

Von der Schulmathematik zur Benfordverteilung

(Taylor's Version)

Aufbau der Präsentation

Wiederholung: Zufall und Wahrscheinlichkeitsverteilungen

Was ist Benfords Gesetz?

Wiederholung: Zahlensystem und Logarithmen

Exponentielle (wissenschaftliche) Schreibweise

Die Benfordverteilung und Logarithmen

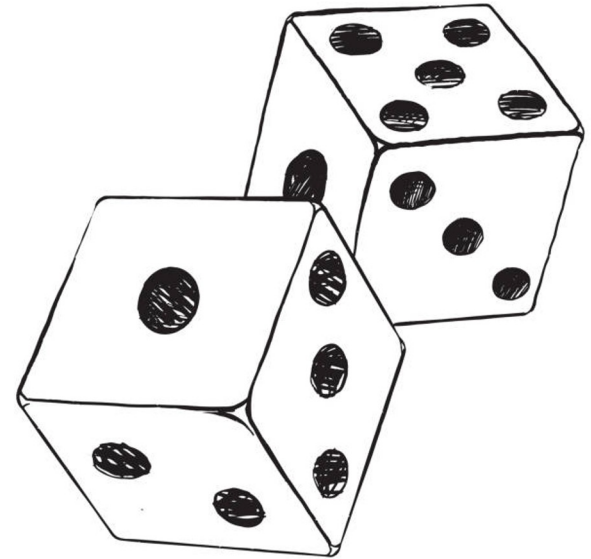
Skalenunabhängigkeit und Logarithmen

Schlussbemerkungen: mathematisches Denken

Wiederholung: Zufall

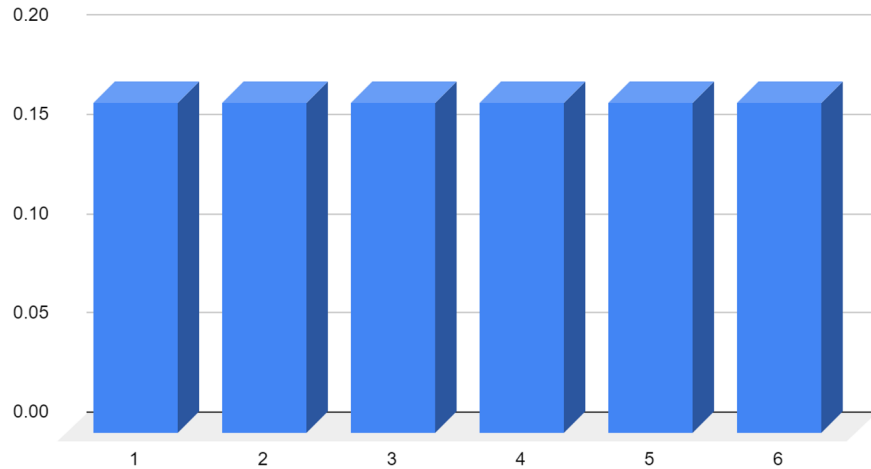
Intuitives Beispiel von Zufall: Würfel

Gezinkter Würfel: Zufall?

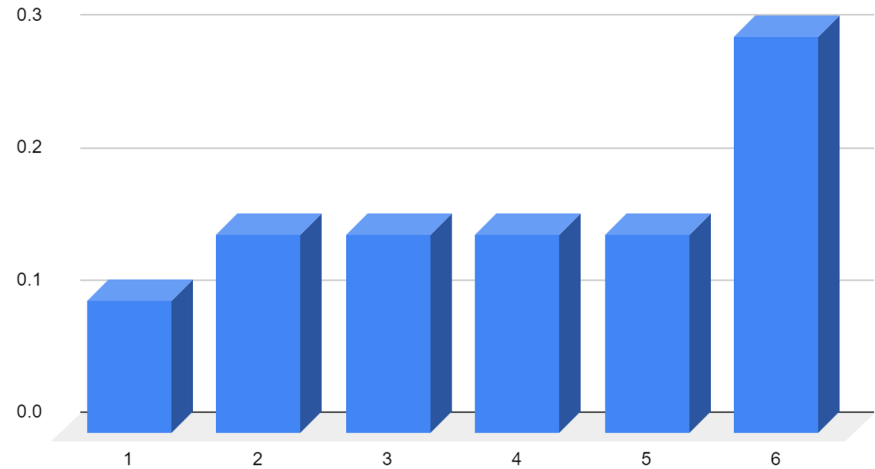


Wiederholung: Wahrscheinlichkeitsverteilung

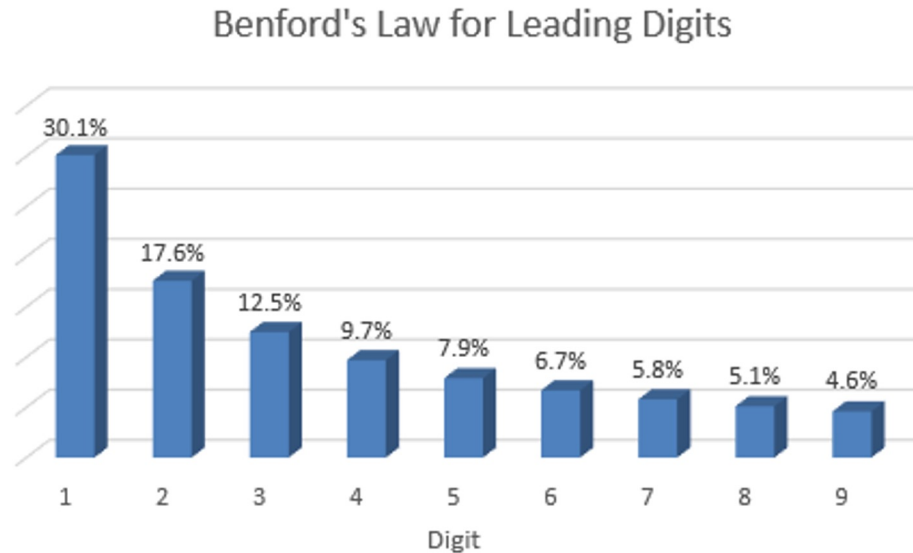
Wahrscheinlichkeitsverteilung eines fairen Würfels



Wahrscheinlichkeitsverteilung eines gezinkten Würfels



Was ist Benfords Gesetz?



<https://statisticsbyjim.com/probability/benfords-law/>

Anwendbarkeit von Benfords Gesetz

Gilt Benfords Gesetz nur für natürliche Datensätze? Oder auch für beliebige menschliche Datensätze?

Beispiel: die Anzahl der Streams der Top 500 Songs von Taylor Swift auf Spotify?

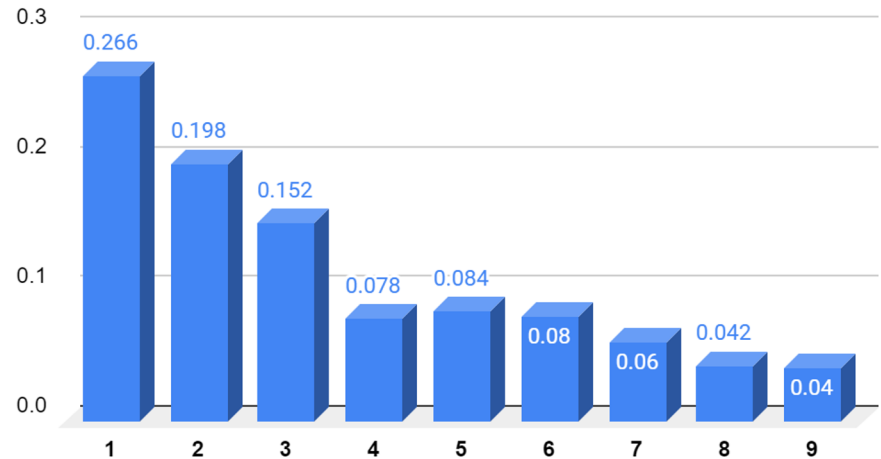
Ja? Nein? Warum?

Probier's mal selber!



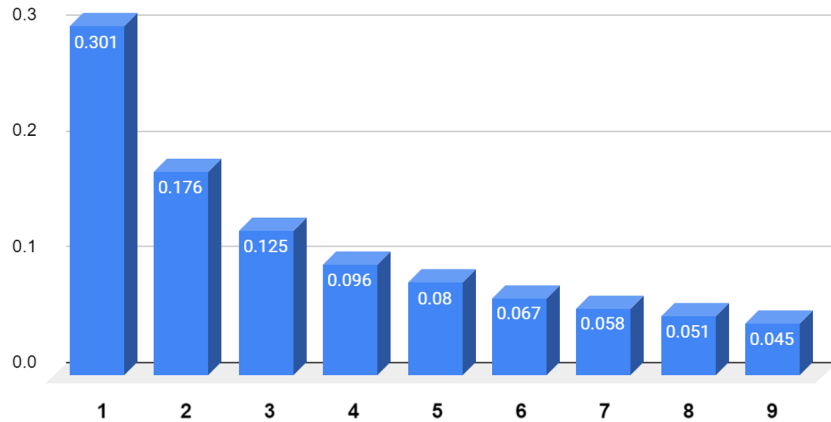
1. Scanne den QR-Code
2. Betrachte die 2. Spalte
3. Überprüfe, ob sich beim Runterscrollen die Anzahl der Einträge je nach führende Ziffer ändert

Führende Ziffer der Top 500 Titel von Taylor Swift

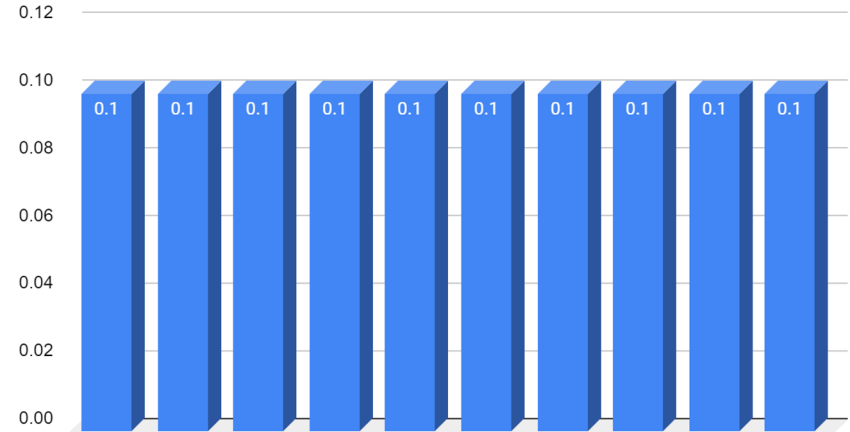


Gleichmäßigkeit im benfordschen Gesetz?

Benfords Gesetz

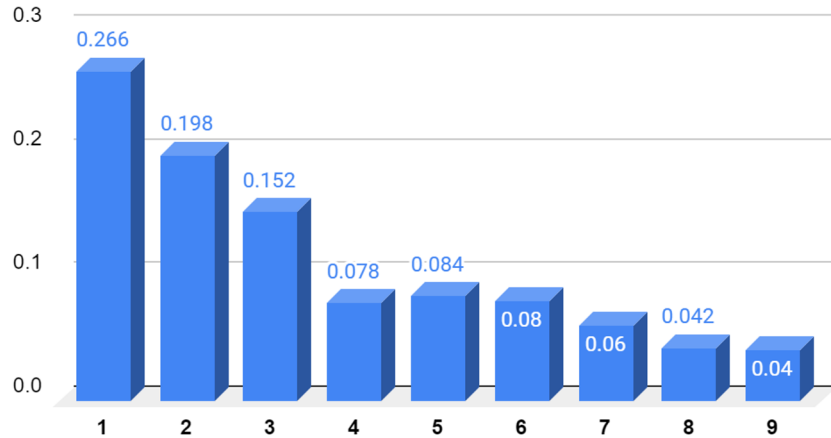


Benfords Gesetz (logarithmisch)

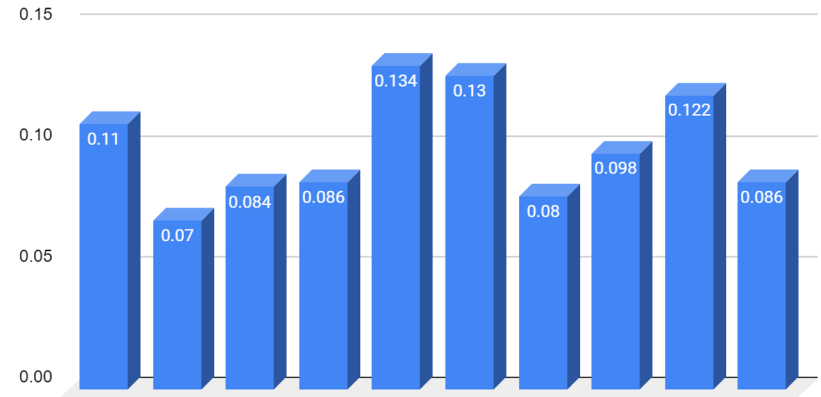


Gleichmäßigkeit (Taylor's version)

Führende Ziffer der Top 500 Titel von Taylor Swift



Führende Ziffer der Top 500 Titel von Taylor Swift
(logarithmisch)



Zahlen und Logarithmen

Dezimalsystem (Basis-10)

10 Ziffern (0, 1, 2, 3, 4, 5, 6, 7, 8, 9)

Stelle bestimmt den Wert (mal 10 multipliziert)

Basis-10 Logarithmus

$\log_{10}(10) = 1$, $\log_{10}(100) = 2$, $\log_{10}(1000) = 3$, ...

Exponentielle (“wissenschaftliche”) Schreibweise

Vereinfachte Zahlendarstellung

“Anti-Hero” : 1.528.306.561 $\rightarrow 1,5 \cdot 10^9$

“Shake It Off” : 1.385.891.279 $\rightarrow 1,4 \cdot 10^9$

“Love Story” : 877.220.622 $\rightarrow 8,8 \cdot 10^8$

Exponentielle Schreibweise und Logarithmen

Wenn man den Logarithmus aus der exp. Schreibweise nimmt:

“Anti-Hero”	:	$\log_{10}(1,5 * 10^9)$	->	9,18
“Shake It Off”	:	$\log_{10}(1,4 * 10^9)$	->	9,14
“Love Story”	:	$\log_{10}(8,8 * 10^8)$	->	8,94

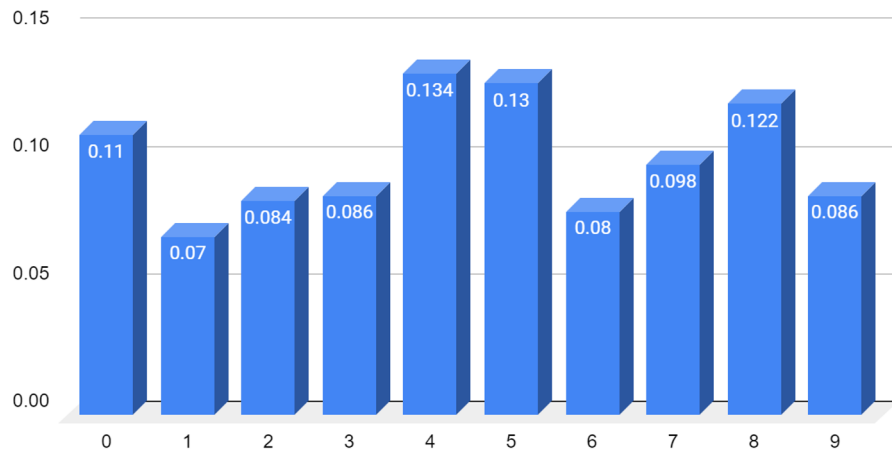
Logarithmen und Benfords Gesetz

“Anti-Hero” : 9,18

“Shake It Off” : 9,14

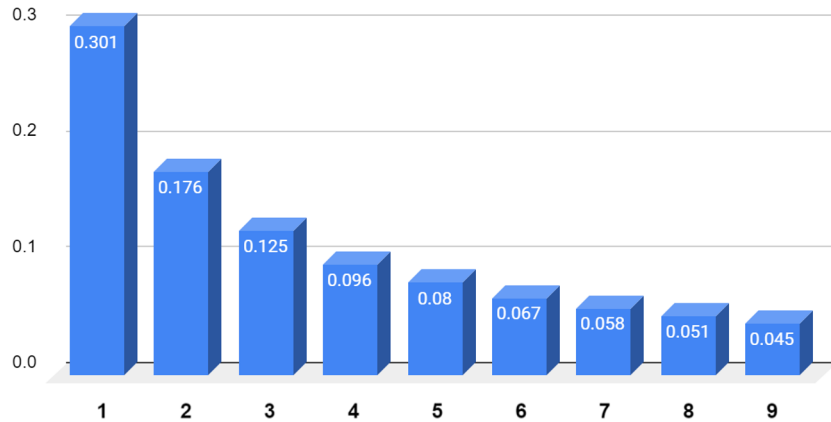
“Love Story” : 8,94

Führende Ziffer der Top 500 Titel von Taylor Swift
(logarithmisch)

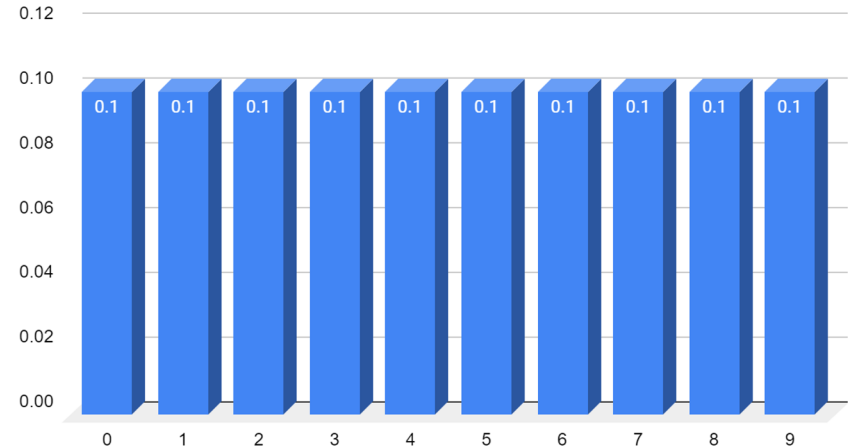


Gleichmäßigkeit im benfordschen Gesetz

Benfords Gesetz



Benfords Gesetz (logarithmisch)



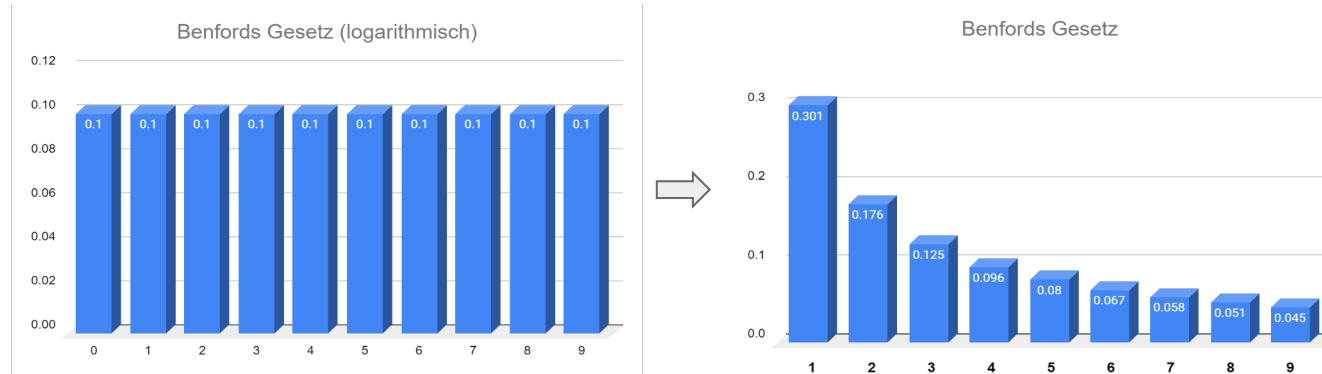
Benfordverteilung und Logarithmen

FRANK BENFORD

The frequency of first digits thus follows closely the logarithmic relation

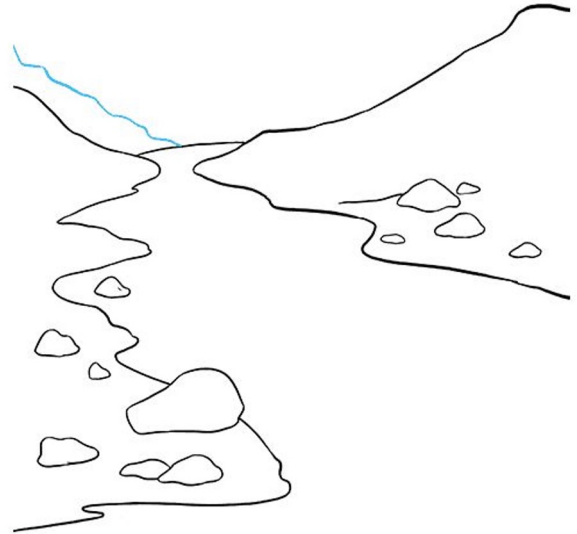
$$F_a = \log \left(\frac{a + 1}{a} \right), \quad (1)$$

where F_a is the frequency of the digit a in the first place of used numbers.



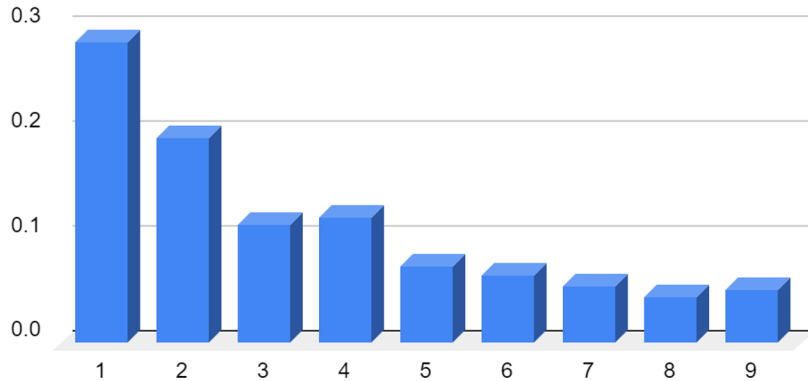
Skalenabhängigkeit ?

Macht es einen Unterschied für die Verteilung der führenden Ziffern, ob man z.B. Flüsse in Meter oder Yard misst?

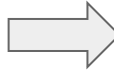


Skalenunabhängigkeit (Taylor's Version)

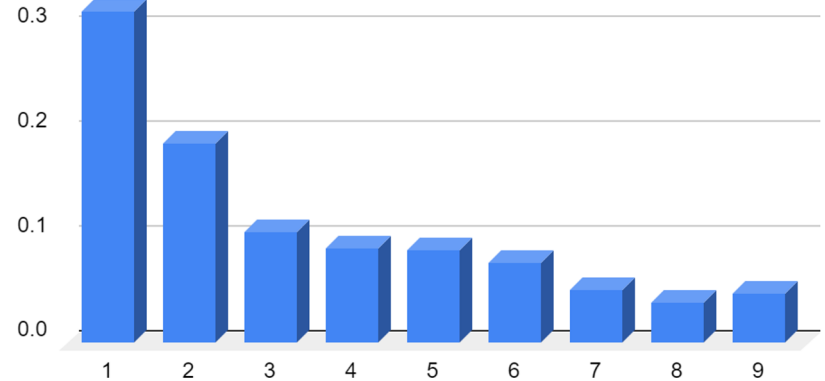
Führende Ziffer der Top 500 Titel von Taylor Swift
(gestreamte Minuten)



* 60



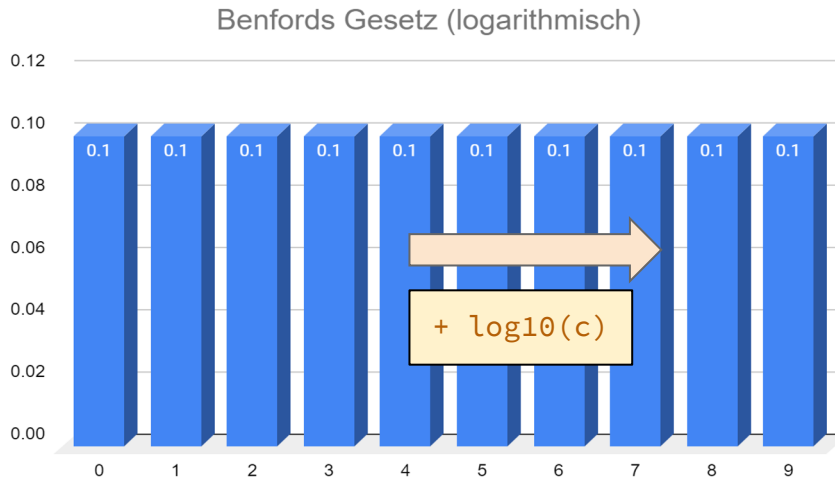
Führende Ziffer der Top 500 Titel von Taylor Swift
(gestreamte Sekunden)



Skalenunabhängigkeit und Logarithmen

Multipliziert man eine Zahl x mit einem Skalierungsfaktor c , erhält man:

$$\log_{10}(c \cdot x) = \log_{10}(x) + \log_{10}(c)$$



Horizontale Verschiebung
= Verteilung bleibt
unverändert!

Schlussbemerkungen: mathematisches Denken

Ausgangslage

Benfordsverteilung unintuitiv
(keine gleichmäßige
Verteilung)

Anwendbarkeit unklar
(nur Naturdatensätze?)

Logik hinter der
Benfordverteilung unsichtbar

Endlage

Benfordsverteilung intuitiver
(gleichmäßige Verteilung der
Logarithmen)

Anwendbarkeit verständlicher
(skalenunabhängige Datensätze)

Logik hinter der
Benfordverteilung klarer