

A 1 In dieser Aufgabe soll die Aussage des Zentralen Grenzwertsatzes für Bootstrap-Daten überprüft werden.

- a) Simuliere 50 bzw. 1000 Datensätze mit jeweils 500 gleichverteilten Daten auf $[0, 1]$, speichere die Mittelwerte im Vektor $Means_{sim}$.
- b) Simuliere einen weiteren Datensatz mit 500 gleichverteilten Daten auf $[0, 1]$ und erzeuge mit Bootstrap 50 bzw. 1000 Replikationen. Speichere die Mittelwerte der Replikationen im Vektor $Means_{boot}$.
- c) Vergleiche die Histogramme und Boxplots von $Means_{sim}$ und $Means_{boot}$, teste jeweils mit dem Shapiro-Wilk Test auf Normalverteilung.
- d) Vergleiche die Mittelwerte von $Means_{sim}$ und $Means_{boot}$ mit einem t-Test.
- e) Interpretiere und begründe Deine Beobachtungen.

A 2 In dieser Aufgabe sollen die Konfidenzintervalle einer linearen Regression mit den aus dem Bootstrap-Verfahren gewonnenen Konfidenzintervallen verglichen werden.

- a) Führe für die Parameter der multiplen linearen Regression aus Labor 1 und 2 jeweils ein Bootstrap-Verfahren mit 50, 100, 1000 und 10000 Replikationen durch. Vergleiche die Verteilungen der Parameter-Schätzer und interpretiere die Befunde (Histogramm, QQ-Plot, t-Test).
- b) Bestimme die Konfidenzintervalle mit Hilfe des Bootstrap Verfahrens mit 50, 100, 1000, und 10000 Replikationen, setze als Option *basic* ein.
- c) Vergleiche die Bootstrap-Konfidenzintervalle mit den Konfidenzintervallen aus der linearen Regression. Welche Anzahl von Replikationen ist empfehlenswert und warum?
- d) Vergleiche die Konfidenzintervalle bei 1000 Replikationen mit der Option *basic* mit den Ergebnissen wenn die Optionen *perc* bzw. *bca* verwendet werden.