. Zo is $\text{Log}_2(16)=4$ want $2*2*2*2 = 2^4=16$



Een tussenstap

Bedenk dat $10^5 * 10^7$ gelijk is aan 10^{12}
Want:10*10*10*10*10 (5 factoren 10 die met elkaar vermenigvuldigd worden)
* (vermenigvuldigt u met)
10*10*10*10*10*10*10 (7 factoren 10 die met elkaar vermenigvuldigd worden)
=
10*10*10*10*10*10*10*10*10*10*10*10
$=10^{12}$ (kort en goed 12 factoren 10)

De machten van eenzelfde grondtal (in dit geval 10) mogen bij vermenigvuldiging worden opgeteld. In het voorbeeld telt u de macht 5 bij de macht 7 op aangezien de twee getallen die u moet vermenigvuldigen hetzelfde grondtal hebben.

De regel die hierbij hoort: $u^{t+r} = u^t \times u^r$

De definitie van een logaritme?

$q = \log_a(x) \leftrightarrow a^q=x$. In woorden: Als q gelijk is aan de a logaritme van x, dan is a tot de macht q gelijk aan x.

Een voorbeeld:
$3 = \log_2(8) \leftrightarrow 2^3=8$. In woorden: 3 is de 2 logaritme van 8, dus is 2 tot de macht 3 gelijk aan 8.

De Log toets van de rekenmachine gebruikt de zogenaamde 10 log, daarnaast is er nog de LN (e^x) toets, ofwel de Natuurlijke Logaritme, die gebruikt als grondtal het getal e = 2,71828182846. Dit getal is een beetje te vergelijken met π, het heeft bijzondere eigenschappen en heeft, net als π, een oneindige decimale ontwikkeling.

Nu terug naar de 10 logaritme, hier een paar voorbeelden:

$\log_{10}(1/10)$	$=-1$	want $10^{-1}=1/10$
$\log_{10}(1)$	$= 0$	want $10^0=1$
$\log_{10}(10)$	$= 1$	want $10^1=10$
$\log_{10}(100)$	$= 2$	want $10^2=100$
$\log_{10}(1000)$	$= 3$	want $10^3=1000$

Laten we in algemene zin eens naar de logaritmen kijken.

- Stel $Log_g (a)+ Log_g (b)=p$
- Gebruik de regel dat als $k+l=m$ dan $g^{k+l}=g^m$
- Dan volgt: $g^{\log_g (a)+\log_g (b}=g^p$
- Gebruik de regel $u^{t+r}=u^t\times u^r$ (zie de tussenstap)
- Dan volgt:$g^{\log_g (a))}\times g^{\log_g (b)}=g^p$
- Per definitie is $g^{Log_g (x)}=x$ (verhef g tot de macht tot welke ik g moet verheffen zodat er x uitkomt)
- Dus: a x b = g^p dan volgens de definitie
- $log_g (a*b)=p$

En daaruit volgt een belangrijke en heel handige rekenregel voor logaritmen:

$Log_g\,(a*b)=Log_g\,(a)+Log_g\,(b)$

- Dus: $\log_{10}(10*100)=\log_{10}(1000)$ mag worden veranderd in:
- $\log_{10}(10)+\log_{10}(100)=\log_{10}(1000)$ oftewel $1 + 2 = 3$
- $\log_{10}(8551)=\log_{10}(10*855{,}1)=\log_{10}(100*85{,}51)$
- Dit kunt u herschrijven als: $\log_{10}(8551)=\log_{10}(10)+\log_{10}(855{,}1)=\log_{10}(100)+\log_{10}(85{,}51)$
- $\log_{10}(8551)=1+\log_{10}(855{,}1)=2+\log_{10}(85{,}51)$

Waarom was een tabellenboek zo handig?

Kijk eens naar de volgende opgave:

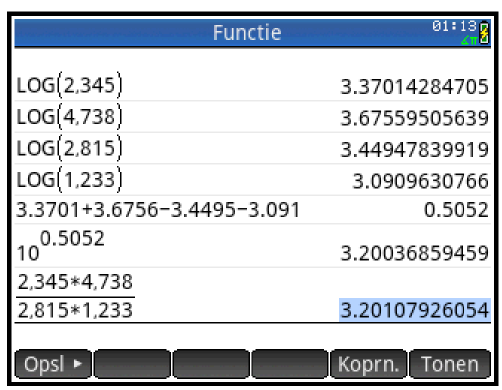
$$\frac{2345*4738}{2815*1233}=?$$

Denk even aan de tijd dat er geen rekenmachines bestonden. Er waren best veel mensen die dit soort sommen moesten oplossen; boekhouders, architecten, bouwkundigen, belastingambtenaren, etc.
Hoe deden ze dit zonder rekenmachine; met een tabellenboekje!
Want deze opgave is relatief snel te benaderen via de volgende operatie:
$\log_{10}(2345)+\log_{10}(4738)-\log_{10}(2815)-\log_{10}(1233)=\log(\text{antwoord})$

In het tabellenboek zijn al deze waardes op te zoeken:
3,3701 + 3,6756 – 3,4495 – 3,0910 = Log(antwoord)
Dus 0,5052 =Log (antwoord)

Nu wordt het tabellenboekje omgekeerd gebruikt, u zoekt bij de uitkomsten naar 5052, u ziet dat er een nul voor staat, dus het getal ligt tussen 1 en 10. Dan kunt u in de tabel de decimalen vinden: 3200, dus de uitkomst is 3,200.

U kunt het met de HP-Prime ook berekenen via $10^{0{,}5052} = 3{,}200$. En u kunt de originele opgave invoeren in de rekenmachine, dan ziet u een klein verschil, maar kijk eens naar dat verschil en bedenk hoe snel u dit met alleen pen, papier en tabellenboekje had kunnen doen. Die tijd moet natuurlijk niet vergelijken worden met het gebruik van de rekenmachine, maar met het uitschrijven van de vermenigvuldigingen en delingen die oorspronkelijk werden gevraagd. Ten opzichte van die tijd was het gebruik van de logaritmetabellen een heel belangrijke doorbraak.



De ontdekking van de wet van Newcombe en Benford

Net als iedere wis- en natuurkundige in die tijden gebruikten Newcombe en Benford dus tabellenboeken en zoals het goede wetenschappers beaamd, merkten zij op dat de pagina's waarop de getallen die begonnen met het getal 1 veel meer bekeken werden door hen zelf en andere gebruikers van de tabellenboeken dan de pagina's die begonnen met de hogere cijfers zoals 8 en 9. Dit konden ze zien doordat sommige pagina's vaker besmeurd en iets meer versleten waren, terwijl andere pagina's weinig slijtage vertoonden.
Benford ging er werk van maken en onderzocht allerlei databestanden met numerieke gegevens: lengte van rivieren, oppervlakte van meren en landerijen, hoogte van bergen, wiskundige tafels, getallen uit kranten en tijdschriften etc. Hiermee vond Benford een steeds betere onderbouwing van zijn vermoeden, het wiskundige bewijs kwam pas veel later, dat is in 1995 gepubliceerd door Hill.

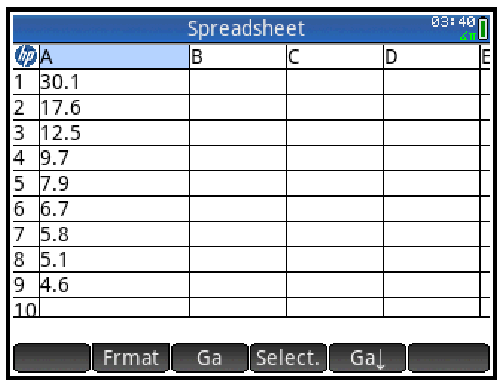
De wet stelt: de waarschijnlijkheid dat het eerste cijfer van een getal uit een lijst met gemeten gegevens (D_1) gelijk is aan d kan worden berekend met de volgende formule:

$$P(D_1 = d) = \log_{10}\left(1 + \frac{1}{d}\right), \, voor \;\, d = 1 \ldots 9$$

Dus $P(D_1 = 1) = log_{10}\left(1 + \frac{1}{1}\right) = log_{10}(2) = 0{,}301$

En $P(D_1 = 2) = log_{10}\left(1 + \frac{1}{2}\right) = log_{10}(1{,}5) = 0{,}176$

$P(D_1 = 3) = log_{10}\left(1 + \frac{1}{3}\right) = log_{10}(1{,}3333) = 0{,}125$



	A	B	C	D	E
1	30.1				
2	17.6				
3	12.5				
4	9.7				
5	7.9				
6	6.7				
7	5.8				
8	5.1				
9	4.6				
10					

Hiernaast een tabel met berekening van de percentages kans op eerste cijfer 1 t/m 9.

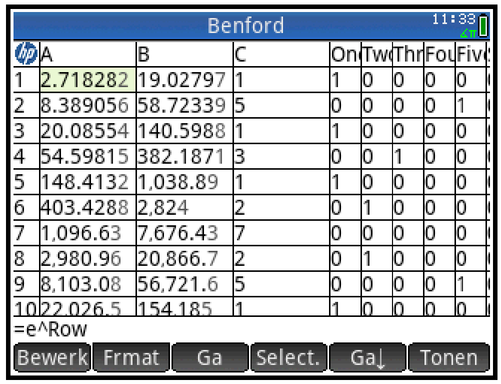
Een belangrijk aspect van het verhaal is dat de gebruikte eenheid niets uitmaakt. Dus of je nu als Europeaan met eenheid de meter de lengte van rivieren meet of je bent een Amerikaan en je gaat diezelfde rivieren meten met yards, inches of mijlen, je krijgt allebei een tabel met heleboel getallen die aan de regel voldoen!

De tabel geeft bij benadering aan hoe vaak een bepaald cijfer als eerste cijfer voor zal komen in een verzameling (bestaande uit genoeg getallen) die 'toevallig' tot stand is gekomen.

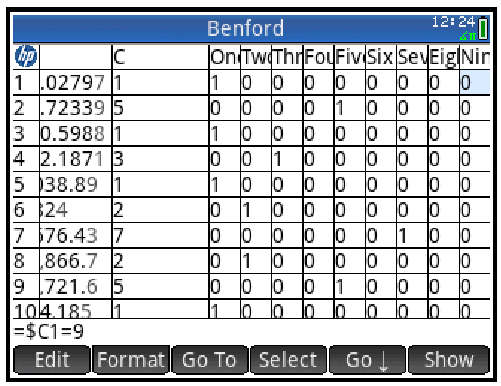
Op de website http://testingbenfordslaw.com staan zeer verschillende datasets waarbij wordt bekeken of ze voldoen aan de wet.
U zult verbaasd zijn hoe gevarieerd de datasets zijn die allemaal voldoen aan de wet. Het feit dat de eenheid waarmee wordt gemeten geen verschil maakt geeft eigenlijk aan dat als je een tabel met data hebt, je de data met elk willekeurig getal kan vermenigvuldigen en dat de nieuwe tabel ook deze verdeling van eerste cijfers zal moeten hebben. DAT KLOPT! En dat is best apart.

Mocht u ooit denken dat u de belastinginspecteur met valse omzetcijfers om de tuin kan leiden, bedenk dat een slimme belastinginspecteur snel kan testen of uw gegevens ‘echt’ zijn of verzonnen!

Laten we eens een lange lijst met getallen maken, bijvoorbeeld met e^{rij}. We gebruiken de spreadsheet applicatie en zetten in de eerste kolom het getal e tot de macht van de rij, dat begint dus bij $e^1=e$ (≈ 2,71828182846.....) en eindigt bij e^{250}≈ 3,75*10108, een onwaarschijnlijk groot getal.

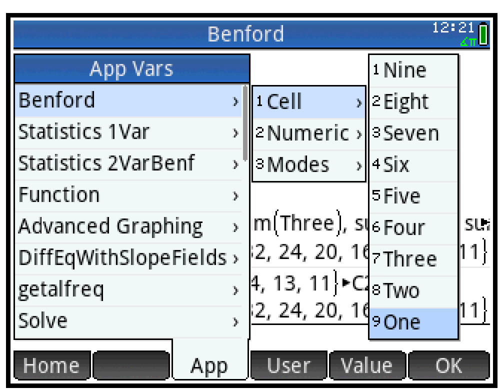


Laadt de Benford applicatie (beschikbaar op www.hp-prime.nl) op uw pc, bijvoorbeeld op het bureaublad. Open nu de HP aansluitkit software en koppel de HP Prime aan de PC via de USB poort. U ziet in het scherm van de aansluitkit de naam van de rekenmachine in het linker gedeelte van het scherm staan. Sleep nu van het bureaublad de Benford app naar de naam van de rekenmachine en laat los. De app wordt geladen op de HP Prime. Open de app, u ziet als de cursor op een cel in de A kolom staat onderin beeld de formule =e^Row.
Oftewel de cel wordt berekend als e tot de macht rij (van de cel), de rij van de cel loopt van 1 tot en met 250.
In cel A251 staat een ‘willekeurig’ getal (in dit geval 7) waarmee de waarde in kolom A wordt vermenigvuldigd en zo ontstaat kolom B, later gaan we die waarde veranderen en zullen we zien dat dit nauwelijks effect heeft op de uitkomst van ons onderzoek.
In de C kolom staat steeds het eerste cijfer van het getal uit kolom B.
In de kolommen daarnaast wordt gecontroleerd welk cijfer (van 1 tot en met 9) in kolom C staat, zo is er de kolom One, daarin staat een 1 als in kolom C een 1 staat en daarin staat een 0 als er geen 1 in kolom C staat. Dat loopt door voor de tweeën, de drieën tot aan de negens.



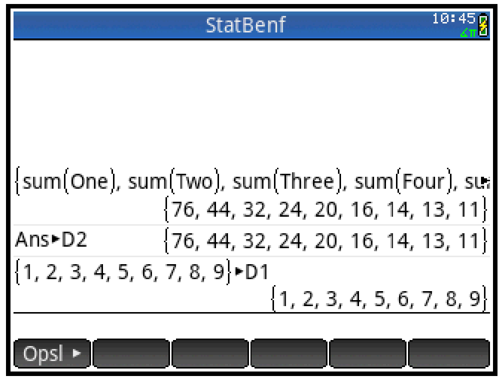
Die kolommen met al die cijfers gaan we in een grafiek bekijken; eerst zetten we de totalen van de kolommen met de enen, tweeën etc. in een lijst. Dat kan op verschillende manieren: Via de emulator (met het toetsenbord van de PC) of met de rekenmachine zelf.

Gebruikt u de emulator, dan de Benford app openen, ga vervolgens naar het Home scherm met [Settings] en typ de opdracht in: (Let wel op de hoofd- en kleine letters!) {sum(One),sum(Two),......,sum(Nine)}
Bevestig met [Enter].

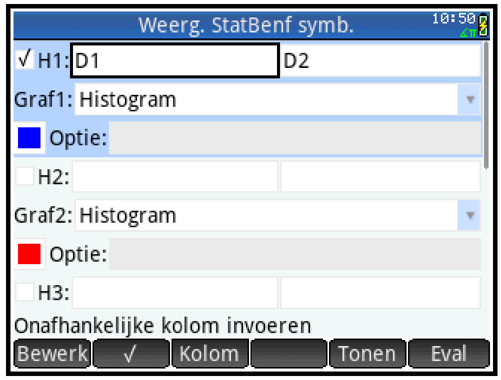


Gebruikt u de rekenmachine dan ook eerst de Benford app openen, vervolgens het Home scherm openen met [Settings] en met [Shift] en [5] plaatst u de accolades, dan via de gereedschapskist [Mem] en de catalogus haalt u het sum commando op. Kies via [Vars] en [Toep.] de variabele One (zie het scherm hiernaast) dan een komma, opnieuw sum en de variabele Two tot aan Nine. Bevestig met [Enter] en nu gebeurt iets dat u waarschijnlijk nog niet kent van de Prime: u moet even wachten! In de rechterbovenhoek van het scherm ziet u een zandloper terwijl de rekenmachine bezig is om in alle 9 de lijsten iedere keer 250 elementen uit die lijsten op te tellen.

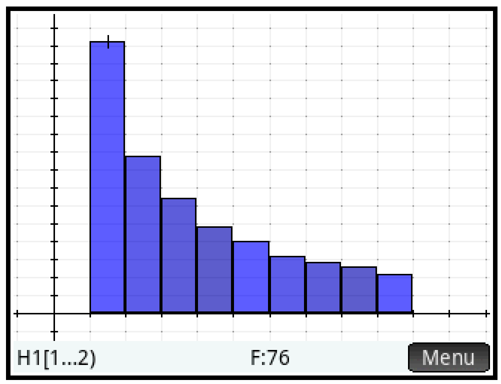
Open nu de 2Var.Statistieken app, maak deze leeg met Shift en Esc Clear.



Open het Home scherm met Settings en kies Opsl ▸ en typ ‘D2’. Het laatste antwoord in het Home scherm (de optelling van de kolommen One tot en met Nine) staat nu in lijst C2 van de 2Var.statistieken app.
Typ vervolgens: ‘{1,2,3,4,5,6,7,8,9} Opsl ▸ D1’.

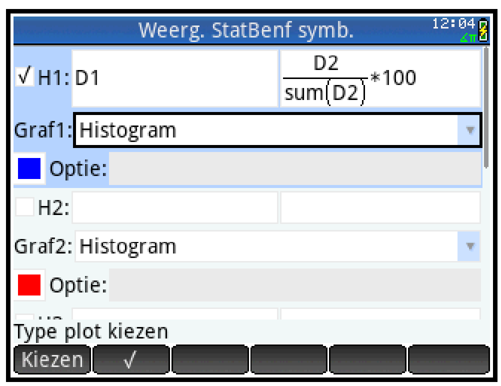


Via het symbolische scherm van de 2Var.Statistieken app stellen we de eigenschappen van de grafiek in. Kies Symb Setup en vul de velden in zoals in het scherm hiernaast. Dus kolom D1 voor de onafhankelijke variabele op de x-as, D2 voor de afhankelijke variabele op de y-as en als type het Histogram.



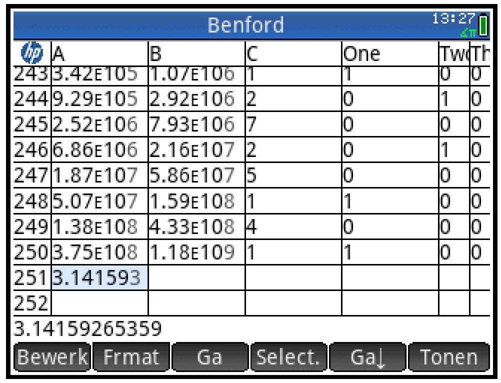
Via View Copy en optie 2 Automatisch schalen opent u de grafiek.

We kunnen deze absolute waarden ok omzetten in een percentage. Open de symbolische instellingen en pas de definitie van de plot aan (zie het scherm hiernaast). Zet de cursor op het vak waar D2 stond en kies Bewerk, typ het commando in, u kunt eventueel ook via de catalogus het commando sum invoeren.



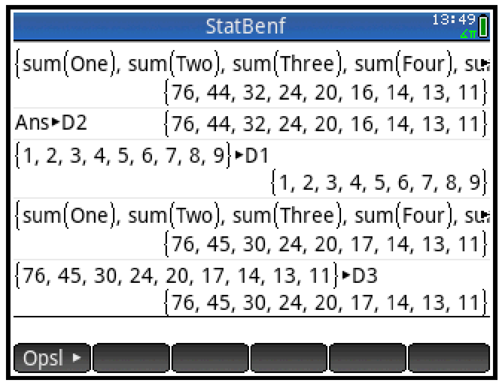
Kies opnieuw View Copy en optie 2 Automatisch schalen voor de grafiek. Nu heeft u de percentages in beeld.

Tot slot de clou van Benford's wet, de data kan met een willekeurig getal worden vermenigvuldigd waardoor de wet niet beinvloed wordt.
Open de Benford app en kies onderin de beeldrand de optie Ga, vul dan in A251 en bevestig met Enter. U had ook met cursorbesturing kunnen wandelen, maar dat duurt nogal lang.

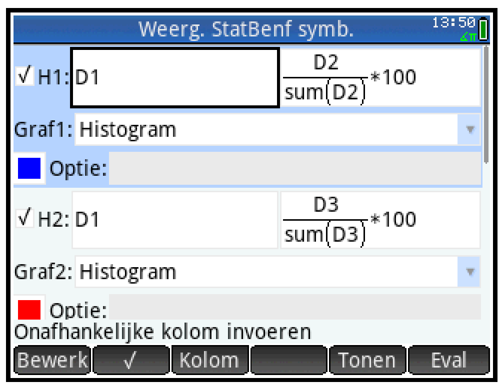


Verander de waarde 7 in een ander getal naar keuze, na de bevestiging hiervan wordt de hele spreadsheet opnieuw doorgerekend; in kolom B staan nu de producten van de getallen in de A kolom en het door u gekozen getal, dit verandert de waarden in de kolommen C, One, Two, etc tot en met Nine.

Open het Homescherm met Settings en dubbelklik op de regel waarin de lijst met aantallen eenen, tweeën, ect. berekend wordt. Hiermee kopieert u deze eerdere invoer naar de opdrachtregel, bevestig met Enter.
Open nu eerst de 2Var.Statistieken app, en sla dan in het Homescherm de zojuist berekende lijst op als lijst D3.



Pas het symbolische scherm aan (zie het scherm hiernaast).



Kies View Copy en optie 2 Automatisch schalen voor de grafieken. Nu heeft u de percentages in beeld van beide datasets. Wandel met cursor langs de kolommen, gebruik ▲ of ▼ om van het ene naar het andere histogram te stappen bij de verschillende waarden langs de x-as.