

Grundlagen der Datenanalyse mit R, 3. Auflage – Errata

Stand: 10. November 2017

Inhaltlich relevante Korrekturen

- Abschn. 2.1.3, S. 99:
“> sd(na.omit(vec~~NA~~)) # um NA bereinigten Vektor übergeben
...
> goodIdx <- !is.na(vec~~NA~~) # Indizes der nicht fehlenden Werte
> mean(vec~~NA~~[goodIdx]) # um NA bereinigten Vektor übergeben
...
> sum(vec~~NA~~)
...
> sum(vec~~NA~~, na.rm=TRUE)”
- Abschn. 2.7.3, S. 66:
“Den Modalwert erhält man zusammen mit seiner Auftretenshäufigkeit durch Indizieren der mit unique() gebildeten **und dann geordneten** Einzelwerte des Vektors mit diesem Index.
> **sort(unique(vec))**[modIdx]”
- Abschn. 2.11.4, S. 100:
“Bei cov(~~<Vektor1>, <Vektor2>~~) und cor(~~<Vektor1>, <Vektor2>~~) bewirkt das Argument na.rm=TRUE, dass ein aus zugehörigen Werten von ~~<Vektor1>~~ und ~~<Vektor2>~~ gebildetes Wertepaar nicht in die Berechnung von Kovarianz oder Korrelation eingeht, wenn wenigstens einer der beiden Werte NA ist.
Bei den Funktionen cov() und cor() stehen zur Behandlung fehlender Werte der **zeilenweise und paarweise** Fallausschluss zur Verfügung, die über das Argument use="complete.obs" bzw. use="pairwise.complete.obs" ausgewählt werden.
Bei der Kovarianz bzw. Korrelation zweier Variablen führen sie zum selben Ergebnis, unterscheiden sich aber bei Kovarianz- bzw. Korrelationsmatrizen von mehr als zwei Variablen.”
- Abschn. 8.1.5, S. 295:
“> **summary(glmFit)**
Call:
glm(formula = postFac DVpre + IV, family = binomial(link = "logit"),
data = dfAncova)”
- Abschn. 9.2.2, S. 330:
Durch eine neue Version des survival-Pakets hat sich die Syntax von survsplit() geändert. Der Aufruf im Beispiel muss nun lauten:
“> dfSurvCP <- survSplit(Surv(obsT, status) ~, cut=seq(30,90,by=30),
+ start="start", id="ID", zero=0, data=dfSurv)”
- Abschn. 10.3.3, S. 376:
“Kappa~~M~~() aus dem Paket DescTools ermittelt den κ -Koeffizienten nach Fleiss”
“> Kappa~~M~~(rateMat, conf.level=0.95)”
(Änderung in Paket DescTools Version 0.99.9)
- Abschn. 10.3.3, S. 376:
“> symmetry_test(cTab~~Bow~~, teststat="quad") # ...”

- Abschn. 14.8.3, S. 558:
“... ist dies eine kategoriale Variable, deren Häufigkeiten von `geom_bar(stat="bin count")` automatisch ausgezählt und als Säulenhöhe visualisiert werden.”
“> pB2 <- pB1 + `geom_bar(stat="bin count", position=position_dodge())`”
(Änderung in Paket `ggplot2` Version 2.0.0)

Weitere Hinweise

- Abschn. 1.1, S. 1, Fußnote 1:
Weitere kommerzielle Anbieter von R-Distributionen sind “Oracle R Enterprise” sowie “HP distributed R”:
<http://www.oracle.com/technetwork/database/database-technologies/r/r-distribution/>
<http://www.vertica.com/hp-vertica-products/hp-vertica-distributed-r/>
- Abschn. 4.2.2, S. 155:
Eine aktuelle Alternative zu `read.table()` zum schnellen Lesen großer Textdateien ist das Paket `readr`:
<https://cran.r-project.org/package=readr>
- Abschn. 4.2.4, S. 159:
Eine neue Alternative zum Paket `foreign` zum Austausch von Daten mit SPSS, SAS und Stata ist das Paket `haven`:
<http://cran.r-project.org/package=haven>
Das Paket `readxl` ist eine aktuelle Alternative zu `XLConnect`:
<https://cran.r-project.org/package=readxl>
- Abschn. 10.3.3, S. 375:
“> (`agree <- sum(diag(prop.table(cTab)))`)”
Paket `DescTools` besitzt für die prozentuale Übereinstimmung seit Version 0.99.9 die Funktion `Agree()`, die eine spaltenweise aus den Urteilen zusammengesetzte Matrix erwartet.
> `library(DescTools)`
> `Agree(cbind(rtr1, rtr2))`
> `Agree(cbind(rtr1, rtr2, rtr3))`”

Tipp- und Druckfehler

- Abschn. 2.3.3, S. 40:
“das kartesische Produkt der übergebenen Grundme`n`gen”
- Abschn. 2.6.5, S. 60:
“... kann ein Faktor in einen geordneten Faktor unter Verwendung des `levels` Arguments un`m`gewandelt werden (s. o.)”
- Abschn. 2.10.5, S. 94:
“... ist es dann notwendig, sie zunächst in ein Format umzuwande`l`n, ...”
- Abschn. 3.3, S. 129:
“Bevor Datensätze analysiert werden können, müssen sie häufig in eine andere als ihre ur`s`prüngliche Form gebracht werden ...”

- Abschn. 4.2.4, S. 164:
“dbRemoveTable(\langle DB-Verbindung \rangle , " \langle table-Name \rangle ") löscht ein table in der verbundenen Datenbank”
- Abschn. 4.3.1, S. 167:
“Dort gibt choose.files() einen Vektor von Pfadnamen zu den ausgewählten Dateien zurück.”
- Abschn. 5.1, S. 171:
“... der griechische Buchstabe des zu schätzenden Parameters”
- Abschn. 7.3.6, S. 230:
“... stellt eine weitere Möglichkeit dar, auf Unterschieden zwischen jeweils zwei Gruppen zu testen.”
- Abschn. 8.3, S. 305, Abbildung 8.4:
“... für die Chance, verglichen mit der baseline Kategorie eine Kategorie g zu erhalten; ...”
- Abschn. 8.5.2, S. 323:
“# Parameterschätzungen mit Standardfehlern und Wald-Tests”
- Abschn. 9, S. 324:
“... wieviel Zeit bis zum Eintreten eines bestimmten Ereignisses verstrichen ist ...”
- Abschn. 9.2.2, S. 328:
“... die hier implizit verwendete Zeitskala flexibel etwa das Alter eines Beobachtungsobjekts sein...”
“... deren Lebensalter zu Beginn und zu Ende einer zehnjährigen Untersuchung ...”
- Abschn. 9.2.2, S. 329:
“... aus zwei Gruppen in einer zehnjährigen Untersuchung”
“... das zugehörige kalendarische Datum oder das Alter eines Beobachtungsobjekts.”
“Bei mehreren Beobachtungszeitpunkten muss der Datensatz ...”
- Abschn. 10.3.1, S. 373:
“... verträglich sind, dass der theoretische Zusammenhang 0 ist.”
- Abschn. 10.5.10, S. 397:
“... ob eine Stetigkeitskorrektur durchgeführt wird (Voreinstellung: TRUE).”
- Abschn. 12.4, S. 447:
“... die die multidimensionale Skalierung auf zwei Dimensionen räumlich anordnen soll.”
- Abschn. 12.6.1, S. 451:
“In der für lm() anzugebenden Formel $\langle AV \rangle \sim \langle UV \rangle \dots$ ”
- Abschn. 13.3, S. 498:
“... um dann einen fehlenden Wert anstatt der ungültigen Optimismus-Schätzung zurückzugeben.”
- Abschn. 14.3.1, S. 506:
“Die gebräuchlichsten von ihnen sind in Tab. 14.2 beschrieben.”
- Abschn. 14.5.10, S. 533:
“Für farbige Bilder ist ein array mit drei Ebenen zu verwenden.”
- Abschn. 14.8.3, S. 559:
“geom_smooth(method= \langle Methode \rangle , se=TRUE, level=0.95, fullrange=FALSE)”

- Abschn. 14.8.6, S. 565:
“Die Ausgabe entspricht der Struktur der übergebenen Matrix, ...”