

Proposal zur Bachelorarbeit

**Bootstrapping Ansätze zur Bestimmung von  
Konfidenzbändern für Verteilungsfunktionen**

Dennis Richter

Gutachter:

Prof. Dr. Peter Buchholz

NaN

Technische Universität Dortmund

Fakultät für Informatik

Lehrstuhl für praktische Informatik (LS 4)

<https://ls4-www.cs.tu-dortmund.de>

# 1 Einleitung

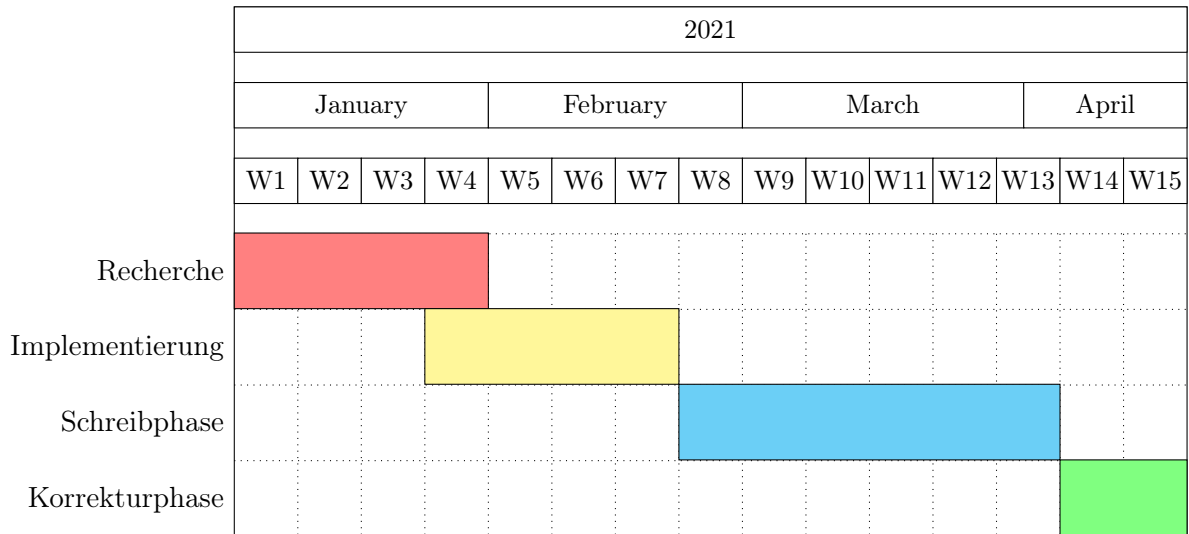
Um bei der statistischen Auswertung die Genauigkeit einer auf Grund von Stichproben ermittelten Schätzfunktion zu bestimmen, ist es oft notwendig, ein Konfidenzintervall für den Schätzer  $\hat{\theta}$  des Parameters  $\theta$  anzugeben. Ziel ist es dieses Intervall so zu konstruieren, dass es den wahren Wert von  $\theta$  mit hoher Wahrscheinlichkeit (oft 95%) überdeckt. Für einzelne Werte der Zufallsvariable kann dies leicht erreicht werden, häufig verwendete Verfahren sind z.B. die Maximum-Likelihood-Methode und die Bootstrap-Methode. In manchen Fällen jedoch benötigt man solche eine Bereichsschätzung für einen ganzen Bereich von Werten der Zufallsvariable. Da der Konfidenzbereich in diesem Fall für mehrere Werte der Zufallsvariable gleichzeitig gilt, spricht man nicht mehr von einem Konfidenzintervall, sondern von einem Konfidenzband. Mit anderen Worten, gesucht ist ein Konfidenzband, welches die gesamte wahre Verteilungsfunktion mit großer Wahrscheinlichkeit überdeckt.

## 2 Problembeschreibung und Ziele

Herkömmliche analytische Verfahren zur Bestimmung von Konfidenzbändern wie z.B. die Bonferroni-Methode, gehen oft davon aus, dass  $\theta$  Normalverteilt ist und die Regressionsfunktion  $\eta(x, \theta)$  linear angenähert werden kann. Ähnlich wie bei Konfidenzintervallen bietet Bootstrapping auch hier eine einfache Alternative in Fällen, in denen etwa eine Verteilungsannahme von  $\theta$  nur zweifelhaft bzw. gar nicht vorausgesetzt werden kann. Für den Einsatz von Bootstrap zur Bestimmung von Konfidenzbändern gibt es verschieden Herangehensweisen, zwei solcher Ansätze werden in [1, 2] diskutiert. Im ersten Fall nimmt man an, dass  $\theta$  Normalverteilt ist, verzichtet aber auf die lineare Approximation von  $\eta(x, \theta)$ . Man spricht daher von parametrischem Bootstrapping. Im zweiten Fall, dem nicht-parametrischen Bootstrapping, wird zusätzlich keine Verteilungsannahme getroffen, sondern die Verteilung von  $\theta$  wird ebenfalls mit Hilfe des Resampling-Verfahrens geschätzt. Diese Variante ist rechenintensiver, kann aber in allgemeineren Fällen angewendet werden.

Obwohl es einige theoretische Ansätze zur Bestimmung von Konfidenzbändern mittels Bootstrap gibt, sind Implementierungen der Verfahren und Umsetzungen als Algorithmen in der Literatur eher selten bis nicht existent. Ich möchte in meiner Arbeit einige solcher Methoden vorstellen und im Kontext von OMNeT++ implementieren. Die Verfahren sollen in C++ implementiert und anschließend nach Möglichkeit in die OMNeT++ IDE integriert werden. Darüber hinaus soll eine empirisch Bewertung der Konfidenzbänder anhand von einfachen Beispielen in Form einer OMNeT++ Simulation erfolgen. Folgenden Zeitplan hab ich für meine Arbeit vorgesehen:

### 3 Zeitplan



**W1 - W2:** Recherche zu Konfidenzbändern und Bootstrapping-Ansätze zur Bestimmung von Konfidenzbändern

**W3:** Inhaltsverzeichnis, Einleitung und Vorstellung der Bootstrapping-Ansätze verfassen

**W4:** Recherche zu Parameterstudien, Auswertung und Darstellung in Kontext von OMNeT++

**W5:** Erstellung einfacher Beispielen als Grundlage für die empirische Bewertung der Verfahren

**W6 - W7:** Implementierung der Verfahren in C++ und Integration in die OMNeT++ IDE

**W8:** Implementierung dokumentieren

**W9 - W11:** Anwendung der Verfahren und empirische Bewertung anhand der Beispiele dokumentieren

**W12 - W13:** Fazit schreiben und erster Entwurf der Bachelorarbeit sollte fertiggestellt werden

**W14 - W15:** Korrekturlesen und Anpassungen vornehmen

## Literatur

- [1] CHENG, RUSSELL: *Bootstrap confidence bands and goodness-of-fit tests in simulation input/output modelling*. Winter Simulation Conference, Seiten 16–30, 2015.
- [2] CHENG, RUSSELL C. H.: *Bootstrapping simultaneous confidence bands*. Winter Simulation Conference, Seiten 275–280, 1997.