

## Statistik II — Aufgabenzettel 6

(Abgabe: 24.06.2024)

### Hinweise zur Abgabe dieses Aufgabenzettels

- es handelt sich um eine Gruppenabgabe, wobei Gruppen mit 2-3 Mitgliedern zulässig sind — **Einzelabgaben sind nicht mehr zulässig**
- Gruppen können tutorienübergreifend gebildet und bei späteren Aufgabenzetteln jederzeit neu zusammengesetzt werden
- jede Person kann in nur einer Gruppe mitwirken und jede Gruppe nur eine Lösung abgeben
- die Abgabe ist bis zum 24.06.2024, 23:59 Uhr über Moodle möglich und von nur *einer* (beliebigen) Person aus jeder Gruppe durchzuführen
- geben Sie in Ihrer Abgabedatei auf der ersten Seite der Lösungen klar und deutlich alle Mitglieder der Gruppe an (Vorname, Name, Matrikelnummer)
- wir empfehlen, dass die abgebende Person sowohl die finale Abgabedatei als auch einen Screenshot der Abgabe in Moodle an die anderen Gruppenmitglieder schickt
- Lösungen können u.a. aus folgenden Komponenten bestehen: R Code und R Output, Screenshots, abfotografierte oder eingescannte handschriftliche Lösungen, mit einem Editor wie MS Word erstellte Lösungen — wir sind hier flexibel, alles was lesbar ist wird akzeptiert
- die einzelnen Komponenten sind für die Abgabe jedoch zu einer *einzigsten* **PDF Datei** zusammenzufügen — hierfür gibt es kostenlose Online-Tools, z.B. Smallpdf
- der Lösungsweg muss immer klar ersichtlich und die Lösung vollständig sein — sollten beispielsweise Grafiken zu erstellen sein, so sollten diese auch Teil Ihrer Abgabe sein
- wir nutzen das folgende Punkteschema (pro Aufgabe):
  - gar nichts gemacht  $\rightsquigarrow$  0 Punkte
  - sich an der Aufgabe versucht, aber sehr wenig richtig gemacht  $\rightsquigarrow$  1 Punkt
  - teilweise richtige Lösungen/Lösungsansätze vorgelegt  $\rightsquigarrow$  2 Punkte
  - die Aufgabe gut bearbeitet, mit kleinen Schönheitsfehlern  $\rightsquigarrow$  3 Punkte
  - die Aufgabe vollständig zufriedenstellend bearbeitet  $\rightsquigarrow$  4 Punkte

(bewertet wird nicht kleinteilig jede Teilaufgabe, sondern nach Gesamteindruck)

- nach Ende der Abgabefrist werden Musterlösungen bei Moodle bereitgestellt, eine Besprechung in den Tutorien wird es nicht geben
- falls während der Bearbeitungszeit Fragen auftreten sollten, dann melden Sie sich jederzeit gerne über die [Pinnwand](#) (eine Antwort kommt in der Regel innerhalb einiger Stunden)

## Aufgabe 1: Sinusvenenthrombosen

Im Zeitraum Anfang Januar bis Ende April 2021 wurden in Deutschland ca. fünf Millionen Menschen mit dem Astrazeneca-Wirkstoff gegen Corona geimpft. In etwa 10 dieser Geimpften erlitten eine Sinusvenenthrombose. Unter normalen Umständen (d.h. ohne Impfung) liegt das Risiko, in einem solchen Zeitraum eine Sinusvenenthrombose zu bekommen, bei ca. 0.000025%. Überprüfen Sie mit einem Binomialtest zum Niveau  $\alpha = 0.01$ , ob die Anzahl an Sinusvenenthrombosen unter den Geimpften dafür spricht, dass das Risiko einer solchen Erkrankung durch die Astrazeneca-Impfung erhöht wird.

## Aufgabe 2: Buy-the-Dip

In dieser Aufgabe sollen Sie untersuchen, ob die durchschnittliche prozentuale Änderung des Deutschen Aktienindex (DAX) an einem Tag systematisch von Null abweicht, wenn der Index am Vortag um 5% oder mehr gefallen ist. Laden Sie hierzu wie folgt die Daten in R:

```
DAX <- read.csv("http://www.rolandlangrock.com/Daten/DAX89-24.csv")$x*100
DAX
```

Der Vektor DAX enthält die prozentualen Änderungen des DAX an jedem Börsentag seit 1987. Führen Sie nun die folgenden Schritte aus:

- identifizieren Sie zunächst alle Beobachtungen kleiner gleich  $-5$ , d.h. alle Börsentage, an denen der DAX um 5% oder mehr gefallen ist;
- speichern Sie anschließend die Beobachtungen an den jeweiligen *Folgetagen* in einem Vektor ab;
- untersuchen Sie zuletzt mit einem  $t$ -Test zum Niveau  $\alpha = 0.05$ , ob die im vorherigen Schritt ermittelten prozentualen Änderungen systematisch von Null verschieden sind;
- geben Sie insb. auch den  $p$ -Wert für den von Ihnen durchgeführten Test an.

Optional (ohne Wertung): Welche naheliegende Strategie ergibt sich aus dem Ergebnis?

## Aufgabe 3: Roulette

Wir wollen bei einem neu konstruierten Rouletterad statistisch überprüfen, ob die 37 möglichen Ausgänge (entsprechend der Zahlen  $0, 1, \dots, 36$ ) wie gewünscht mit jeweils gleicher Wahrscheinlichkeit eintreten. Dazu drehen wir das Rouletterad 1000 Mal und notieren jeweils das Ergebnis. Die so erhaltenen Beobachtungen können Sie wie folgt in R laden:

```
Roulette <- read.csv("http://www.rolandlangrock.com/Daten/RouletteDaten.csv")$x
Roulette
```

Überprüfen Sie mit einem Chi-Quadrat-Anpassungstest zum Niveau  $\alpha = 0.05$ , ob diese Daten Evidenz gegen die Annahme einer Gleichverteilung der 37 Ausgänge liefern. Geben Sie auch den  $p$ -Wert an.

## Aufgabe 4: Einstiegsgehälter

Laden Sie wie folgt die Umfragedaten zum Geschlecht sowie dem erwarteten monatlichen Einstiegsgehalt in R:

```
dat <- read.csv("http://www.rolandlangrock.com/Daten/Gehaelter.csv")
dat
```

Untersuchen Sie mit einem (approximativen) Zwei-Stichproben-Gaußtest zum Niveau  $\alpha = 0.01$ , ob sich die Erwartungen hinsichtlich des Einstiegsgehalts zwischen weiblichen und männlichen Studierenden unterscheiden. Geben Sie in Ihrer Lösung den Wert der Teststatistik, den Ablehnungsbereich sowie den  $p$ -Wert an.