

## Fallstudien I

### Projekt 2 - Vergleich zweier Verteilungen Platzverweise in der deutschen Fußball-Bundesliga

Das Ziel dieses Projektes ist es zu untersuchen, ob sich die Anzahl an Platzverweisen (gelb-rote und rote Karten) in der deutschen Fußball-Bundesliga zwischen Heim- und Auswärtsteams unterscheidet. Zu erwarten ist zwar, dass die Anzahl der Platzverweise für unterschiedliche Mannschaften variiert, allerdings ist unklar ob Teams als Heim- bzw. Auswärtsmannschaft auch generell zu einer unterschiedlich aggressiven Spielweise neigen.

Um diese Fragestellung zu untersuchen, liegen in der Datei `buli.Rdata` Daten zu allen 3672 Bundesligaspielen der Saisons 2006/07 bis 2017/18 vor. Die Daten wurden von der Internetseite <http://www.football-data.co.uk/germanym.php> bezogen. Im Datensatz enthalten sind das Datum des Spiels (`Date`), der Name der Heim- (`HomeTeam`) bzw. der Auswärtsmannschaft (`AwayTeam`), die Anzahl der Heimtore (`FTHG`), die Anzahl der Auswärtstore (`FTAG`), das Endergebnis (`FTR`; H für Heimsieg, D für Unentschieden, A für Auswärtssieg), die Anzahl der roten Karten des Heim- (`HR`) bzw. des Auswärtssteams (`AR`), die Anzahl der gelben Karten des Heim- (`HY`) bzw. des Auswärtssteams (`AY`), sowie die zugehörige Saison.

Führen Sie zunächst eine deskriptive Analyse des Datensatzes durch und betrachten Sie dabei zum Beispiel folgende Fragestellungen (hier dürfen Sie aber gerne auch weitere Aspekte untersuchen, die Ihnen relevant erscheinen):

- Unterscheidet sich die Anzahl an roten Karten zwischen Heim- und Auswärtssteams?
- Welche Teams neigen am häufigsten zu Platzverweisen? Gibt es Unterschiede bei den Teams zwischen Heim- und Auswärtsauftritten?
- Hat sich die Anzahl an roten Karten über die verschiedenen Saisons verändert?

Analysieren Sie daraufhin detailliert, inwiefern sich die Anzahl an roten Karten zwischen Heim- und Auswärtssteams unterscheidet. Verwenden Sie dazu sinnvolle statistische Verfahren um folgende Teilfragen zu untersuchen:

- Mit welcher Verteilung kann die Anzahl der Platzverweise pro Spiel für Heim- bzw. Auswärtssteams (`HR` und `AR`) geeignet modelliert werden? Diskutieren Sie, ob die zu treffen-

den Verteilungsannahmen gerechtfertigt werden können und vergleichen Sie die beiden angepassten Verteilungen.

- Wie gut ist die Anpassung der entsprechenden Verteilungen bei den vorliegenden Daten? (Vergleichen Sie dazu z.B. die beobachteten relativen Häufigkeiten mit den theoretischen Wahrscheinlichkeitsfunktionen der angepassten Verteilungen)
- Überprüfen Sie mit einem geeigneten Testverfahren, ob sich die Anzahl an Platzverweisen zwischen Heim- und Auswärtsteams zu einem im Vorfeld gewählten Signifikanzniveau signifikant unterscheidet.

**Vorgeschlagene Methoden:** Diskrete Wahrscheinlichkeitsverteilungen, Parameterschätzung / Anpassung, statistische Hypothesentests (parametrisch / nichtparametrisch)

## Literatur

- [1] Herbert Büning und Götz Trenkler. *Nichtparametrische statistische Methoden*. De Gruyter, 2013.
- [2] Hans-Friedrich Eckey, Reinhold Kosfeld und Matthias Türck. *Wahrscheinlichkeitsrechnung und Induktive Statistik*. Springer, 2005.
- [3] Ludwig Fahrmeir, Christian Heumann, Rita Künstler, Iris Pigeot und Gerhard Tutz. *Statistik: Der Weg zur Datenanalyse*. 8. Auflage, Springer, 2016.
- [4] Bernhard Rüger. *Test- und Schätztheorie, Band II: Statistische Tests*. Oldenbourg, 2002.