



Statistik II

Übungsblatt 10

Abgabe: Do. 18.1.2007, 10.00 Uhr, Briefkasten: Statistik II.

Bei jeder Aufgabe können maximal 5 Punkte erreicht werden.

1. Betrachtet wird der Datensatz "Math Students". Er enthält Informationen über Studenten der Mathematik an der LMU München.
 - (a) Stellen Sie die Daten graphisch dar
 - i. Den Zusammenhang zwischen den Einflussvariablen.
 - ii. Die Diplomnote in Abhängigkeit der Einflussvariablen.
 - (b) Erstellen Sie ein geeignetes Modell für die Diplom Note, und interpretieren Sie es.
2. Lowess kann man in Streudiagrammen in Mondrian einzeichnen lassen (solange Rserve und R auf dem Rechner installiert sind). Vergleichen Sie die Benutzung von Lowess in den zwei Software Paketen an hand vom Streudiagramm SALBEG gegen AGE im Bank Datensatz.
3. Stellen Sie die "Total Bill" und "Tip in \$" aus dem "Tipping" Datensatz in Histogrammen dar.
 - (a) Berechnen Sie für beide Variablen Kerndichteschätzer für deren Verteilungen.
 - (b) Sind Sie zufrieden mit der Form der Schätzer am linken Rand? Wo liegt das Problem und was könnte man dagegen tun?
 - (c) Zeichnen Sie ein Streudiagramm für die Variablen. Berechnen Sie eine Regressionslinie und eine Lo(w)ess Glättung. Welchen Span würden Sie empfehlen und warum? Besprechen Sie Ihre Resultate.
4. Wir betrachten erneut den Datensatz "Tipping":
 - (a) Berechnen Sie den Median $x_{0.5}$ der "Tipping Rate" (hier werden wieder alle Werte betrachtet).
 - (b) Erstellen Sie 100 Bootstrap Samples der Tipping Rate (z.B. mit dem R-Befehl `sample`) und berechnen Sie jeweils $x_{0.5(i)}$, $i = 1, \dots, 50$.
 - (c) Wie groß ist die durchschnittliche Abweichung der Bootstrap Stichproben vom Median der Ursprungstichprobe, und wie groß ihr Standardfehler?
 - (d) Wie interpretieren Sie die Ergebnisse aus (c)?

5. In einem Experiment werden 9 Studenten zufällig ausgewählt und ihre Körpergröße bestimmt. Die Körpergröße sei normalverteilt mit Erwartungswert $\mu = 180$ cm und Standardabweichung $\sigma = 8.4$ cm. Simulieren Sie die Ergebnisse dieses Experiments und berechnen Sie für μ ein 95 % Konfidenzintervall. Wiederholen Sie diese Simulation insgesamt 100 mal.
- (a) Zeichnen Sie die erhaltenen Konfidenzintervalle in ein gemeinsames Koordinatensystem. Haben alle Intervalle die gleiche Länge?
 - (b) Wieviele der Intervalle enthalten μ ? Würden Sie erwarten, daß alle Intervalle μ enthalten?
 - (c) Führen Sie eine Bootstrapanalyse mit der ersten Ihrer Stichproben durch. Besprechen Sie die Unterschiede zwischen diesen Resultaten und den aus Ihren 100 Simulationen.