

Vorlesung "Empirische Bewertung in der Informatik"

Freie Universität Berlin, Institut für Informatik, Arbeitsgruppe Software Engineering
Prof. Dr. Lutz Prechelt, Stephan Salinger
Übungsblatt 1 SS 2007 Abgabe 2007-04-30 bis 09:00Uhr

Aufgabe 1-1: (Grundbegriffe von R erlernen)

- **Installieren** Sie sich R, bevorzugt in Version 2.4.1 (Referenzversion für die Vorlesung).
 - siehe <http://www.r-project.org>
 - Auf den öffentlichen Rechnerpools ist bereits eine Installation vorhanden (sowohl auf Windows als auch auf Linux)
- Der **Start** des Programms erfolgt über die Kommandos *Rterm* oder *Rgui* (bei Windows) oder *R* (bei Linux).
 - Die empfehlenswerteste Umgebung für die Benutzung von R ist der iESS-Mode, der von dem Zusatzpaket ESS (Emacs speaks statistics) im XEmacs-Editor bereitgestellt wird (wie in der Übung gesehen). Der Editor fungiert dann als Shell. ESS ist ebenfalls über [r-project.org](http://www.r-project.org) auffindbar.
- Laden Sie sich die Datei ***R_intro_session.txt*** von der Vorlesungswebseite herunter.
 - Diese Datei enthält die Kommandos, die in der R-Demo in der Übung benutzt wurden.
- Führen Sie alle diese Kommandos der Reihe nochmals aus und **vollziehen Sie nach**, was dabei passiert und warum.
 - Rufen Sie jeweils die Hilfe auf, um sich mehr Detailinformationen über die benutzten Funktionen und Operatoren zu beschaffen.
- Sie sollten die folgenden **Ziele** erreichen:
 - Für jedes Kommando erläutern können, was dort passiert und warum es passiert (ausgenommen die wiederholte Auswertung des Arguments bei 'replicate').
 - Die wichtigsten Operatoren von R kennen und verwenden können: Arithmetik, Vergleich, logische Verknüpfungen.
 - Ein allgemeines Verständnis der Behandlung von Vektoren, Matrizen und Listen in R erwerben: Bildung, Verwendung in Ausdrücken und Indizierung.
 - Eigene Funktionen definieren und verwenden können.
 - Mit den Hilfsfunktionen und der Dokumentation von R umgehen können.
- Lesen Sie nötigenfalls **Hintergrundwissen** in den Dokumenten "An Introduction to R" oder "The R language definition" nach, die Sie beide über `help.start()` finden.
- **Aufgabe:** Geben Sie drei Bereiche in der Dokumentation an (in drei verschiedenen Dokumenten), in denen die Indizierung von R erklärt wird.

Senden Sie die Lösung im Body einer Email **bis Montag, den 30.04.2007, 09:00 Uhr, an [salinger\[klam*mer*af*fe\]inf.fu-berlin.de](mailto:salinger[klam*mer*af*fe]inf.fu-berlin.de)** und bringen Sie einen Ausdruck in die nächste Übung mit. Die Email muss folgenden Betreff (Topic) haben: [UE Empir07-1.1][<Familienname>,<Vorname>]

Aufgabe 1-2: (Kleine R-Aufgabe)

Simulieren Sie **100000 Würfe mit einem Würfel** und zählen Sie nach, wie oft dabei die Folge 3, 4, 5 auftrat. Der Erwartungswert ist $(100000 - (3-1))/6^3$

Verwenden Sie eine möglichst gut zu R passende Formulierung (d.h. z.B. ohne explizite Schleife); nicht viel länger als zwei Zeilen.

Sie benötigen für die Lösung lediglich die Funktionen *sample* und *sum* sowie Indizierung und Operatoren -- es gibt aber natürlich auch noch andere Lösungsmöglichkeiten.

Senden Sie die Lösung (d.h. die R-Kommandos, nicht deren Resultat) per Email **als Attachment bis Montag, den 30.04.2007, 09:00 Uhr, an salinger[klam*mer*af*fe]inf.fu-berlin.de** und bringen Sie einen Ausdruck in die nächste Übung mit.

Das Attachment *muss* den Namen `<Familienname.Vorname>_ue1.2.R` haben.

Die Email muss folgenden Betreff (Topic) haben:

[UE Empir07-1.2][<Familienname>, <Vorname>]