

*Hinweis:* Nach einer Wiederholung verschiedener Modellklassen sollen diese Aufgaben im Tutorium selbstständig gelöst werden. Die Lösung wird am Ende vorgestellt.

**Aufgabe 3** (Modellklassen Beispiele)

Welche statistischen Modelle würden Sie in den folgenden Situationen verwenden?

- (a) Die Bundesregierung möchte sich ein Bild von der Einkommenssituation in Deutschland verschaffen. Dazu werden von 100.000 zufällig ausgewählten Personen das monatliche Einkommen und einige Kovariablen erhoben. Von Interesse ist dabei der Zusammenhang von Einkommen und Einflussgrößen wie Schulbildung, Migrationshintergrund, Beruf, Alter und Geschlecht.
- (b) Eine BWL-Professorin möchte wissen, ob gleiche Anteile von Männern und Frauen in einem Unternehmen tatsächlich zu besseren Unternehmensbilanzen führen. Dazu betrachtet sie die Jahresbilanzen der Dax-Unternehmen der letzten 20 Jahre und erhebt zusätzlich den Anteil Frauen pro Unternehmen und Jahr. Ihre Hypothese ist, dass sich ein ausgewogenes Verhältnis von weiblichen zu männlichen Mitarbeitern positiv auf den Unternehmenserfolg auswirkt, über den genauen Verlauf möchte sie aber keine zu restriktiven Annahmen treffen.
- (c) Nach einem verlorenen Fußballspiel herrscht in der Kabine großer Streit. Die Abwehrspieler werfen den Mittelfeldspielern vor, dass diese sich zu wenig an der Defensivarbeit beteiligen würden. Die Mittelfeldspieler finden diesen Vorwurf ungerechtfertigt und behaupten, dass Mittelfeldspieler generell viel mehr laufen müssten als Abwehrspieler oder Stürmer. Um diese Aussage zu überprüfen, wird am nächsten Bundesligaspieltag die Laufleistung aller Spieler in allen Partien gemessen. Da davon ausgegangen wird, dass es Partien gibt, in denen generell mehr bzw. weniger gelaufen wird, werden die Laufleistungen vor der Analyse noch standardisiert. Nach der Standardisierung sollen die Laufleistungen innerhalb und zwischen den Mannschaften als unabhängig angenommen werden.
- (d) Um die Wirksamkeit eines neuen Medikaments zu testen wird an 100 Patienten an 10 aufeinanderfolgenden Tagen gemessen, ob sie Kopfweg haben oder nicht. Dabei erhält die eine Hälfte der Patienten das neue Medikament und die andere Hälfte ein Placebo.
- (e) Ein Sportwissenschaftler interessiert sich für die Frage, wie die Anzahl Turniersiege von Tennisspielern von deren Alter abhängt. Seine Hypothese ist, dass die Anzahl der Turniersiege zunächst mit der steigenden Erfahrung steigt und gegen Ende der Karriere, wo das Verletzungsrisiko zunimmt, wieder sinkt.

- (f) Im Verlauf einer Skisprungsaison werden bei jedem Wettkampf die Weiten der 50 besten Skispringer aus dem Vorjahr miteinander verglichen. Von Interesse ist, inwiefern die Leistung der Skispringer von Kovariablen wie Größe, Gewicht, Trainingsintensität, Nationalität oder Familienstand abhängt. Zur besseren Vergleichbarkeit werden die Weiten dabei pro Wettkampftag standardisiert um die Unterschiede zwischen den Schanzen weitestgehend zu eliminieren.
- (g) Im Münchner Mietspiegel wird der Mietpreis pro Quadratmeter untersucht. Von Interesse sei hier der Zusammenhang von Mietpreis und Wohnungsgröße, wobei davon ausgegangen wird, dass die Kovariable „Wohnungsgröße“ einen nichtlinearen Einfluss hat.
- (h) Eine Wildtierökologin interessiert sich dafür, wo im Schwarzwald wie viele Rotfüchse leben. Da sie nicht im gesamten Schwarzwald alle Füchse zählen kann, beschränkt sie sich auf einige Probegebiete und möchte dann aufgrund des dort festgestellten Zusammenhangs zwischen der Anzahl gesichteter Füchse und bestimmten Landschaftsparametern Vorhersagen für den gesamten Schwarzwald treffen. (Die Erhebung der Landschaftsparameter für den gesamten Schwarzwald stellt nämlich keine Schwierigkeit dar.) Dabei geht sie davon aus, dass eine der Kovariablen, nämlich der Anteil Wald im Umkreis von 1000 m, einen nichtlinearen Einfluss hat. Das begründet sie damit, dass die Füchse einerseits den Wald brauchen um sich zurückziehen zu können und andererseits die Nähe zu Siedlungen und Äckern schätzen, da sie dort größere Chancen haben, Nahrung zu finden. In extrem walddreichen bzw. walddarmen Gebieten erwartet sie also weniger Füchse als in heterogenen Gebieten.

#### **Aufgabe 4** (Modellklassen Zusammenhänge)

Erstellen Sie eine Art Mind Map, aus dem die Beziehungen der besprochenen Modelle ersichtlich werden.