

Ergebnisse der Analyse: harmonic-ana-3b

harmonic-ana-3b-zipf-regression-major.R

Dieses Programm liest alle annotierten Beethoven Sonaten ein (im .tsv Format) und erstellt ein Dataframe, welches jedem Akkord (chord) einen Rang (rank) und die relative Häufigkeit (relative_frequency) zuordnet. Dann werden die Akkorde in Dur (Major) ausgefiltert.

Weiter wird die Zipf Funktion wie folgt definiert: $f(r) = \frac{a}{(b+r)^c}$

Anschliessend wird mit der Funktion 'nls' aus dem Paket 'stats' eine Optimierung der Variablen a , b und c durchgeführt mit der Methode der kleinsten Quadrate. Dies ermöglicht eine Maximierung des Bestimmtheitsmass R^2 .

Es ergibt sich folgendes:

```
> summary(zipf_fit)
Formula: chord_freq_df$relative_frequency ~ zipf_function(chord_freq_df$rank, a, b, c)
Parameters:
  Estimate Std. Error t value Pr(>|t|)
a 0.52530   0.03474   15.12  <2e-16 ***
b 1.67828   0.09469   17.72  <2e-16 ***
c 1.33423   0.02286   58.38  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.0009972 on 837 degrees of freedom
Number of iterations to convergence: 7
Achieved convergence tolerance: 1.993e-06

[1] "R^2 = 0.980712560621019"
```

Zuletzt wurde ein Plot Rang vs. rel. Häufigkeit mit dem Paket 'ggplot2' erstellt, in welchem die Funktion über die Datenpunkte gelegt wurde. (Abbildung 1)

harmonic-ana-3b-zipf-regression-minor.R

Dieses Programm liest alle annotierten Beethoven Sonaten ein (im .tsv Format) und erstellt ein Dataframe, welches jedem Akkord (chord) einen Rang (rank) und die relative Häufigkeit (relative_frequency) zuordnet. Dann werden die Akkorde in Dur (Major) ausgefiltert.

Weiter wird die Zipf Funktion wie folgt definiert: $f(r) = \frac{a}{(b+r)^c}$

Anschliessend wird mit der Funktion 'nls' aus dem Paket 'stats' eine Optimierung der Variablen a , b und c durchgeführt mit der Methode der kleinsten Quadrate. Dies ermöglicht eine Maximierung des Bestimmtheitsmass R^2 .

Es ergibt sich folgendes:

```
> summary(zipf_fit)
Formula: chord_freq_df$relative_frequency ~ zipf_function(chord_freq_df$rank, a, b, c)
Parameters:
  Estimate Std. Error t value Pr(>|t|)
a 0.471908  0.013374   35.28  <2e-16 ***
b 1.664821  0.042370   39.29  <2e-16 ***
c 1.290851  0.009668   133.52  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.0004356 on 496 degrees of freedom
Number of iterations to convergence: 6
Achieved convergence tolerance: 3.027e-06

[1] "R^2 = 0.997536484227347"
```

Zuletzt wurde ein Plot Rang vs. rel. Häufigkeit mit dem Paket 'ggplot2' erstellt, in welchem die Funktion über die Datenpunkte gelegt wurde. (Abbildung 2)

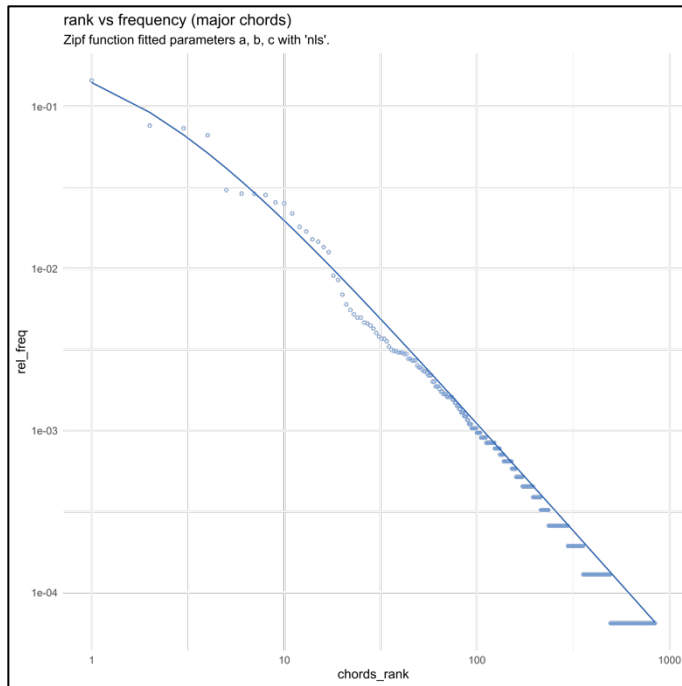


Abbildung 1

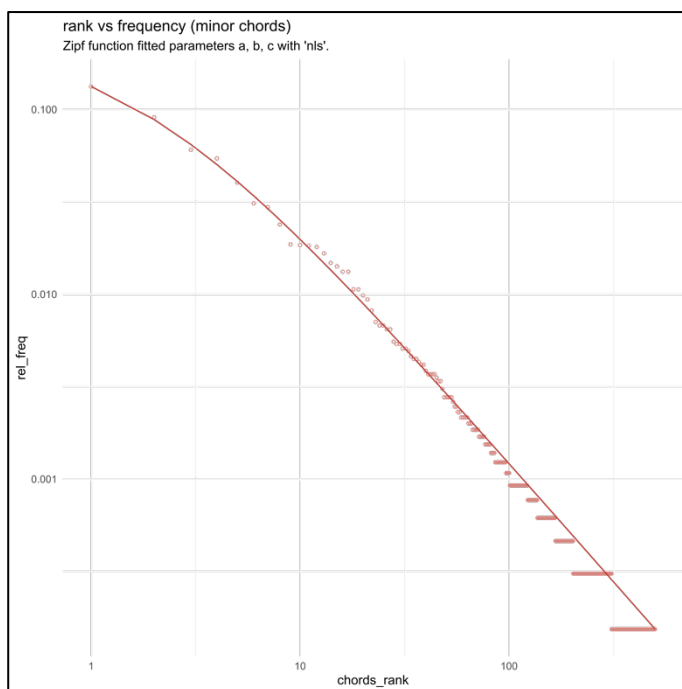


Abbildung 2