

R Intro

Dr. C.J. Luchsinger

1 Vorbemerkungen

1. Wir geben hier keine vollständige Einführung sondern zeigen, wie man sich immer selber helfen kann. Bei Fragen: jobs@math-jobs.com oder 076 392 03 20 oder Literatur im Computerraum: Dalgaard resp. Lam.

2. R ist angelehnt an das kommerzielle Statistikpaket S-PLUS. Wenn man R kann, kann man auch S-PLUS. S-PLUS bietet auch einfache Eingabemasken, welche R nicht bietet. Weitere Statistikpakete sind SAS, SPSS, Minitab, etc.

3. Alle sollten innert 2 Wochen Kapitel 1 und 2 in Dalgaard selbständig durcharbeiten (auch diejenigen Personen, welche mit S-PLUS arbeiten; alternativ: p. 119-164 im Lam). Kapitel 3 (Beschreibende Statistik) sollte auch während des Semesters oder in den Semesterferien durchgearbeitet werden.

2 Intro-Session

```
> R
> q()
```

Mittelwerte, Varianzen

```
> a <- rnorm(5)
> a
> a <- rnorm(5)
> rnorm(5)
> sort(a)
> max(a)
> min(a)
> range(a)
> median(a)
> mean(a)
> mean
>?mean ("wie schneide ich die äussersten Werte ab"?)
> mean(a, 0.4)
> var(a)
>?var (teilt durch  $(n - 1)$  statt durch  $n$ )
> b <- a + 5
> b
> mean(b)
> var(b)
> d <- 5 * a
> var(d)
> 25 * var(a) (vgl. Lemma 3.7 a))
> ls()
> rm(a)
> ls()
```

Kovarianz/Korrelation

```
>?rexp ( $\lambda = 1$  ist der sogenannte Default; bei  $\mathcal{N}(\mu, \sigma^2)$  ist  $\mathcal{N}(0, 1)$  der Default)
> e <- rexp(10'000, 3)
> f <- rexp(10'000, 4)
> plot(e, f)
> cov(e, f) (vgl. Lemma 3.8 a); aber nicht exakt 0 da endliche Stichprobe)
> cor(e, f)
> g <- 5 + 10 * e
> cor(e, g) (vgl. Lemma 3.10 b))
> h <- 1/g
> plot(h, g) (mit Fingern rechts unten kopieren gehen mit Taste "↑")
> cor(g, h)
> mean(h)
> 1/mean(g) (vgl. Bemerkung 2 zu Lemma 3.5 [Jensen-Ungleichung])
```

Histogramm

```
> w <- rnorm(10)
> hist(w)
> w <- rnorm(100)
> hist(w)
> w <- rnorm(1'000'000)
> hist(w)
>?hist ("möchte feiner unterteilen")
> hist(w, 100)
```

**Ausdrucken (von Plots und Sessions) je nach System (Mac, Windows, Linux/Unix) anders;
Notfall mit Copy/Paste.**