

GEBURTSTAGSPROBLEM

In der Vorlesung wurde das Geburtstagsproblem vorgestellt und stochastisch untersucht. Wir wollen in dieser Übung überprüfen, ob das theoretische Modell die Wirklichkeit ausreichend genau beschreiben kann.

Auf der Website finden Sie eine Liste aller Mitglieder des 18. Deutschen Bundestags mit Namen und Geburtsdatum.

1. Schreiben Sie ein Programm in C, dass die Liste einliest und in einem Array speichert. Es handelt sich um kommaseparierte Werte (CSV), die mithilfe der Funktionen **fopen**, **fgets** und **strtok** eingelesen werden können.
2. Die eingelesenen Daten sollen nun analysiert werden. Erweitern Sie ihr Programm, sodass es für jedes $k \in \{1..365\}$ die Wahrscheinlichkeit angibt, dass mindestens zwei von k Abgeordneten am selben Tag Geburtstag haben. Wählen Sie dazu mehrmals k zufällige Einträge aus der Liste und überprüfen Sie, wie oft mindestens zwei dieser k Personen am selben Tag Geburtstag haben.
3. Die statistisch ermittelten Werte aus Aufgabe 2 sollen anschließend mit den theoretischen Ergebnissen aus der Vorlesung verglichen werden. Erstellen Sie dazu ein Diagramm mithilfe von R, dass die in Aufgabe 2 ermittelten Werte und die nach folgender Formel berechneten Werte visualisiert.

$$P(n, k) > 1 - e^{-k \times (k-1) / 2n}$$

4. Sind die theoretischen und statistischen Ergebnisse gleich? Wenn nicht, geben Sie mögliche Ursachen für die Differenz an.