

Einführung in R

11. Aufgabenblatt

Präsenzaufgabe 1

Der Datensatz `survey` aus dem Paket `MASS` enthält die Antworten von 237 Studierenden auf verschiedene Fragen. Unter anderen wurde das Rauchverhalten (Smoke) abgefragt. Das ordinal-skalierte Merkmal `Smoke` hat die Ausprägungen: `Heavy`, `Regul`, `Occas`, `Never`.

- a. Testen Sie die Nullhypothese

$$c(P_{Heavy}, P_{Regul}, P_{Occas}, P_{Never}) = c(0.05, 0.10, 0.10, 0.75)$$

zum Niveau $\alpha = 0.05$.

- b. Testen Sie die Unabhängigkeit der Merkmale `Smoke` und `Sex` im Datensatz `survey` zum Niveau $\alpha = 0.05$.

Präsenzaufgabe 2

Laden Sie den Datensatz `heartatk.txt` aus `Stud.IP`. Testen Sie die Abhängigkeit der Merkmale `SEX` und `DIED` zum Signifikanzniveau 5%. Zur Beantwortung der Frage benutzen Sie einmal den χ^2 -Test und einmal den Fisher-Test.

Präsenzaufgabe 3

In einer Studie nahmen Ärzte über 5 Jahre hinweg entweder regelmäßig Aspirin oder ein Placebo, ohne zu wissen zu welcher Gruppe sie gehören. Gezählt wurde in jeder Gruppe die Anzahl der Herzinfarkte.

	Infarkt	Ja	Nein
Gruppe			
Placebo		189	10845
Aspirin		104	10933

Besteht ein signifikanter Zusammenhang zwischen Herzinfarkt und der Einnahme von Aspirin? Verwenden Sie $\alpha = 0.05$ für Ihren Test. Zur Beantwortung der Frage benutzen Sie einmal den χ^2 -Test und einmal den Fisher-Test.

Präsenzaufgabe 4

Der Datensatz `Titanic` im Paket `datasets` enthält Daten über das Geschlecht, Alter, den Status und das Überleben von Fahrgästen des *Titanic*-Schiffs in Form eines 4-dimensionalen *array*. Testen Sie mit dem χ^2 -Test, ob es eine Korrelation zwischen dem Überleben der Fahrgäste und den anderen Merkmalen gibt.

Hinweis: nutze `ftable(Titanic, row.vars = "Class", col.vars = "Survived")`

Hausaufgabe 1 (6 Punkte)

In einer Studie wird nach dem Rauchverhalten von Männern und Frauen gefragt. Die folgende Tabelle zeigt die beobachteten Häufigkeiten:

	Nicht-Raucher	Raucher
Mann	14	1
Frau	26	12

Testen Sie mit dem *exakten Test von Fisher* und dem χ^2 -Test, ob es einen Zusammenhang zwischen Rauchen und Geschlecht gibt. Welcher der Tests ist hier besser geeignet und warum?

Hausaufgabe 2 (6 Punkte)

Laden Sie den Datensatz [airquality](#). Untersuchen Sie den Zusammenhang zwischen dem Merkmal [Ozone](#) und den Merkmalen [Solar.R](#), [Wind](#) und [Temp](#) in diesem Datensatz mit geeigneten Tests. Verwenden Sie $\alpha = 0.05$ für Ihre Tests.

Hausaufgabe 3 (8 Punkte)

Betrachten Sie im Datensatz [survey](#) aus dem Paket [MASS](#) die Variablen [Pulse](#) (Pulsrate, Schläge pro Minute) und [Smoke](#) (Rauchverhalten, verschiedene Kategorien). Gibt es einen statistisch signifikanten Zusammenhang zwischen der Pulsrate und dem Rauchverhalten? Verwenden Sie die Funktion [cor.test](#) für einen Test auf Basis von Kendalls Tau. Betrachten Sie die Variable [Smoke](#) als numerisch, d.h. ordnen Sie den Ausprägungen Never, Occas, Regul und Heavy die Zahlenwerte 1 bis 4 in einer sinnvollen Reihenfolge zu. (Never < Occas < Regul < Heavy) Verwenden Sie $\alpha = 0.05$ für Ihre Tests.

Abgabe der Lösungen: bis **Dienstag 13.01.2020**,

maendle@uni-bremen.de