MAGISTERARBEIT

Titel der Magisterarbeit

„Analyse von Werbedaten in Österreich mittels moderner statistischer Methoden“

Verfasserin

Nina Huber, Bakk. rer. soc. oec.

angestrebter akademischer Grad

Magistra der Sozial- und Wirtschaftswissenschaften (Mag. rer. soc. oec.)

Wien, im August 2009

Studienkennzahl lt. Studienblatt: A 066 951
Studienrichtung lt. Studienblatt: Magisterstudium Statistik
Betreuer: Ao. Univ.-Prof. Mag. Dr. Andreas Futschik

Zusammenfassung

Die vorliegende Arbeit untersucht den Ad-Bench-Datensatz der Firma „Focus Institut Marketing Research Ges.m.b.H.“, der sich mit österreichischen Werbungen aus den Bereichen „Hörfunk“, „Infoscreen“, „Online“, „Plakat“, „Print, „Prospekt“ und „TV“ beschäftigt. Es wird der Einfluss des Aufwandes, dessen Veränderung über die Zeit und, wenn vorhanden, auch die Wirkung der Länge auf die Bekanntheit, auf den Markenimpact und auf die Gefälligkeit eines Sujets untersucht. In diesem Zusammenhang soll auch geklärt werden, ob sich Männer von Frauen und die verschiedenen Altersgruppen untereinander unterscheiden.
Im weiteren Verlauf wird analysiert, ob die Wirtschaftsbereiche bezüglich der Beurteilung der Sujets differieren, wie die diversen Werbeträger wahrgenommen werden und welche von ihnen besonders sympathisch wirken. Außerdem wird geprüft, welcher Werbeträger sich durch Effizienz auszeichnet und ob die einzelnen Längenkategorien sich bezüglich der Effizienz unterscheiden.

Die Analyse der ersten Fragestellung wird mit linearer und logistischer Regression durchgeführt. Die restlichen Untersuchungen basieren auf nichtparametrischen Gruppenvergleichen mittels Kruskal-Wallis-Tests und anschließenden Post-Hoc-Tests mit Bonferroni-Korrektur.

Die Länge eines Sujets beeinflusst alle drei abhängigen Variablen, der Aufwand bei den Fernsehspots ebenfalls, in den restlichen Werbeträgern nur die beiden Variablen *Bekanntheit* und *Markenimpact*. Der Verlauf über die Zeit ist nicht eindeutig. Es gibt sowohl positive als auch negative Veränderungen. Die Zielgruppen unterscheiden sich nicht sehr stark voneinander.
Bei den Wirtschaftsbereichen liegen die Kategorien „Touristik“ und „Getränke“ in der Gefälligkeit hoch, die Kategorie „Reinigung“ befindet sich dagegen in einem niedrigen Bereich.
TV-Werbungen sind die bekanntesten Werbungen und werden von den Befragten auch als besonders sympathisch eingestuft. Prospekte zeichnen sich sowohl durch hohe Werte bei der Wiedererkennung als auch bei der Gefälligkeit aus.
Der billigste und auch effizienteste Werbeträger ist der Infoscreen, am ineffizientesten ist das Printmedium. Am effizientesten sind die kürzesten Spots, weil sie am wenigsten Aufwand erfordern. Interessant ist allerdings, dass sich Spots mit einer Dauer von 20 bis 25 Sekunden von jenen mit einer Dauer von 15 bis 20 Sekunden nicht unterscheiden. Der Mehraufwand, der bei teureren Spots natürlich anfällt, wird durch eine höhere Bekanntheit ausgeglichen.

Inhaltsverzeichnis

1 Einleitung **1**

2 Methoden **4**
2.1 Einführung und einführende Begriffe 4
2.2 Hypothesentests 6
2.2.1 Wald-Test 8
2.2.2 Likelihood-Ratio-Test 9
2.2.3 Tests auf Gleichheit zweier Mittelwerte 10
2.2.3.1 T-Test 10
2.2.3.2 Wilcoxon-Rangsummentest 11
2.2.4 Tests auf Gleichheit der Varianzen 12
2.2.4.1 Levene-Test 12
2.2.4.2 Bartlett-Test 13
2.2.5 Kolmogorov-Smirnov-Test 14
2.3 Regressionsanalyse 14
2.3.1 Verallgemeinerte lineare Modelle 15
2.3.1.1 Exponentialfamilie von Verteilungen 15
2.3.1.2 Modellformulierung 17
2.3.1.3 Schätzung der Parameter 18
2.3.1.4 Interpretation der Parameter 19
2.3.1.5 Wechselwirkungen 20
2.3.1.6 Nominale Variablen in der Regression 21
2.3.1.7 Nichtlineare Terme 22
2.3.1.8 Spline-Funktionen 22
2.3.1.9 Variablenselektion 26
2.3.2 Einfache lineare Regression 26
2.3.2.1 Schätzung der Parameter 29
2.3.2.2 Interpretation der Parameter 30
2.3.2.3 Modellanpassung 30
2.3.3 Logistische Regression 32
2.3.3.1 Link-Funktionen 33
2.3.3.2 Interpretation der Parameter 35
2.3.3.3 Gütekriterien 36
2.4 Gruppenvergleiche 36
2.4.1 Varianzanalyse 36
2.4.2 Kruskal-Wallis-Test 39
2.4.3 Korrektur nach Bonferroni 40
2.5 Datentransformation 41
2.6 Fehlende Werte 43

2.6.1 Singuläre Imputation 43
2.6.2 Multiple Imputation 44

3 Auswertung **45**
3.1 Software 45
3.2 Beschreibung des Datensatzes 46
3.3 Einfluss von Aufwand, Länge und Zeit 52
3.3.1 Einfluss im Werbeträger „Hörfunk“ 56
3.3.2 Einfluss im Werbeträger „Infoscreen“ 66
3.3.3 Einfluss im Werbeträger „Print“ 76
3.3.4 Einfluss im Werbeträger „TV“ 83
3.4 Bewertungsvergleich der Wirtschaftsbereiche 90
3.5 Vergleich der Werbeträger 95
3.6 Effizienzvergleich der Werbeträger 101
3.7 Sympathievergleich der Werbeträger 105
3.8 Effizienzvergleich der Längenkategorien 107

4 Interpretation und Zusammenfassung **111**

Abbildungsverzeichnis **119**

Tabellenverzeichnis **121**

Literatur **124**

Lebenslauf **127**

1 Einleitung

Diese Arbeit beschäftigt sich mit einem Datensatz, der von der Firma „Focus Institut Marketing Research Ges.m.b.H." erhoben wird. Er untersucht den Einfluss und die Wirkung von Werbesujets in Österreich. Dazu zählen Werbeauftritte im Hörfunk oder Fernsehen, auf Webseiten im Internet, in Printmedien oder Prospekten, auf Plakaten oder Infoscreens. Dies geschieht für einige dieser Werbeträger schon seit Anfang des Jahres 1998. Folgende Fragestellungen sollen in dieser Arbeit beantwortet und diskutiert werden:

- Hat die Höhe der Werbeausgaben einen Einfluss auf die Bekanntheit, die Bewertung oder den Wiedererkennungswert eines Sujets?
- Beeinflusst die Länge eines Spots die oben genannten Variablen?
- Verändert sich dieser Einfluss über die Jahre?
- Gibt es in den Zielgruppen unterschiedliche Auswirkungen des Aufwandes, der Zeit oder der Länge?
- Unterscheiden sich die einzelnen Wirtschaftsbereiche in der Beurteilung der Sujets?
- Wie werden die diversen Werbeträger wahrgenommen?
- Wie effizient sind die einzelnen Werbeträger?
- Welche Werbeträger eignen sich besonders für den Aufbau von Sympathie?
- Welche Kosten-Nutzen-Relation von Spotlängen gibt es?

Die Intention der Erhebung ist nicht die Beantwortung solcher Fragen, der Datensatz soll vielmehr das Durchsetzungsvermögen eines Sujets im Konkurrenzumfeld und dessen Effizienz messen (siehe Sujet Focus, „Focus Institut Marketing Research Ges.m.b.H."). Bei der Untersuchung wird das interessierende Sujet mit den restlichen relevanten Werbungen in Zusammenhang gesetzt und verglichen. Fragen, wie die oben genannten, wurden mit diesen konkreten Daten noch nie behandelt.
Der Datensatz steht nur in aggregierter Form zur Verfügung, würde er auf Personenebene vorliegen, könnte man noch weitere Fragen behandeln bzw. Details untersuchen. Außerdem werden in einigen Werbeträgern interessante Informationen wie der Aufwand nicht erhoben, weshalb man dessen Einfluss

nicht kalkulieren sowie auch keinen Effizienzvergleich anstellen kann. Außerdem unterliegen die Prospekt-Werbungen einem alternativen Abfragesystem, weshalb sie mit den restlichen Werbeträgern nicht vergleichbar sind. Zusätzliche Fragen, die im Laufe der Analyse aufgetaucht sind, findet man in der Zusammenfassung.

Im Zuge der Auswertung fiel auf, dass die Bewertung eines Sujets nur bei den TV-Werbungen mit dem Aufwand in Verbindung steht. Die Aussage „Je höher der Aufwand, desto besser ist eine Werbung“ ist also in diesem konkreten Datensatz bei den Hörfunk- und Infoscreen-Spots sowie auch bei den Print-Sujets nicht gültig. Wiedererkennungswert und Bekanntheit hängen aber sehr wohl mit dem Aufwand zusammen. Dieser Einfluss ist in einigen Gruppen sogar negativ.
Ein längerer Spot führt im Allgemeinen auch zu höheren Werten in den interessierenden Größen, auf keinen Fall sinken die Werte für länger dauernde Sujets. Der Einfluss verändert sich sehr wohl mit der Zeit. Diese Veränderung ist aber weder eindeutig positiv noch eindeutig negativ.

Die einzelnen Wirtschaftsbereiche weisen Unterschiede in ihrer Beurteilung auf. Es gibt ein eindeutiges Ranking der besten bzw. der schlechtesten Sujets. Im mittleren Bereich ist dies ein bisschen schwieriger zu verdeutlichen. Die am höchsten bewerteten Sujets gehören zu den Wirtschaftsbereichen „Touristik“ und „Getränke“, die schlechtesten sind Spots aus dem Bereich „Reinigung“.

Im Vergleich der Werbeträger hinsichtlich ihrer Bekanntheit liegen die TV-Spots am besten platziert, den größten Wiedererkennungswert haben Prospekte, die höchste Bewertung für die Gefälligkeit haben wiederum Prospekte, dicht gefolgt von Infoscreen-Werbungen.

Den effizientesten Werbeträger stellt der Infoscreen dar, da sie sehr billig ist. Sehr ineffizient sind die Print-Werbungen, in dem das Sujet mit der höchsten Effizienz, nämlich 46.371,17, liegt. Dies ist ein Sujet aus dem Wirtschaftsbereich „Finanzen“ und der Warengruppe „Sparen/Anlagen/Anleihen“ mit einem Aufwand von 668.468,7€ und einer Bekanntheit von lediglich 14,42 %. Am effizientesten sind natürlich die kürzesten Spots. Doch eine Sonderstellung haben Spotlängen zwischen 20 und 25 Sekunden. Sie liegen in der Effizienz in etwa gleich wie die etwas kürzeren Spots, die nur 15 bis 20 Sekunden dauern. Der ineffizienteste Spot in diesem Zusammenhang ist eine Fernsehwerbung aus dem Wirtschaftsbereich „Haus & Garten“ und der Warengruppe „Alarmanlagen“. Er hat eine Effizienz von 21.526,72.

Fernsehwerbungen werden von den Befragten als besonders sympathisch gekennzeichnet, dicht hinter den Sympathiewerten der TV-Spots liegen die Infoscreen-Spots. Es gibt ein Sujet, das von keiner befragten Personen als sympathisch bezeichnet wurde. Dies ist eine Print-Werbung aus dem Wirtschaftsbereich „Haus & Garten“ und der Warengruppe „Wäschetrockner“.

Im folgenden Kapitel wird ein kurzer Einblick in die Methodik gegeben. Es werden die in der Analyse verwendeten Verfahren beschrieben. Die ersten vier Punkte der Fragestellungen werden mit linearen und logistischen Regressionsmodellen behandelt. Es wird für jede Zielgruppe ein eigenes Modell angepasst, wobei die Variablen *Aufwand* und *Zeit* als restringierte kubische Splines und die Variable *Länge* als Dummy-Variable modelliert werden. Mit Likelihood-Ratio- und Wald-Tests wird überprüft, ob diese Variablen einen Einfluss haben oder nicht. Die restlichen fünf Fragen untersucht man mit dem Kruskal-Wallis-Test, welcher eine nichtparametrische Alternative zur Varianzanalyse darstellt. Er vergleicht auch bei Verletzung wichtiger Voraussetzungen mehrere Gruppen miteinander. Post-Hoc-Tests werden als Wilcoxon-Test gerechnet und konservativ mit der Bonferroni-Methode korrigiert.
Die Theorie, die hinter der Auswertung steckt, wird in dieser Arbeit nur kurz angesprochen, es wird also nur das für das Verständnis Wesentliche erklärt. Für weiterführende Details wird auf relevante Literatur verwiesen.
In Kapitel 3 werden die eingangs genannten Fragestellungen ausgewertet und beschrieben. Im anschließenden vierten Kapitel werden die Ergebnisse zusammengefasst und interpretiert. Dabei werden auffällige oder unerwartete Effekte genauer untersucht.

2 Methoden

2.1 Einführung und einführende Begriffe

Der erste Schritt in einer statistischen Analyse ist die Erhebung der Daten. Auf die Versuchsplanung wird in dieser Arbeit aber nicht näher eingegangen. Die Einführung in die Methoden wird nach den Büchern Bleymüller, Gehlert und Gülicher (2002), Schlittgen (2003) und Rinne (1995) vorgenommen.

Man befindet sich also in einer Situation, in der die Daten bereits vorliegen. Angenommen, man erhebt n Untersuchungseinheiten und k Variablen. Diese Informationen werden in einer Tabelle mit n Zeilen und k Spalten, der sogenannten Datenmatrix, zusammengefasst. Die Daten in einer Datenmatrix können nun entweder in gruppierter oder in ungruppierter Form vorliegen. Gruppierte Daten liegen dann vor, wenn Ausprägungen einzelner Variablen bei mehreren Untersuchungseinheiten vorkommen. Man braucht dann nicht die gesamte Datenmatrix angeben, sondern es reicht, wenn man den Mittelwert bzw. bei binär codierten Variablen die Summe der Ausprägungen angibt. Bei ungruppierten Daten entspricht jede Zeile einer Untersuchungseinheit, außerdem sind die Ausprägungen der k Variablen für die meisten Untersuchungseinheiten unterschiedlich.

Variablen werden auf unterschiedlichen Skalenniveaus gemessen. Man unterscheidet Nominal-, Ordinal-, Intervall- und Verhältnisskala, die im Folgenden kurz beschrieben werden.
Zur Nominalskala zählen Variablen, wenn die einzelnen Ausprägungen ausschließlich zur Klassifizierung beitragen. Die Ausprägungen haben keine natürliche Reihenfolge. Beispiele für nominalskalierte Variablen sind Haarfarbe, Automarke oder Religion.
Das nächsthöhere Messniveau stellt die Ordinalskala dar. Auf dieser Skala gibt es eine Rangordnung zwischen den Ausprägungen. Die Abstände zwischen den Ausprägungen sind allerdings nicht quantifizierbar. Ordinalskalierte Variablen sind zum Beispiel Schulnoten oder Rangplätze einer Fußballliga.
Auf der Intervallskala kann der Abstand zwischen den Merkmalsausprägungen im Unterschied zur Ordinalskala gemessen werden. Ein typisches Beispiel für eine intervallskalierte Variable ist die Messung der Temperatur in Grad Celsius. Es kann eine Größer-als-, allerdings keine Doppelt-so-groß-wie-Aussage getroffen werden. Diese Quantifizierung ist erst auf der Verhältnis- oder auch Ratioskala möglich.
Verhältnisskalierte Variablen haben einen natürlichen Nullpunkt, wie etwa beim Messen der Temperatur in Grad Kelvin. Weitere Beispiele wären Län-

ge, Gewicht, Alter und Einkommen.
Variablen, die auf der Intervall- oder Verhältnisskala gemessen werden, bezeichnet man als metrische Variablen. Nominal- oder ordinalskalierte Variablen nennt man kategorial. Zu den kategorialen Variablen zählen auch metrische Variablen, die in Kategorien eingeteilt werden. Eine binäre Variable ist eine kategoriale Variable, die nur zwei verschiedene Werte annehmen kann. Diese werden typischerweise mit $0-1$ codiert. Man kann natürlich auch Namen vergeben, wie „männlich"–„weiblich" oder „krank"–„gesund".

Für die verschiedenen Typen von Variablen können nun Maßzahlen berechnet werden, welche die vorliegenden Daten $X_1, \ldots, X_n$ charakterisieren sollen. Zuerst werden zwei Gruppen behandelt, nämlich Lage- und Streuungsmaße. Zu den Lagemaßen zählt das arithmetische Mittel $\bar{X}$, das auch oft nur als Mittelwert bezeichnet wird, und der Median $\tilde{X}$. Ersteres kann nur für metrische Variablen berechnet werden und ist definiert durch:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1}$$

Werden die Daten der Größe nach geordnet, so ist der Median jener Wert, der in der Mitte liegt. Er kann auch für ordinalskalierte Merkmale berechnet werden. Als Modus wird jener Wert bezeichnet, der am häufigsten vorkommt. Er ist für alle Skalenniveaus definiert.
Ein Maß für die Streuung der Daten ist die Spannweite. Sie ist die Differenz von kleinstem und größtem Wert. Da eine Größer-als-Beziehung vorausgesetzt ist, kann sie für ordinal-, intervall- und verhältnisskalierte Variablen berechnet werden. Sie ist zwar sehr einfach zu berechnen, hat aber den Nachteil, dass sie nur zwei Werte zur Berechnung benötigt und die restlichen Werte gar nicht berücksichtigt. Eine Maßzahl, die alle Beobachtungen in der Berechnung benötigt, ist die Stichprobenvarianz:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2) \tag{2}$$

Die Quadratwurzel dieser Größe nennt man die Standardabweichung der Stichprobe, man bezeichnet sie mit s.

Quantile gewähren Rückschlüsse auf die empirische Verteilung der Daten, besonders darauf, wo der Großteil der Verteilungsmasse liegt. Das α-Quantil ist jener Wert, unter dem α % der Daten liegen. Der Median kann somit als das 50%-Quantil interpretiert werden. Bei symmetrischen Verteilungen stimmen

Median und arithmetisches Mittel ungefähr überein. Liegt der Median links des Mittelwerts, so nennt man die Verteilung rechtsschief bzw. linkssteil. Der Großteil der Datenmasse liegt hierbei auf der linken Seite. Umgekehrt gilt dies für linksschiefe bzw. rechtssteile Verteilungen. Die Schiefe ist definiert durch

$$m_3 = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{s} \right)^3, \tag{3}$$

wobei $\bar{X}$ das arithmetische Mittel ist und wie in Gleichung 1 berechnet wird. Ist der Wert der Schiefe kleiner als 0, so handelt es sich um eine linkssteile Verteilung. Eine symmetrische Verteilung weist einen Wert von 0 auf und eine rechtssteile Verteilung einen streng positiven Wert.

Nach dieser kurzen Einführung über grundlegende statistische Maßzahlen und deren Charakterisierung wird in Kapitel 2.2 zuerst die allgemeine Testsituation beschrieben. Dann wird noch konkret auf einige Tests, die im Rahmen der Auswertung benötigt werden, eingegangen. Im darauffolgenden Kapitel wird die Regressionsanalyse beschrieben. Man beschränkt sich auf verallgemeinerte lineare Modelle, auf deren Modellierung und auf die Schätzung und Interpretation der Parameter.
Im Anschluss (Kapitel 2.4) werden zwei Methoden für Gruppenvergleiche vorgestellt. Die erste Methode ist die Varianzanalyse, die unter der Voraussetzung der Unabhängigkeit, der Normalverteilung und der Varianzhomogenität die Mittelwerte mehrerer Gruppen vergleicht. Eine Alternative bei Verletzung einer der Voraussetzungen ist der Kruskal-Wallis-Test.
Abschließend wird erläutert, wie man mit fehlenden Werten in der Analyse umgehen kann. Im Zuge dessen werden drei Imputationsverfahren vorgestellt.

2.2 Hypothesentests

Ein statistischer Test dient dazu, vorliegende Hypothesen auf ihre Richtigkeit zu überprüfen. Im Grunde unterscheidet man zwei Arten von Tests: zum einen die Parametertests, zum anderen die Verteilungstests. Bei Parametertests ist man an bestimmten Parametern in der Grundgesamtheit interessiert, wie etwa dem Mittelwert oder einem Anteil. Verteilungstests dienen der Entscheidung, ob die empirische Verteilung einer vorliegenden Stichprobe einer theoretischen Verteilung genügt.
Um eine Entscheidung zu treffen, muss man eine Entscheidungsregel definieren. Will man etwa überprüfen, ob ein Parameter gleich einem vorher

festgelegten Wert ist, so wird man die Nullhypothese annehmen, wenn die Schätzung des Parameters nicht „allzu weit“ von dieser spezifizierten Größe abweicht. Die Bedeutung von „allzu weit“ muss aber noch konkretisiert werden.

In einer Testsituation kann man zwei Fehler begehen. In Wahrheit gilt die Hypothese, man entscheidet sich aber dagegen. Dies nennt man den Fehler 1. Art oder auch α-Fehler. Ein weiterer möglicher Fehler ist der β-Fehler oder auch Fehler 2. Art. In Wahrheit gilt die Hypothese nicht, man entscheidet sich aber dafür. Die Größe $1 - \beta$ wird die Power oder Güte eines Tests genannt. Sie gibt an, mit welcher Wahrscheinlichkeit eine wahre Alternativhypothese als solche erkannt wird. In der Theorie will man einen Test haben, der eine hohe Power hat, also mit großer Wahrscheinlichkeit eine falsche Nullhypothese auch entdeckt, und zwar bei möglichst kleinem Fehler 1.Art. Die beiden Fehler sind aber leider antagonistisch: Dies bedeutet, wenn der α-Fehler klein gehalten wird, so steigt der β-Fehler. Eine typische Wahl für die Wahrscheinlichkeit des Fehlers 1. Art ist $\alpha = 0{,}05$. Besonders in medizinischen Studien wird α, das sogenannte Signifikanzniveau, noch kleiner gewählt, meist bei $\alpha = 0{,}01$, da man sich nicht fälschlicherweise gegen die Hypothese entscheiden will.

Bei der Durchführung eines Tests geht man grob nach folgendem Schema vor:

1. Aufstellen der Null- und Alternativhypothese
2. Wahl des Signifikanzniveaus α, wobei $0 < \alpha < 1$
3. Festlegung einer geeigneten Teststatistik und Bestimmung der Verteilung dieser Teststatistik unter der Nullhypothese
4. Bestimmung des Ablehnbereichs, das heißt Bestimmung des kritischen Werts als das α-, $(1 - \alpha)$- oder $(1 - \frac{\alpha}{2})$-Quantil der entsprechenden Verteilung, je nachdem ob man eine ein- oder zweiseitige Alternative formuliert hat
5. Berechnung der Teststatistik
6. Entscheidung und Interpretation

Angenommen, $\boldsymbol{\theta}$ ist ein Vektor von Parametern der Länge k und die Null-

und Alternativhypothese lauten:

$$H_0 : r(\boldsymbol{\theta}) = \boldsymbol{\theta}_0$$
$$H_1 : r(\boldsymbol{\theta}) \neq \boldsymbol{\theta}_0$$

Die Funktion $r(\cdot)$ ist eine Funktion, die dem Parametervektor Restriktionen vorschreibt, und $\boldsymbol{\theta}_0$ ist ein k-Vektor.
Um diese Null- gegen die zweiseitige Alternativhypothese zu testen, werden in den folgenden beiden Abschnitten zwei Parametertests vorgestellt, die auf der Likelihood-Theorie basieren. Außerdem werden ihre Vor- und Nachteile kurz diskutiert. Im Anschluss werden der Levene-Test und der Bartlett-Test sowie der Kolmogorov-Smirnov-Test behandelt. Die ersten beiden testen, ob die Varianzen in mehr als zwei Gruppen homogen sind. Der Kolmogorov-Smirnov-Test ist ein Verteilungstest.

2.2.1 Wald-Test

Der Wald-Test geht von der Idee aus, dass $\hat{\boldsymbol{\theta}}_{ML}$, der unrestringierte Maximum-Likelihood-Schätzer des Parametervektors, die oben genannte Einschränkung $r(\cdot)$ erfüllt. Die Größe $r(\hat{\boldsymbol{\theta}}_{ML}) - \boldsymbol{\theta}_0$ sollte für jede Komponente des Vektors also nahe 0 liegen. Die Teststatistik des Wald-Tests (Greene, 2000, Seite 154) ist:

$$W = (r(\hat{\boldsymbol{\theta}}_{ML}) - \boldsymbol{\theta}_0)'(Var(r(\hat{\boldsymbol{\theta}}_{ML}) - \boldsymbol{\theta}_0))^{-1}(r(\hat{\boldsymbol{\theta}}_{ML}) - \boldsymbol{\theta}_0) \tag{4}$$

Die Verteilung der Teststatistik W ist unter der Nullhypothese für große Stichproben annähernd eine χ^2-Verteilung mit Δ Freiheitsgraden, wobei Δ die Anzahl der Restriktionen ist.
Beim Wald-Test muss der unrestringierte Maximum-Likelihood-Schätzer und dessen Varianz berechnet werden. Die Varianz wird üblicherweise über die Informationsmatrix geschätzt.
Die Fisher-Informationsmatrix $\boldsymbol{I}$ des Parametervektors $\boldsymbol{\theta}$ (siehe Harrell, 2001, Seite 186) ist eine symmetrische $k \times k$-Matrix, deren Eintrag in der i-ten Zeile und j-ten Spalte folgendermaßen definiert ist:

$$\boldsymbol{I}_{ij}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right], \tag{5}$$

wobei $L(\boldsymbol{\theta})$ die Likelihood-Funktion des Parametervektors $\boldsymbol{\theta}$ bezeichnet, auf die in 2.3.1.3 genauer eingegangen wird. Die Fisher-Informationsmatrix wird also als der negative Erwartungswert der zweiten partiellen Ableitungen der logarithmierten Likelihood-Funktion berechnet. In der Praxis setzt man in

die Informationsmatrix den Maximum-Likelihood-Schätzer ein und erhält dadurch eine gute Schätzung der Varianz (siehe Harrell, 2001, Seite 187):

$$\widehat{Var(\boldsymbol{\theta})} = \boldsymbol{I}(\hat{\boldsymbol{\theta}}_{ML})$$

2.2.2 Likelihood-Ratio-Test

Der Likelihood-Ratio-Test, auch Likelihood-Quotienten-Test genannt, vergleicht die Maximum-Likelihood-Schätzer des unrestringierten und des restringierten Modells. Er geht von der Idee aus, dass sich die Maximum-Likelihood-Schätzungen nicht allzu sehr voneinander unterscheiden, wenn der wahre Parameter die Restriktion erfüllt.
Der unrestringierte Maximum-Likelihood-Schätzer des Parameters $\boldsymbol{\theta}$ wird mit $\hat{\boldsymbol{\theta}}_U$, der restringierte mit $\hat{\boldsymbol{\theta}}_R$ bezeichnet. Der Quotient der Likelihood-Funktionen, ausgewertet an diesen beiden Schätzungen $L(\hat{\boldsymbol{\theta}}_R)/L(\hat{\boldsymbol{\theta}}_U)$, liegt zwischen den Werten 0 und 1, da beide Likelihoods positiv sind und der Zähler auf jeden Fall kleiner als der Nenner ist, da ein restringiertes Optimum niemals größer sein kann als ein unrestringiertes. Ein kleiner Wert des Quotienten weist darauf hin, dass die Restriktion nicht stimmt.

Die Teststatistik des Likelihood-Ratio-Tests ist aber nicht der Quotient der Likelihoods, sondern das (-2)-fache des natürlichen Logarithmus (siehe Greene, 2000, Seite 152):

$$\begin{aligned} LR &= -2 \ln \left[\frac{L(\hat{\boldsymbol{\theta}}_R)}{L(\hat{\boldsymbol{\theta}}_U)} \right] \\ &= -2[\ln L(\hat{\boldsymbol{\theta}}_R) - \ln L(\hat{\boldsymbol{\theta}}_U)] \end{aligned} \tag{6}$$

Die Teststatistik LR ist in großen Stichproben annähernd χ^2-verteilt mit Δ Freiheitsgraden, wobei Δ wiederum die Anzahl der Restriktionen ist.

Der Wald-Test und der Likelihood-Ratio-Test sind einander sehr ähnlich. In großen Stichproben, wenn also $n \to \infty$, führen sie theoretisch zum gleichen Ergebnis. Im Allgemeinen ist der Wald-Test leichter zu berechnen, da man nur die Likelihood des unrestringierten Parameters berechnen muss. Beim Likelihood-Ratio-Test werden beide Parameter, also zusätzlich zum unrestringierten auch der restringierte geschätzt (siehe Greene, 2000, Seite 152), was natürlich in einer komplizierteren Berechnung endet. Aber genau deshalb, weil er mit beiden Schätzungen am meisten Information aus den Daten einbezieht, ist er vom statistischen Standpunkt her der bessere Test. Der Wald-Test reagiert äußerst empfindlich, auf welche Weise ein Parameter in

das Modell eingebunden wird. Außerdem ist die Teststatistik unter der Nullhypothese bei kleinen Stichproben nicht mehr χ^2-verteilt und daher kann es zu falschen Entscheidungen kommen. Die Likelihood-Ratio-Teststatistik gibt auch in kleineren Stichproben noch sichere und stabile Ergebnisse (siehe Harrell, 2001, Seite 190).
Beide Tests eignen sich besonders zum Vergleich von hierarchisch genesteten Regressionsmodellen (siehe 2.3). Zwei Modelle sind genestet, wenn die Menge der Parameter des kleineren, das heißt restringierten Modells, eine echte Teilmenge der Menge der Parameter des größeren, das heißt unrestringierten Modells, ist. Das bedeutet, das komplexere Modell unterscheidet sich vom einfacheren Modell nur durch Ergänzung von einem oder mehreren Parametern. Die Nullhypothese lautet dann

H_0 : Das restringierte ist nicht schlechter als das unrestringierte Modell.

Angenommen, das unrestringierte und das restringierte Regressionsmodell lauten

$$C(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$
$$C(Y|X) = \beta_0 + \beta_3 X_3$$

Die Nullhypothese lautet dann formal

$$H_0 : \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Kann man die Nullhypothese nicht verwerfen, so kann man davon ausgehen, dass man in der Praxis das restringierte Modell verwenden kann, da es nicht schlechter ist als das unrestringierte. Anders gesagt: Wird die Nullhypothese nicht angenommen, so kann man davon ausgehen, dass die Variablen X_1 und X_2 einen signifikanten Einfluss auf die Größe $C(Y|X)$ haben.

2.2.3 Tests auf Gleichheit zweier Mittelwerte

2.2.3.1 T-Test

Dieser Test vergleicht die Mittelwerte zweier unverbundener Stichproben. X_{1_i} und X_{2_j} bezeichnen die Beobachtungen der beiden Stichproben, wobei $i = 1, \ldots, n_1$ und $j = 1, \ldots, n_2$. Der Test setzt voraus, dass beide Stichproben einer Normalverteilung mit gleicher Varianz $\sigma_1^2 = \sigma_2^2$, aber möglicherweise unterschiedlichen Erwartungswerten μ_1 bzw. μ_2 folgen. Die Nullhypothese des T-Tests lautet

$$H_0 : \mu_1 = \mu_2$$

Sind die Varianzen σ_1^2 und σ_2^2 der beiden Stichproben zum Beispiel aus vorigen Studien bekannt, dann handelt es sich streng genommen und einen Gauß-Test (siehe Bleymüller, Gehlert und Gülicher, 2002, Seite 110), da die Teststatistik

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{7}$$

unter der Nullhypothese einer Standardnormalverteilung folgt. $\bar{X}_1$ und $\bar{X}_2$ sind die arithmetischen Mittel der beiden Stichproben (vgl. Gleichung 1). Sind die Varianzen nicht bekannt, so müssen diese aus der Stichprobe geschätzt werden und die Teststatistik wird abgewandelt zu

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}}, \tag{8}$$

wobei s_1^2 und s_2^2 wie in Gleichung 2 geschätzt werden und vorausgesetzt wird, dass sie gleich sind. T unterliegt unter der Nullhypothese einer T-Verteilung[1] mit $n_1 + n_2 - 2$ Freiheitsgraden. Sind die Varianzen nicht homogen, so muss ein Welch-Test gerechnet werden. Hierbei wird bei der Berechnung der Teststatistik wie in Gleichung 7 vorgegangen, aber die Varianzen werden durch die Stichprobenvarianzen s_1 und s_2 ersetzt. Auch die Anzahl der Freiheitsgrade für die T-Verteilung wird etwas anders bestimmt (siehe Rinne, 1995, Seite 396).

2.2.3.2 Wilcoxon-Rangsummentest

Der Wilcoxon-Rangsummentest ist das nichtparametrische Pendant zum T-Test. Die Annahme der Normalverteilung schwächt der Wilcoxon-Test insofern ab, als er nur fordert, dass die Daten aus einer stetigen Verteilung stammen, die die gleiche Form haben. Er verwendet zur Berechnung der Teststatistik nicht die Originaldaten, sondern deren Ränge. Die beiden Stichproben werden gemeinsam betrachtet und der Größe nach geordnet. Dann werden den Datenpunkten ihre Ränge zugeordnet und die Rangsumme für beide Gruppen wird separat gebildet.
Angenommen, man hat zwei Stichproben aus einer stetig verteilten Grundgesamtheit mit jeweils n_1 bzw. n_2 Beobachtungen. Die Nullhypothese lautet:

H_0 : Die Stichproben unterscheiden sich hinsichtlich ihres Lageparameters.

[1] Eine T-Verteilung strebt mit steigender Anzahl an Freiheitsgraden gegen eine Standardnormalverteilung, deshalb kann man bei großen Stichproben anstatt eines T-Tests einen Gauß-Test rechnen.

Mit $X_{1_1}, \ldots, X_{1_{n_1}}$ werden die Beobachtungen aus Stichprobe 1 bezeichnet, die Beobachtungen aus der zweiten Stichprobe mit $X_{2_1}, \ldots, X_{2_{n_2}}$. Nun ist $X_1, X_2 \ldots, X_{n_1+n_2}$ eine Aufzählung aller beobachteten Werte, man speichert aber noch zusätzlich, zu welcher Gruppe die Beobachtung gehört. Diese Werte werden nun aufsteigend sortiert und es werden ihnen die Ränge zugeordnet. Sind zwei Beobachtungen gleich, wird der Rang durch entsprechende Rangmittel ersetzt. Mit $R_1, \ldots, R_{n_1}$ bezeichnet man die Ränge aus der ersten Stichprobe. Die Rangsumme W_1 ist dann also $W_1 = \sum_{i=1}^{n_1} R_i$. Mit $S_1, \ldots, S_{n_2}$ bezeichnet man die Ränge der zweiten Stichprobe, deren Rangsumme dann $W_2 = \sum_{j=1}^{n_2} S_j$ ist. Die Teststatistik des Wilcoxon-Rangsummentests (siehe Lehmann und D'Abrer, 1975) ist dann

$$W = W_1 \tag{9}$$

Die Verteilung dieser Teststatistik kann für kleine Stichproben exakt durch Kombinatorik ermittelt werden, für große Stichproben ist dies allerdings sehr aufwändig. Man kann sie dann aber durch eine Normalverteilung approximieren.

2.2.4 Tests auf Gleichheit der Varianzen

In der Varianzanalyse (Kapitel 2.4.1) müssen die Daten die Voraussetzung der Varianzhomogenität erfüllen. Zwei Tests, die die Nullhypothese der Varianzengleichheit überprüfen, sind der Levene-Test und der Bartlett-Test.

2.2.4.1 Levene-Test

Der Levene-Test geht auf Howard Levene zurück, der diesen Test in *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (1960) vorgestellt hat. Angenommen, es liegen n Beobachtungen vor, die k Gruppen mit jeweils n_i, $i = 1, \ldots, k$ Einheiten zugeordnet sind. Die j-te Beobachtung der i-ten Gruppe wird mit X_{ij} bezeichnet. Die Null- und Alternativhypothese sind

$$\begin{aligned} &H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 \\ &H_1 : \sigma_i^2 \neq \sigma_j^2 \text{ für mindestens ein Paar } i \neq j \ , \end{aligned} \tag{10}$$

wobei mit σ_i^2 die Varianz der i-ten Gruppe gemeint ist.

Die Levene-Teststatistik (siehe Levene, 1960) ist definiert durch:

$$L = \frac{(n-k)\sum_{i=1}^{k} n_i(\bar{Z}_{i.} - \bar{Z}_{..})^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Z_{ij} - \bar{Z}_{i.})^2} \tag{11}$$

Sie ist unter der Nullhypothese ungefähr F-verteilt mit $k-1$ und $n-k$ Freiheitsgraden. Dabei können die Z_{ij} folgendermaßen definiert sein:

$$Z_{ij} = |X_{ij} - \bar{X}_{i.}|, \text{ mit } \bar{X}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} \text{ als Mittelwert der } i\text{-ten Gruppe}$$

$$Z_{ij} = |X_{ij} - \tilde{X}_{i.}|, \text{ mit } \tilde{X}_{i.} \text{ als Median der } i\text{-ten Gruppe}$$

$$Z_{ij} = |X_{ij} - \check{X}_{i.}|, \text{ mit } \check{X}_{i.} \text{ als } \alpha\text{-getrimmtes Mittel}^2 \text{ der } i\text{-ten Gruppe}$$

Streng genommen betrachtet Levene nur die erste Definition. Die beiden robusteren Methoden gehen auf Brown und Forsythe (1974) zurück.

2.2.4.2 Bartlett-Test

Der Bartlett-Test überprüft wie der Levene-Test die Nullhypothese der Varianzhomogenität in mehr als zwei Gruppen (vgl. Gleichung 10).
Wie in 2.2.4.1 gibt es n Beobachtungen, die in k Gruppen zu je $n_i, i = 1, \ldots, k$ Einheiten zusammengefasst werden. X_{ij} bezeichnen die $j = 1, \ldots, n_i$ Beobachtungen in den k Gruppen und s_i^2 die Stichprobenvarianz in der i-ten Gruppe.
Die Teststatistik des Bartlett-Tests ist:

$$B = \frac{(n-k)\ln(s_p^2) - \sum_{i=1}^{k}(n_i - 1)\ln(s_i^2)}{1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\left(\frac{1}{n_i - 1}\right) - \frac{1}{n-k}\right)}, \tag{12}$$

wobei die Größe s_p^2 die gepoolte Stichprobenvarianz ist und folgendermaßen berechnet wird

$$s_p^2 = \frac{1}{n-k}\sum_{i=1}^{k}(n_i - 1)s_i^2$$

B folgt unter der Nullhypothese einer χ^2-Verteilung mit $k-1$ Freiheitsgraden.

Der Bartlett-Test sollte nur verwendet werden, wenn die zugrunde liegenden Daten einer Normalverteilung folgen (siehe Box, 1953). Bei einer Verletzung dieser Annahme kann er zu verzerrten Ergebnissen führen. In solch einem Fall sollte besser der Levene-Test durchgeführt werden. Von diesem wurde gezeigt, dass er relativ robust gegen die Verletzung der Normalverteilungsannahme ist (siehe Glass, 1966).

[2] Das α-getrimmte Mittel ist das arithmetische Mittel, wenn die α % kleinsten und α % größten Werte gestrichen werden.

2.2.5 Kolmogorov-Smirnov-Test

Wie schon in der Einleitung beschrieben, überprüft der Kolmogorov-Smirnov-Test, ob eine Stichprobe einer theoretischen Verteilung folgt. Die Verteilung kann sowohl stetig als auch diskret sein. Der Test geht auf die beiden sowjetischen Mathematiker Andrei Kolmogorov und Nikolai Smirnov zurück und kann auch schon bei kleinen Stichproben angewendet werden.
Die Teststatistik des Tests ist der größte Abstand, den die empirische Verteilung der Daten zur theoretischen Verteilung aufweist. Bezeichnet man mit F^e die empirische Verteilung und mit F^t die theoretische Verteilung, dann ist die Teststatistik gegeben durch

$$K = \sup_x |F^e(x) - F^t(x)| \tag{13}$$

Kolmogorov und Smirnov haben herausgefunden, dass die Verteilung der Teststatisik K nicht von der in der Nullhypothese spezifizierten Verteilung abhängt, sondern nur vom Stichprobenumfang, einen Beweis findet man in Feller (1948). Die Verteilung der Kolmogorov-Smirnov-Teststatistik liegt in tabellierter Form vor. Diese Tabelle findet man unter anderem in Birnbaum (1952). Für große Stichproben ($n > 100$) kann der kritische Wert c bei einem Signifikanzniveau von $\alpha = 0,05$ mit der Formel $c = 1,358/\sqrt{n}$ bzw. bei einem Signifikanzniveau von $\alpha = 0,01$ mit der Formel $c = 1,628/\sqrt{(n)}$ angenähert werden (siehe Birnbaum, 1952, Seite 427).
Die Parameter der theoretischen Verteilung sollten bekannt sein. Müssen sie aus der Stichprobe geschätzt werden, so ist der Test konservativ[3].

2.3 Regressionsanalyse

Bei der Regressionsanalyse handelt es sich um ein multivariates Analyseverfahren. Es wird der Zusammenhang zwischen einer abhängigen und einer oder mehreren unabhängigen Variablen untersucht. Im Unterschied zur Korrelationsanalyse, bei der nur festgestellt wird, ob ein Zusammenhang besteht, ist die Richtung des Zusammenhangs bei der Regressionsanalyse eindeutig, nämlich der Einfluss der unabhängigen Variablen auf die abhängige Variable. Die abhängige Variable wird auch Regressand oder erklärte Variable, die unabhängigen Variablen werden Regressoren oder erklärende Variablen genannt. Sie werden mit Y bzw. $X = (X_1, X_2, \cdots, X_k)$ bezeichnet. Die Einteilung der Variablen in abhängige und unabhängige Variablen muss bereits

[3]Ein Test wird konservativ genannt, wenn er eine Abweichung von der Nullhypothese nur in seltenen Fällen als signifikant ansieht, er also die Nullhypothese „eher“ annimmt.

vor der Analyse feststehen.

Die folgenden Abschnitte halten sich an das Buch von Harrell (2001). Wird ein anderes Buch oder ein Artikel benutzt, so steht dies explizit im Text.

2.3.1 Verallgemeinerte lineare Modelle

Verallgemeinerte lineare Modelle (generalized linear models, GLM) wurden von Nelder und Wedderburn (1972) eingeführt. Es handelt sich hierbei um Modelle, die die Regressionsfunktion an eine breite Klasse von Verteilungen anpassen. Außerdem fordern sie keine strikten Modellvoraussetzungen.
Der Vektor $\beta = (\beta_0, \beta_1, \cdots, \beta_k)$ fasst die Regressionskoeffizienten, auch Parameter genannt, zusammen. Der Parameter β_0 ist optional und wird Intercept genannt. Der Ausdruck $X\beta$ ist eine gewichtete Summe der X, also $X\beta = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$.
Ein verallgemeinertes lineares Modell wird eindeutig durch drei Komponenten charakterisiert:

- einer Verteilungsfunktion f
- dem linearen Prädiktor
- der Link-Funktion

Die Verteilungen, die man mit verallgemeinerten linearen Modellen behandeln kann, fasst man unter dem Begriff *Exponentialfamilie von Verteilungen* zusammen. Sie haben die Eigenschaft, dass sie in einer bestimmten Form angeschrieben werden können. Diese Verteilungen werden im folgenden Abschnitt kurz beschrieben.

2.3.1.1 Exponentialfamilie von Verteilungen

Zu einer Exponentialfamilie von Verteilungen gehören jene Verteilungen, deren Dichte- bzw. Wahrscheinlichkeitsfunktionen einer Beobachtung y sich in folgender Form darstellen lassen (McCullagh und Nelder, 1989, Seite 20):

$$f(y|\theta, \phi) = exp(\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right) \tag{14}$$

In dieser Darstellung sind $a(\phi), b(\theta)$ und $c(y, \phi)$ bekannte Funktionen. θ ist der natürliche Lageparameter und ϕ wird der Dispersionsparameter genannt. Die Momente der Verteilung hängen direkt mit diesen Funktionen und Parametern zusammen. Der Erwartungswert ist die erste Ableitung von $b(\theta)$ nach

θ, die Varianz ist die zweite Ableitung von $b(\theta)$ nach θ, gewichtet mit $a(\phi)$.

Bekannte Verteilungen, die zu dieser Verteilungsfamilie gehören, sind die Normal- und die Binomialverteilung.
Die Normalverteilung ist eine stetige Verteilung und geht auf Carl Friedrich Gauß zurück, weshalb sie auch Gauß-Verteilung genannt wird. Sie ist die wohl bekannteste und wichtigste statistische Verteilung. Unter gewissen Voraussetzungen können beliebige Verteilungen durch die Normalverteilung approximiert werden. Ihre Dichtefunktion mit den Parametern μ und σ^2 ist:

$$f(y|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right)$$

Diese Funktion lässt sich folgendermaßen umschreiben:

$$\begin{aligned} f(y|\mu,\sigma^2) &= \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 - ln(\sigma\sqrt{2\pi})\right) = \\ &= \exp\left(-\frac{y^2-2y\mu+\mu^2}{2\sigma^2} - \frac{1}{2}ln(2\pi\sigma^2)\right) = \\ &= \exp\left(\frac{y\mu-\frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\left(\frac{y^2}{\sigma^2} - ln(2\pi\sigma^2)\right)\right) \end{aligned}$$

Betrachtet man die Darstellung in Gleichung 14, so gilt $\theta = \mu$, $b(\theta) = \frac{\mu^2}{2}$, $a(\phi) = \sigma^2$ und $c(y,\phi) = -\frac{1}{2}\left(\frac{y^2}{\sigma^2} + ln(2\pi\sigma^2)\right)$. Die Normalverteilung bildet also eine Exponentialfamilie von Verteilungen. Ihre beiden Parameter stellen gleichzeitig die Momente dar, μ ist der Erwartungswert und σ^2 die Varianz.

Die Binomialverteilung ist eine diskrete Verteilung, die eine Folge von Bernoulli-Experimenten beschreibt. Ein Bernoulli-Experimemt ist ein Versuch, bei dem es nur zwei mögliche Ausgänge gibt. Die Erfolgswahrscheinlichkeit p in einem Bernoulli-Versuch ist konstant über alle Versuche, außerdem sind die Versuche voneinander unabhängig. Ein typisches Beispiel für ein Bernoulli-Experiment stellt der Münzwurf dar. Es gibt nur zwei mögliche Ausgänge, nämlich Kopf oder Zahl. Vorausgesetzt, die Münze ist fair, haben beide Ausgänge die gleiche Eintrittswahrscheinlichkeit, nämlich $p = 0,5$. Diese Wahrscheinlichkeit ist konstant, sie gilt für den ersten wie auch für den zehnten Wurf. Außerdem sind die Versuche untereinander unabhängig. Der Ausgang des ersten Versuchs hat keinen Einfluss auf den Ausgang des zweiten Versuchs.

Die Wahrscheinlichkeitsfunktion der Binomialverteilung mit den Parametern n und p, wobei n die Anzahl der Wiederholungen des Bernoulli-Experiments und p die Wahrscheinlichkeit des positiven Ausgangs ist, lautet:

$$\begin{aligned} f(y|n,p) &= \binom{n}{y} p^y (1-p)^{n-y} = \\ &= \exp\left(\ln\binom{n}{y} + y\ln(p) + (n-y)\ln(1-p)\right) = \\ &= \exp\left(\ln\binom{n}{y} + y(\ln(p) - \ln(1-p)) + n\ln(1-p)\right) = \\ &= \exp\left(y\ln\left(\frac{p}{1-p}\right) + n\ln(1-p) + \ln\binom{n}{y}\right), \end{aligned}$$

Setzt man in der Darstellung der letzten Zeile $\ln\left(\frac{p}{1-p}\right) = \theta$, $-n\ln(1-p) = b(\theta)$, $a(\phi) = 1$ und $\ln\binom{n}{y} = c(y,\phi)$, so erkennt man, dass auch die Binomialverteilung zur Exponentialfamilie der Verteilungen zählt. Ihr Erwartungswert ist np und ihre Varianz $np(1-p)$.

Weitere Verteilungen dieser Klasse sind die Poisson- und die negative Binomialverteilung als diskrete Verteilungen und die Gamma- und die Inverse-Gauß-Verteilung als stetige Verteilungen. Die Exponentialverteilung gehört als Spezialfall der Gammaverteilung natürlich auch dazu.

2.3.1.2 Modellformulierung

Die Regressionsfunktion ist jene Funktion, die die Abhängigkeit der Regressoren zum Regressand herstellt. Mathematisch sieht diese Beziehung folgendermaßen aus:

$$C(Y|X) = g(X\beta).$$

In dieser Arbeit werden nur lineare Beziehungen behandelt. Die Linearität wird im Allgemeinen erst nach einer Transformation erreicht

$$h(C(Y|X)) = X\beta,$$

wobei $h(u) = g^{-1}(u)$, also die inverse Funktion[4] ist. Die Funktion h wird Link-Funktion genannt. Im Folgenden wird anstelle von $h(C(Y|X))$ auch $C'(Y|X)$

[4]Die Inverse einer Funktion f ist jene Funktion f^{-1}, für die gilt: $f(f^{-1}) = f^{-1}(f) = id_A$, wobei id_A die Identitätsfunktion bezeichnet. Dies ist jene Funktion, die ihr Argument selbst zurückgibt: $id_A : A \mapsto A$ mit $id_A(x) = x$

Tab. 2.1: Kanonischer Link für einige Verteilungen

Verteilung	*kanonischer Link*	*Bezeichnung*
Normal	$\theta = \mu$	Identitätslink
Binomial	$\theta = \ln\left(\frac{p}{1-p}\right)$	Logit-Link
Poisson	$\theta = \ln(\mu)$	Log-Link
Gamma	$\theta = \mu^{-1}$	reziproker Link
Inverse Gauß	$\theta = \mu^{-2}$	reziproker Link mit Power 2

geschrieben. Meist wird der Strich auch weggelassen und man schreibt nur $C(Y|X)$. Somit bezeichnet man mit $C(Y|X)$ die Funktion, die eine lineare Beziehung zwischen $Y|X$ und X herstellt.
Die Wahl der Link-Funktion in der Regression hängt von der Art der Verteilung ab. Für Verteilungen aus der Exponentialfamilie gibt es einen sogenannten natürlichen bzw. kanonischen Link. Dieser Link setzt den natürlichen Parameter der Exponentialfamilie direkt mit dem linearen Prädiktor in Zusammenhang:

$$\theta = X\beta$$

In Tabelle 2.1 ist der kanonische Link für einige Verteilungen aufgelistet (siehe McCullagh und Nelder, 1989, Seite 24):
Manchmal verwendet man aber trotz mathematisch angenehmer Eigenschaften nicht den kanonischen Link. Für die logistische Regression werden im Kapitel 2.3.3 auch zwei nicht-kanonische Link-Funktionen kurz diskutiert.

2.3.1.3 Schätzung der Parameter

Die Maximum-Likelihood-Methode ist ein übliches Verfahren, um Parameter auch in Nichtstandard-Situationen zu schätzen. Sie beruht auf der Likelihood-Funktion und sucht jene Parameter, welche die Likelihood-Funktion maximieren. Mit anderen Worten sucht sie jene Parameter, die die beobachteten Daten am wahrscheinlichsten machen.
Angenommen, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ ist ein Vektor von unbekannten Parametern. Y ist eine Zufallsvariable, deren Dichtefunktion f von den unbekannten Parametern θ abhängt. Weiters gibt es eine Zufallsstichprobe der Größe n. Die Likelihood-Funktion ist definiert als die gemeinsame Verteilungsfunktion der Beobachtungen $Y_1, \ldots, Y_n$:

$$L(Y_1, \ldots, Y_n|\boldsymbol{\theta}) = f(Y_1, \ldots, Y_n|\boldsymbol{\theta}) \tag{15}$$

Liegen n unabhängige Beobachtungen vor und bezeichnet man mit $f_i(Y_i|\boldsymbol{\theta})$ die Dichte der i-ten Beobachtung, dann ist die Likelihood-Funktion das Produkt der Randdichten:

$$L(Y_1, \ldots, Y_n|\boldsymbol{\theta}) = \prod_{i=1}^{n} f_i(Y_i|\boldsymbol{\theta}) \tag{16}$$

Der Maximum-Likelihood-Schätzer $\hat{\boldsymbol{\theta}}_{ML}$ ist nun jener Vektor, der die Likelihood-Funktion maximiert. Für die Berechnung der Maximalstelle ist es nicht von Bedeutung, ob man die Funktion selbst oder die logarithmierte Funktion betrachtet, da die Logarithmus-Funktion eine monotone Transformation ist (Collet, 1991, Seite 50). Da die logarithmierte Funktion wünschenswerte Eigenschaften beim Maximieren hat, betrachtet man die Log-Likelihood:

$$l(Y_1, \ldots, Y_n|\boldsymbol{\theta}) = \log(L(Y_1, \ldots, Y_n|\boldsymbol{\theta})) \tag{17}$$

Für unabhängige Beobachtungen ist die Log-Likelihood dann nicht mehr der Logarithmus des Produkts der Randdichten, sondern die Summe der logarithmierten Randdichten:

$$l(Y_1, \ldots, Y_n|\boldsymbol{\theta}) = \log \prod_{i=1}^{n} f_i(Y_i|\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f_i(Y_i|\boldsymbol{\theta}) \tag{18}$$

Die Maximalstelle der Log-Likelihood-Funktion wird gefunden, indem man die Ableitungen der Log-Likelihood-Funktion gleich null setzt und das Gleichungssystem löst. Meist muss die Lösung aber durch *Versuch und Irrtum* gefunden werden. Dies ist eine heuristische Methode, die verschiedene Lösungsmöglichkeiten durchprobiert, bis schließlich die richtige Lösung gefunden wird.

Im Spezialfall des einfachen linearen Modells (siehe 2.3.2) wird noch eine weitere Technik beschrieben, nämlich die Kleinst-Quadrate-Methode (Ordinary Least Squares, OLS).

2.3.1.4 Interpretation der Parameter

In allgemeinen Regressionsmodellen beschreibt β_j die Änderung der Einheiten von $C(Y|X)$ pro Änderung der Einheiten von X_j auf $X_j + 1$, wenn die restlichen unabhängigen Variablen gleich bleiben:

$$\beta_j = C(Y|X_1, \ldots, X_j + 1, \ldots X_k) - C(Y|X_1, \ldots, X_j, \ldots X_k) \tag{19}$$

Die Interpretation der Parameter in speziellen Modellen, wie zum Beispiel in der einfachen linearen oder in der logistischen Regression, wird direkt in den Kapiteln 2.3.2 und 2.3.3 behandelt. Die Interpretation bei der Berücksichtigung von Wechselwirkungen wird im folgenden Abschnitt beschrieben.

2.3.1.5 Wechselwirkungen

In einigen Fällen kann der Effekt einer Variablen nicht separat betrachtet werden. Der Einfluss einer Variablen X_i auf $C(Y|X)$ hängt von einer weiteren Variablen X_j ab. Diese beiden Variablen stehen in einer Wechselwirkung zueinander. Eine einfache Lösung, dies in einem Regressionsmodell zu berücksichtigen, ist das Hinzufügen der konstruierten Variablen X_iX_j. Im Modell muss also ein zusätzlicher Parameter β_{k+1} geschätzt werden. Die Änderung in der Regressionsfunktion bei einer Erhöhung von X_i auf $X_i + 1$, wenn alle anderen Variablen, inklusive X_j, konstant gehalten werden, hängt nun nicht mehr allein von β_i, sondern zusätzlich von X_j ab:

$$
\begin{aligned}
&C(Y|X_1, \ldots, X_i + 1, \ldots, X_j, \ldots, X_k) - C(Y|X_1, \ldots, X_i, \ldots, X_j, \ldots, X_k) = \\
&= (\beta_0 + \beta_1 X_1 + \ldots + \beta_i(X_i + 1) + \ldots + \beta_j X_j + \ldots + \beta_{k+1}(X_i + 1)X_j) - \\
&-(\beta_0 + \beta_1 X_1 + \ldots + \beta_i X_i + \ldots + \beta_j X_j + \ldots + \beta_{k+1} X_i X_j) = \\
&= \beta_i + \beta_{k+1} X_j
\end{aligned}
$$

Um zu prüfen, ob eine Wechselwirkung tatsächlich besteht, testet man die Hypothese

$$H_0 : \beta_{k+1} = 0$$

gegen die Alternative

$$H_1 : \beta_{k+1} \neq 0$$

Kann die Nullhypothese verworfen werden, so besteht eine Wechselwirkung zwischen den Variablen X_i und X_j.

Es können auch mehr als zwei Variablen in Wechselwirkungen zueinander stehen. Man fügt zum Beispiel bei einer dreifachen Wechselwirkung zwischen den Variablen X_i, X_j und X_l die künstliche Variable $X_iX_jX_l$ ein. Die Interpretation wird dann natürlich um einiges komplizierter.
Betrachet man ein Modell mit Wechselwirkungen, müssen die Haupteffekte, also die Variablen an sich, auf jeden Fall ins Modell aufgenommen werden, da die Interpretation der Parameter sonst keinen Sinn ergäbe.

2.3.1.6 Nominale Variablen in der Regression

Wenn man den Einfluss einer nominalen Variablen auf eine abhängige Variable modellieren will, muss man im Modell binäre Dummy-Variablen einführen. Angenommen, eine nominale Variable T kann k verschiedene Werte annehmen, nämlich $A_1, A_2, \ldots A_k$. Die Merkmalsausprägungen stehen in keiner Beziehung zueinander, sie können also nicht gereiht werden. Diese Variable kann vollständig durch $k-1$ Dummy-Variablen, also Variablen, die nur die Werte 0 oder 1 annehmen können, beschrieben werden. Das Modell lässt sich folgendermaßen schreiben:

$$\begin{aligned} C(Y|X = A_1) &= \beta_0 \\ C(Y|X = A_2) &= \beta_0 + \beta_1 \\ &\vdots \\ C(Y|X = A_k) &= \beta_0 + \beta_1 + \cdots + \beta_{k-1} \end{aligned}$$

Die Regressionsparameter $\beta_0, \ldots, \beta_{k-1}$ beschreiben die k möglichen Ausgänge, wobei β_0 die Referenzkategorie, in diesem Fall A_1, kennzeichnet.
Das Modell kann auch folgendermaßen formuliert werden:

$$C(Y|X) = X\beta = \beta_0 + \beta_1 X_1 + \cdots + \beta_{k-1} X_{k-1},$$

wobei

$$\begin{aligned} X_1 &= \begin{cases} 1 & \text{wenn } T = A_2 \\ 0 & \text{sonst} \end{cases} \\ X_2 &= \begin{cases} 1 & \text{wenn } T = A_3 \\ 0 & \text{sonst} \end{cases} \\ &\vdots \\ X_{k-1} &= \begin{cases} 1 & \text{wenn } T = A_k \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Für die Referenzkategorie A_1 sind alle $X_i = 0$, $i = 1, \cdots, k-1$ und daher $C(Y|X) = \beta_0$.

Will man nun testen, ob es einen Unterschied zwischen den Merkmalsausprägungen gibt, testet man die Nullhypothese

$$H_0 : \beta_0 = \beta_2 = \ldots = \beta_{k-1} = 0$$

Dieses Modell wird auch varianzanalytisches Modell genannt. Detaillierte Informationen zu diesem Modell findet man in Kapitel 2.4.1.

2.3.1.7 Nichtlineare Terme

In einigen Fällen hat eine unabhängige Variable keinen linearen Einfluss auf die abhängige Variable. Angenommen, das Regressionsmodell schaut folgendermaßen aus:

$$C(Y|X) = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k$$

Diese lineare Funktion beschreibt den Einfluss von $X_1, \ldots, X_k$ nur ungenügend, wie man bei der Validierung des Modells bemerkt hat. Eine mögliche Ursache kann ein nichtspezifizierter, nichtlinearer Einfluss einer Variablen sein. Nimmt man zum Beispiel an, dass die Variable X_1 einen quadratischen Einfluss hat, so fügt man die konstruierte Variable X_i^2 dem Modell hinzu:

$$C(Y|X) = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \beta_{k+1} X_1^2$$

Dieses Modell ist, auch wenn es den quadratischen Term X_1^2 enthält, ein lineares Regressionsmodell, da es linear in den Parametern β_i ist.
Um zu überprüfen, ob die Variable X_1 nun einen quadratischen oder lediglich einen linearen Einfluss hat, testet man die Hypothese

$$H_0 : \beta_{k+1} = 0.$$

Natürlich können auch höhere Potenzen von X_1 ins Modell aufgenommen werden. Man muss sich aber im Klaren sein, dass das Modell mit wachsender Anzahl an Parametern schwieriger zu interpretieren wird.

2.3.1.8 Spline-Funktionen

Eine weitere Möglichkeit, nichtlineare Terme in der Regression zu berücksichtigen, bieten Splines. Splines sind stückweise Polynome vom Grad $n \geq 1$, deren Funktionswerte und deren erste $n-1$ Ableitungen an vorher festgelegten Stützstellen übereinstimmen. Ist das Polynom vom Grad 1, so handelt es sich um stückweise lineare Funktionen.

Angenommen, in einem Modell gibt es nur eine erklärende Variable X. Die k Stützstellen befinden sich an den Punkten $s_1, \ldots, s_k$. Es sind $k+2$ Parameter (inklusive Intercept) nötig, um die Funktion zu modellieren. Die lineare Spline-Funktion ist durch die Formel

$$f(X) = \beta_0 + \beta_1 X + \beta_2 (X - s_1)_+ + \beta_3 (X - s_2)_+ + \ldots + \beta_{k+1} (X - s_k)_+ \quad (20)$$

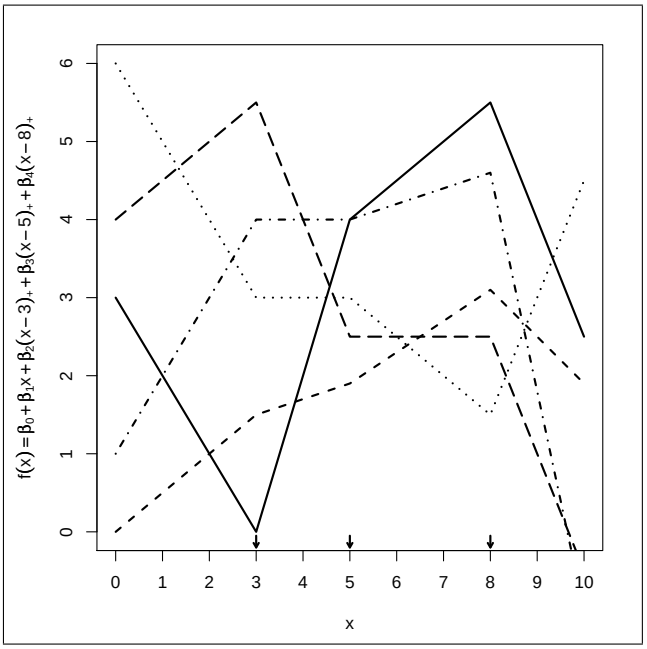


Abb. 2.1: Lineare Spline-Funktionen mit Stützstellen $s_1 = 3$, $s_2 = 5$ und $s_3 = 8$ (gekennzeichnet durch Pfeile)

gegeben, wobei $(c)_+ = \max(c, 0)$ bedeutet. Diese Spline-Funktion ist linear in den Parametern β_i, $i = 0, 1, \ldots, k+1$. Um globale Linearität zu überprüfen, kann die Nullhypothese

$$H_0 : \beta_2 = \ldots = \beta_{k+1} = 0$$

getestet werden. Ein Beispiel, wie so eine Funktion ausschaut, ist in Abbildung 2.1 dargestellt.
Lineare Spline-Funktionen können aber komplexere Strukturen in den Variablen oft nicht annähern. Deshalb sind gelegentlich Polynome höheren Grades notwendig.
Kubische Splines sind stückweise Polynome dritten Grades, deren Funktionswerte und deren erste und zweite Ableitungen[5] an den Stützstellen übereinstimmen. Solch eine Funktion mit Knoten an den Stellen $s_1, \ldots, s_k$ ist durch

$$f(X) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \beta_4 (X - s_1)_+^3 + \ldots + \beta_{k+3} (X - s_k)_+^3 \quad (21)$$

[5]Die Ableitung einer Funktion $f(x)$ an einer Stelle x ist die infinitesimale Änderung der Funktion im Punkt x.

gegeben, wobei wiederum $(c)_+ = \max(c, 0)$ ist. Bei k Stützstellen müssen demnach $k + 4$ Parameter (inklusive Intercept) geschätzt werden.

Es wurde gezeigt (vgl Stone und Koo, 1985), dass sich kubische Splines an den Enden, also vor dem ersten und nach dem letzten Knoten, relativ unkoordiniert verhalten können. Eine Methode, um dem entgegenzuwirken, ist die Wahl von restringierten kubischen Splines. Diese werden gezwungen, dass sie in den äußeren Bereichen linear sind. Ein weiterer Vorteil ist, dass nur k Parameter (inklusive Intercept) geschätzt werden müssen. Die restringierte kubische Spline-Funktion mit k Knoten an den Stellen $s_1, \ldots, s_k$ hat die Form

$$f(X) = \beta_0 + \beta_1 X + \beta_2 X_2 + \ldots + \beta_{k-1} X_{k-1} \tag{22}$$

mit $X_1 = X$ und für $j = 1, \ldots, k-2$ gilt

$$X_{j+1} = (X - s_j)_+^3 - \frac{(X - s_{k-1})_+^3 (s_k - s_j)}{s_k - s_{k-1}} + \frac{(X - s_k)_+^3 (s_{k-1} - s_j)}{s_k - s_{k-1}}$$

Man kann zeigen, dass diese Funktion vor dem Knoten s_1 und nach dem Knoten s_k linear in X ist.
Beispiele für restringierte kubische Spline-Funktionen mit gleichen Stützstellen, aber verschiedenen Parametern zeigt Abbildung 2.2.

Bis jetzt wurde angenommen, dass die Stützstellen vor der Analyse feststehen. Wenn man im Vorhinein schon weiß, dass die Variable X ihre Krümmung an der Stelle s ändert, so kann man s als eine Stützstelle wählen. So ein Vorwissen ist aber im Allgemeinen nicht üblich. In solch einem Fall können die Positionen der Knoten auch als Parameter in das Modell aufgenommen werden. Das Modell ist dann aber nicht mehr linear und bringt somit alle Nachteile einer nichtlinearen Regression mit sich. Insbesondere kann nicht mehr auf Standard-Software zurückgegriffen werden.
Eine in der Praxis beliebtere Methode ist die Wahl der Stützstellen an fixen Quantilen. Stone hat herausgefunden, dass die Wahl der Position der Stützstellen in restringierten kubischen Splines nicht ausschlaggebend ist. Wichtiger ist die Wahl der Knotenanzahl. Tabelle 2.2 zeigt die von Harrell (siehe Harrell, 2001, Seite 23) empfohlenen fixen Quantile bei unterschiedlicher Anzahl an Stützstellen.
Man kann die Funktion dann nur noch modifizieren, indem man die Anzahl der Stützstellen ändert. Für kleine Stichproben nimmt man üblicherweise $k = 3$ Stützstellen, bei größeren Datenmengen kann man auch $k = 7$ wählen. Stone (siehe Harrell, 2001, Seite 23) hat gezeigt, dass sich die Wahl von k nur

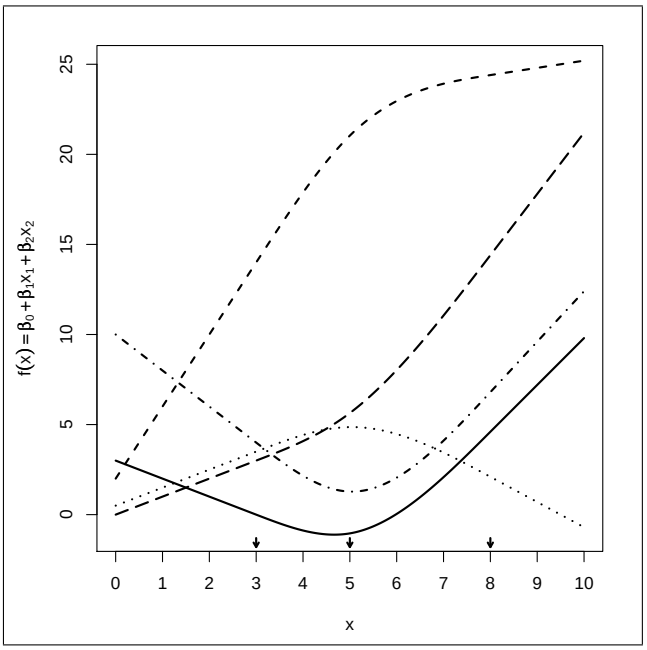


Abb. 2.2: Restringierte kubische Spline-Funktionen mit Stützstellen $s_1 = 3$, $s_2 = 5$ und $s_3 = 8$ (gekennzeichnet durch Pfeile)

Tab. 2.2: Empfohlene Quantile bei unterschiedlicher Anzahl an Stützstellen

k	*Quantile*						
3			0,10	0,50	0,90		
4			0,05	0,35	0,65	0,95	
5		0,05	0,275	0,50	0,725	0,95	
6	0,05	0,23	0,41	0,59	0,77	0,95	
7	0,025	0,1833	0,3417	0,50	0,6583	0,8167	0,975

auf $k = 3$, $k = 4$ und $k = 5$ zu beschränken braucht. Mehr als 5 Stützstellen sind in den seltensten Fällen notwendig. In den meisten Datensätzen ist die Wahl von $k = 4$ ein guter Mittelweg zwischen Flexibilität der Funktion und Vermeidung einer Überanpassung wegen zu vieler Parameter.

2.3.1.9 Variablenselektion

Welche Variablen in ein Modell aufgenommen werden, muss bereits vor der Analyse feststehen. Oft kann im Vorhinein aber nicht beurteilt werden, welche Variablen wichtig sind und Einfluss haben.
Eine einfache, jedoch äußerst aufwändige Methode ist das Ausprobieren aller möglichen Modelle, angefangen vom Nullmodell, das keine Variablen enthält, bis hin zum vollen, gesättigten Modell, das alle zur Verfügung stehenden Variablen beinhaltet. Diese Methode gibt dem Analysten die meiste Information über die Art des Zusammenhangs der abhängigen Variablen von den unabhängigen Variablen. Bei k Variablen gibt es 2^k mögliche Modelle, die untersucht werden müssen. Bei sechs erklärenden Variablen müsste man also $2^6 = 64$ Gleichungen anpassen, bei sieben Variablen steigt die Anzahl auf $2^7 = 128$. Einen Kompromiss zwischen vollständiger Kenntnis aller Modelle und rechnerischem sowie zeitlichem Aufwand bieten die sogenannten schrittweisen Prozeduren.
Diese Verfahren werden von vielen Statistikern allerdings nicht empfohlen, da sie viele Regeln des statistischen Schätzens und Testens verletzen. Sie führen im Allgemeinen zu erhöhten Werten des Bestimmtheitsmaßes R^2 (siehe 2.3.2.3), die $F-$ und χ^2-Teststatistiken (siehe 2.2) haben nicht mehr die geforderte Verteilung, Standardfehler der Regressionskoeffizienten werden zu klein und Konfidenzintervalle[6] zu eng geschätzt.

Da es nicht Ziel dieser Arbeit ist, ein Modell für Prognosen zu erstellen, sondern sich lediglich auf Hypothesentests zu beschränken, werden die Verfahren zur Variablenselektion nicht weiter behandelt. Bei Interesse kann man bei Fahrmeir und Tutz (1994, Seite 122 ff.) nachlesen.

2.3.2 Einfache lineare Regression

Die einfachste Form der Abhängigkeit in der Regression ist die lineare Abhängigkeit. Dies bedeutet, dass die unabhängige Variable durch die abhängigen Variablen mit Hilfe einer linearen Funktion $f(x)$ erklärt werden kann. Die

[6] Ein Konfidenzintervall ist ein zufälliger Bereich, in dem der Parameter mit einer vorgegebenen Wahrscheinlichkeit liegt. Man spricht von einem $(1-\alpha)$-Konfidenzintervall, wenn diese Wahrscheinlichkeit $1-\alpha$ sein soll.

Erläuterungen in diesem Kapitel gehen im Wesentlichen auf Eckey, Kosfeld und Dreger (2001, Kapitel 2.1 und 2.2) und Johnston (1984) zurück. Werden andere Quellen benutzt, so werden diese im Text explizit angegeben.
Formal lässt sich die Regression bei ungruppierten Daten folgendermaßen beschreiben: Es gibt n Beobachtungen einer abhängigen Variablen Y, die in einem Spaltenvektor der Länge n

$$Y = (y_1, y_2 \ldots, y_n)'$$

zusammengefasst werden. Die jeweils n Beobachtungen der k unabhängigen Variablen $X_1, X_2, \cdots, X_k$ werden in die Datenmatrix geschrieben:

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1k} \\ x_{21} & x_{22} & \ldots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nk} \end{pmatrix}$$

Die Werte von X werden als deterministisch, das heißt nicht zufällig, vorausgesetzt. $u = (u_1, u_2, \ldots, u_n)'$ ist ein Vektor der Länge n, der die unabhängigen, identisch verteilten Fehler in dem Modell zusammenfasst.
Der Vektor β ist der Vektor der unbekannten Parameter. Wird der optionale Intercept β_0 in das Modell aufgenommen, wird in der X-Matrix, die dann auch Designmatrix genannt wird, als erster Spaltenvektor der Vektor $(1, 1, \ldots, 1)^t$ eingefügt. Die Designmatrix hat dann die Dimension $n \times (k+1)$.

Die Voraussetzungen bzw. Annahmen der einfachen linearen Regression können in fünf Punkten zusammengefasst werden:

- $E(u) = 0$
- $E(uu') = \sigma^2 I_n$
- X ist nichtstochastisch
- rank$(X)= k$
- $u \sim N(0, \sigma^2)$

In der einfachen linearen Regression wird die abhängige Variable Y auf einem metrischen Skalenniveau beobachtet. Es wird angenommen, dass Y einer Normalverteilung mit Erwartungswert $X\beta$ und einer konstanten Varianz

σ^2 folgt. Dies ergibt sich direkt aus der letzten Annahme, dass die u_i normalverteilt sind. Das Regressionsmodell der einfachen linearen Regression in Matrixnotation schaut folgendermaßen aus:

$$\begin{aligned} Y =& X\beta + u \\ Y =& \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + u. \end{aligned} \tag{23}$$

Betrachtet man gruppierte Daten, so kann man die identischen Einträge zusammenfassen und muss angeben, wie oft diese Einträge beobachtet wurden. Angenommen, es gibt g verschiedene Einträge in den Variablen X, das heißt, es gibt g unterschiedliche Zeilen in der Daten- bzw. Designmatrix X. Zusätzlich zu den Vektoren Y und X muss also noch die Information $(n_1, \ldots, n_i, \ldots, n_g)'$ gespeichert werden.
Der Vektor Y beinhaltet nur noch den Mittelwert der vorher beobachteten Daten und hat die Dimension $g \times 1$. Ist Y binär, so wird meist nicht der Mittelwert, was ja dem Prozentsatz an Einsern entspricht, sondern die absolute Anzahl an Einsern angegeben (siehe Fahrmeir und Tutz, 1994, Seite 16 ff.).

$$Y = (\bar{Y}_1, \cdots, \bar{Y}_i, \cdots, \bar{Y}_g)'$$

Die Datenmatrix X besteht nur noch aus g Zeilen und nach wie vor k Spalten, da sich die Anzahl der Variablen nicht ändert:

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \ldots & x_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ x_{g1} & x_{g2} & \ldots & x_{gk} \end{pmatrix}$$

Der Fall ungruppierter Daten ist genau genommen ein Spezialfall von gruppierten Daten mit $n_1 = \cdots = n_i = \cdots = n_g = 1$. Man kann auch ungruppierte Daten gruppieren, um die Dimension der Daten zu reduzieren. Dies ist kein Problem, wenn nur kategoriale Variablen vorliegen. Bei metrischen Variablen ist es oft nur sehr schwer möglich, da es wenige identische Einträge geben wird.

Wie schon im Kapitel 2.3.1 angesprochen, wird jedes Modell durch drei Komponenten, nämlich einer Verteilungsfunktion, einem linearen Prädiktor und einer Link-Funktion beschrieben. Im Falle des einfachen linearen Modells ist die Verteilungsfunktion die Normalverteilung, der lineare Prädiktor ist der Erwartungswert ($C(Y|X) = E(Y|X) = \mu$) und die Link-Funktion ist die Identitätsfunktion. Die Identitätsfunktion ist in diesem Fall auch der kanonische Link (vgl. Tabelle 2.1).

2.3.2.1 Schätzung der Parameter

Eine weit verbreitete Schätzmethode ist die Kleinst-Quadrate-Methode (Ordinary Least Squares, OLS). Im Modell der Regressionsanalyse wird angenommen, dass zwischen den Werten $C(Y|X)$ und X Fehler auftreten, die im Regressionsmodell nicht berücksichtigt werden können. Bei der Kleinst-Quadrate-Methode sollen jene Parameter gefunden werden, welche die Quadratsumme dieser Fehler minimieren:

$$\hat{\beta}_{OLS} = \text{argmin} \left(\sum_{i=1}^{n} u_i^2 \right)$$

Schreibt man dies in Matrixnotation an und setzt das Modell ein, ergibt sich:

$$\begin{aligned} \hat{\beta}_{OLS} &= \text{argmin } (u'u) \\ \hat{\beta}_{OLS} &= \text{argmin } (Y - X\beta)'(Y - X\beta). \end{aligned} \tag{24}$$

Zur Lösung der Minimierungsaufgabe muss man die Funktion nach β ableiten und 0 setzen. Dies ergibt:

$$X'X\hat{\beta}_{OLS} = X'Y$$

Diese Gleichung wird oft Normalgleichung genannt. Hat die Matrix X vollen Rang[7], so ist die Matrix $X'X$ positiv definit und besitzt eine eindeutige Inverse[8]. Das Gleichungssystem kann dann eindeutig gelöst werden und der Kleinst-Quadrate-Schätzer ist gegeben durch:

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y \tag{25}$$

Man kann nachweisen, dass der Maximum-Likelihood-Schätzer $\hat{\beta}_{ML}$ in diesem Modell gleich dem Kleinst-Quadrate-Schätzer ist. Dies wird an der Stelle aber nicht durchgeführt. Bei Interesse kann man dies zum Beispiel bei Eckey, Kosfeld und Dreger (2001, Seite 57–58) nachlesen.
Der Kleinst-Quadrate-Schätzer spielt in der Theorie des einfachen linearen Modells eine große Rolle. Das Gauß-Markov-Theorem, benannt nach Carl Friedrich Gauß und Andrey Markov, besagt, dass dieser Schätzer **BLUE**

[7]Eine Matrix A hat dann vollen Rang, wenn keine Spalte als Linearkombination der restlichen Spalten geschrieben werden kann. Analog gilt dies für die Zeilen einer Matrix. Der Zeilenrang ist gleich dem Spaltenrang.

[8]Die Inverse einer Matrix A ist jene Matrix A^{-1}, die die Gleichung $AA^{-1} = I$ erfüllt, wobei $I = \text{diag}(1, \ldots, 1)$ die richtig dimensionierte Einheitsmatrix ist, die in ihrer Diagonale als Einträge 1 und sonst 0 hat.

(**B**est, **L**inear, **U**nbiased **E**stimator) ist. Er ist linear, da er nur eine gewichtete Summe der Y ist und die Gewichte $(X'X)^{-1}X'$ feste Zahlen sind. Er ist der beste Schätzer in dem Sinne, dass die Varianz des Kleinst-Quadrate-Schätzers kleiner oder gleich der Varianz eines beliebigen anderen linearen Schätzers $\tilde{\beta}$ ist. Die Gleichheit gilt nur für den Fall, wenn $\hat{\beta}_{OLS} = \tilde{\beta}$ gilt. Der Schätzer ist unverzerrt, manchmal sagt man auch erwartungstreu, da der Erwartungswert des Schätzers gleich dem Parameter ist. Kurz zusammengefasst:

- $E(\hat{\beta}_{OLS}) = \beta$
- $\hat{\beta}_{OLS} = DY$, wobei $D = (X'X)^{-1}X'$ eine nicht zufällige $(k+1) \times n$ Matrix ist
- $Var(\hat{\beta}_{OLS}) \leq Var(\tilde{\beta})$ für jeden beliebigen, linearen Schätzer $\tilde{\beta}$

2.3.2.2 Interpretation der Parameter

Im einfachen linearen Regressionsmodell ist die Bedeutung der Parameter leicht verständlich. Da in der Darstellung der Gleichung 19 in Kapitel 2.3.1.4 $C(Y|X) = E(Y|X)$ und $E(Y|X) = Y$ gilt, gibt der Parameter β_j an, um wie viele Einheiten sich die Variable Y ändert, wenn sich die Variable X_j um eine Einheit ändert und alle restlichen Variablen gleich bleiben.

$$
\begin{aligned}
&E(Y|X_1, \ldots, X_j + 1, \ldots, X_k) - E(Y|X_1, \ldots, X_j, \ldots, X_k) = \\
&= (\beta_0 + \beta_1 X_1 + \ldots + \beta_j(X_j + 1) + \ldots + \beta_k X_k) - \\
&-(\beta_0 + \beta_1 X_1 + \ldots + \beta_j X_j + \ldots + \beta_k X_k) = \beta_j
\end{aligned}
$$

2.3.2.3 Modellanpassung

Wie gut ein lineares Regressionsmodell die Daten beschreibt, kann man anhand der Residuenanalyse bewerten. An den Residuen-Plots kann man erkennen, ob eventuell Heteroskedastizität vorliegt oder ob eine Variable einen quadratischen Einfluss hat. Heteroskedastizität bedeutet, dass die Varianz der u_i nicht für alle $i = 1, \ldots, n$ gleich ist, sondern:

$$Var(u_i) = \sigma_i^2 \ ,i = 1, \ldots, n$$

Da man annimmt, dass die Residuen normalverteilt sind, kann man die standardisierten Fehler betrachten. Diese sollten, wenn die Annahme der Normalverteilung stimmt, innerhalb des Intervalls $[-3, 3]$ liegen. Liegt ein Wert außerhalb des Intervalls, dann bezeichnet man ihn als Ausreißer. Entweder

handelt es sich wirklich um einen extremen und untypischen Wert oder es liegt ein Messfehler vor.
Betrachtet man die gefitteten Werte

$$\hat{Y} = X\hat{\beta}_{OLS} \tag{26}$$

und plottet sie gegen die Residuen, so sollten die Datenpunkte zufällig um null streuen. Wenn keine Struktur zu erkennen ist, dann kann man von einer guten Modellanpassung sprechen. Erkennt man jedoch eine U-Form, so lässt dies darauf schließen, dass ein bestehender quadratischer Einfluss nicht ins Modell genommen wurde. Ist die Struktur keilförmig, so liegt meist Heteroskedastizität vor (siehe Schlittgen, 2003, Kapitel 20).

Das Bestimmtheitsmaß R^2 gibt an, wie gut die Regression die Varianz in den Daten erklärt. Es ist der Quotient aus erklärter Varianz und gesamter Varianz. Die Varianz wird mit der Quadratsumme bestimmt:

$$\begin{aligned}
\text{TSS} &= \sum_{i=1}^{n}(Y_i - \bar{Y_i})^2 \\
\text{ESS} &= \sum_{i=1}^{n}(\hat{Y_i} - \bar{Y_i})^2 \\
\text{RSS} &= \sum_{i=1}^{n}(Y_i - \hat{Y_i})^2,
\end{aligned}$$

wobei $\bar{Y_i}$ den Mittelwert der Beobachtungen und $\hat{Y}$ die gefitteten Werte bezeichnen.
TSS (**T**otal **S**um of **S**quares) ist die Summe der Abweichungsquadrate von den einzelnen Datenpunkten zum Mittelwert, sie misst also die gesamte Variation in den Daten. ESS (**E**xplained **S**um of **S**quares) ist die Summe der Abweichungsquadrate der geschätzten Werte zum Mittelwert, also die Variation, die durch die Regression erklärt wird. RSS (**R**esidual **S**um of **S**quares) ist die Quadratsumme der Residuen und somit die Variation, die vom Modell nicht erklärt werden kann. Das Bestimmtheitsmaß ist definiert durch:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y_i} - \bar{Y_i})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y_i})^2} = \frac{\text{ESS}}{\text{TSS}} = 1 - \frac{\text{RSS}}{\text{TSS}} \tag{27}$$

Es ist zu beachten, dass R^2 nur für ein Modell mit Intercept auf diese Weise berechnet werden kann. Man kann zeigen, dass das Bestimmtheitsmaß mit dem quadrierten Korrelationskoeffizienten zwischen X und Y, also r_{XY},

übereinstimmt (siehe Schlittgen, 2003, Seite 442). Als quadrierten Korrelationskoeffizient kann man R^2 immer bestimmen.
Liegt R^2 nahe bei eins, so spricht man von einer guten Modellanpassung, da der Großteil der Varianz vom Modell erklärt wird. In der Praxis wird man aber ein R^2 von eins nie beobachten. Gute Werte liegen zwischen 0,5 und 0,8.

2.3.3 Logistische Regression

Im Modell der einfachen linearen Regression ist die Variable Y eine metrische, normalverteilte Variable. Oft ist die abhängige Variable aber nicht metrisch, sondern kategorial oder sogar nur binär, wie zum Beispiel überlebt „ja/nein“ oder Herzinfarkt „ja/nein“. Nimmt sie also nur die Werte 0 oder 1 an, so ist Y Bernoulli-verteilt mit der Erfolgswahrscheinlichkeit $\pi = P(Y = 1|X)$ und es gilt:

$$\begin{aligned} E(Y|X) &= \pi = P(Y = 1|X) \\ Var(Y|X) &= \pi(1-\pi) = P(Y = 1|X)(1 - P(Y = 1|X)) \end{aligned}$$

Liegen die Daten gruppiert vor und bezeichnet $\bar{Y}$ die relative Häufigkeit an Einsern von m unabhängigen Bernoulli-Beobachtungen, dann ist die Variable $m\bar{Y}$ binomialverteilt mit den Parametern m und π (siehe Fahrmeir und Tutz, 1994, Seite 24 f.). Sie nimmt die Werte $0, 1, 2 \ldots, m$ mit konstanter Erfolgswahrscheinlichkeit π an. Der Erwartungswert und die Varianz sind dann

$$\begin{aligned} E(m\bar{Y}|X) &= m\pi \\ Var(m\bar{Y}|X) &= m\pi(1-\pi) \end{aligned}$$

Die relative Häufigkeit $\bar{Y}$ folgt dann einer skalierten Binomialverteilung, die die Werte $0, 1/m, 2/m, \ldots, 1$ mit gleicher Erfolgswahrscheinlichkeit wie $m\bar{Y}$ annimmt. Der Erwartungswert und die Varianz sind dann

$$\begin{aligned} E(\bar{Y}|X) &= \pi \\ Var(\bar{Y}|X) &= \frac{\pi(1-\pi)}{m} \end{aligned}$$

Die Vorhersage des Modells kann als Vorhersage der Wahrscheinlichkeit, dass ein gewisses Ereignis eintritt, interpretiert werden. Es kann in einem Modell dann jedoch vorkommen, dass es Werte kleiner als 0 bzw. größer als 1 zulässt. Eine Wahrscheinlichkeit kann aber nicht außerhalb des Intervalls $[0, 1]$ liegen. Die Funktion $C(Y|X) = P(Y = 1|X)$ muss also so transformiert werden, dass die Regressionsfunktion nur Werte im Intervall $[0, 1]$ zulässt. Auf drei Link-Funktionen, die dies realisieren, wird im folgenden Kapitel eingegangen. Die

Beschreibung der Link-Funktionen folgt anhand des Buches Fahrmeir und Tutz (1994, Seite 25–27).

2.3.3.1 Link-Funktionen

Logit-Modell

In Kapitel 2.3.1.2 wurde schon angemerkt, dass der sogenannte Logit-Link der kanonische bzw. natürliche Link für die Binomialverteilung ist. Man setzt also den natürlichen Parameter θ (siehe Tabelle 2.1) direkt in Beziehung zum linearen Prädiktor $X\beta$:

$$\theta = \log\left(\frac{P(Y=1|X)}{1-P(Y=1|X)}\right) = X\beta \tag{28}$$

Die Odds sind zwar nach oben unbeschränkt, sie nehmen auch Werte größer als 1 an, doch nähern sie sich nach unten asymptotisch null an. Durch Transformation der Odds in die Logg-Odds, oder auch Logits genannt, ist der Wertebereich unbeschränkt.

$$\text{Logit}(Y=1|X) = \log(\text{Odds}(Y=1|X) = \log\left(\frac{P(Y=1|X)}{1-P(Y=1|X)}\right)$$

Setzt man die Logits in Beziehung zum linearen Prädiktor und formt die Gleichung um, so erhält man das logistische Regressionsmodell:

$$\begin{aligned}
\log\left(\frac{P(Y=1|X)}{1-P(Y=1|X)}\right) &= X\beta \\
\frac{(Y=1|X)}{1-P(Y=1|X)} &= \exp(X\beta) \\
P(Y=1|X) = \frac{\exp(X\beta)}{1+\exp(X\beta)} &= \frac{1}{1+\exp(-X\beta)}
\end{aligned} \tag{29}$$

Die Wahrscheinlichkeit $P(Y=1|X)$ wird durch diese Transformation auf den Wertebereich $[0,1]$ beschränkt.

Da die Verteilungsfunkton der logistischen Verteilung mit Parameter α und $\beta > 0$ folgende Form hat,

$$F(x) = \frac{1}{1+\exp(-\frac{x-\alpha}{\beta})}$$

kann man den Logit-Link auch als logistische Verteilungsfunktion mit den Parametern $\alpha = 0$ und $\beta = 1$ auffassen. Als Verteilungsfunktion hat die Link-Funktion natürlich die Eigenschaft, dass sie zwischen den Werten 0 und 1 liegt. Außerdem ist diese Verteilungsfunktion symmetrisch um ihren Erwartungswert $\alpha = 0$.
Wie schon in Kapitel 2.3.1.2 angesprochen, verwendet man gelegentlich auch nicht-kanonische Link-Funktionen, wovon zwei spezielle in den beiden nächsten Abschnitten behandelt werden.

Probit-Modell

Die Link-Funktion im Probit-Modell ist die Verteilungsfunktion der Standardnormalverteilung. Sie wird meist mit $\Phi(\cdot)$ bezeichnet.
Das Modell lässt sich folgendermaßen formulieren:

$$C(Y|X) = P(Y = 1|X) = \Phi(X\beta). \tag{30}$$

Die Standardnormalverteilung ist, wie die logistische, eine um ihren Erwartungswert $\mu = 0$ symmetrische Verteilung. Die logistische Verteilung hat aber schwerere Enden, was bedeutet, dass die Dichte mehr Masse an den Seiten hat. Die logistische Verteilungsfunktion wächst schneller als die Standardnormalverteilung, im Punkt 0 kreuzen sie sich, dann ist sie flacher, liegt also graphisch gesehen darunter. Da die Verteilungsfunktion der Standardnormalverteilung nicht als geschlossene Funktion angeschrieben werden kann, sondern nur in tabellierter Form vorliegt, ist es aufwändiger und rechnerisch anspruchsvoller als das Logit-Modell. Deshalb wird das Logit-Modell in der praktischen Anwendung oft bevorzugt.

Komplementäres Log-Log-Modell

Wie schon in den vorigen beiden Modellen fungiert auch in diesem Modell eine Verteilungsfunktion als Link-Funktion, nämlich die der Extremwertverteilung. Das logistische Regressionsmodell mit komplementärem Log-Log-Link schaut folgendermaßen aus:

$$C(Y|X) = P(Y = 1|X) = 1 - \exp(-\exp(X\beta)) \tag{31}$$

Formt man die Gleichung nach $X\beta$ um, so sieht man auch, woher die Link-Funktion ihren Namen hat:

$$X\beta = \log(-\log(1 - P(Y = 1|X)))$$

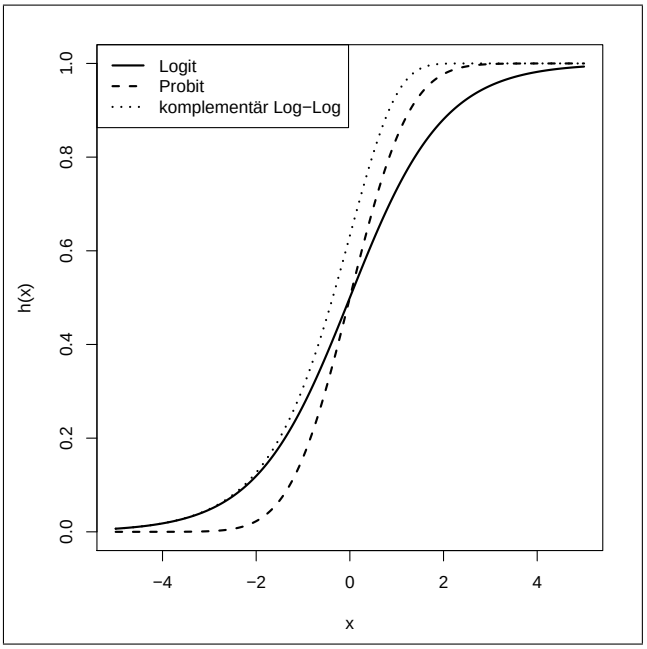


Abb. 2.3: Vergleich der Link-Funktionen des Logit-, Probit- und komplementären Log-Log-Modells

Die Extremwertverteilung ist anders als die beiden vorigen nicht symmetrisch. Sie ist der logistischen Verteilung für kleine Werte sehr ähnlich, in diesem Bereich aber deutlich größer als die Standardnormalverteilung, steigt dann schneller an und flacht später wieder ab. Einen graphischen Vergleich der drei Link-Funktionen bietet Abbildung 2.3.

2.3.3.2 Interpretation der Parameter

In allgemeinen Regressionsmodellen beschreibt β_j die Änderung der Einheiten von $C(Y|X)$ pro Änderung der Einheiten von X_j auf $X_j + 1$ (siehe Gleichung 19).
In der logistischen Regression mit Logit-Link bedeutet demnach eine Erhöhung der Variablen X_j um eine Einheit eine Erhöhung der Log-Odds um β_j Einheiten. Da diese Interpretation nicht intuitiv ist, betrachtet man meist die Größe $\exp(\beta_j)$. Sie gibt an, wie sich die Odds verändert, wenn sich die Variable X_j um eine Einheit erhöht. Die Odds kann man als Chance interpretieren, da sie das Verhältnis der Wahrscheinlichkeit zu ihrer Gegenwahrscheinlichkeit ist: $P(Y = 1|X)/(1 - P(Y = 1|X))$.
Erhöht man X_j um mehr als eine Einheit, also zum Beispiel von $X_j = a$ auf $X_j = b$, so ändert sich die Chance um den Faktor $\exp(\beta_j)^{b-a}$. Keine

Veränderung würde also ein $\exp(\beta_j) = 1$ bewirken, was bedeutet, dass der Regressionskoeffizient $\beta_j = 0$ ist. Ein positiver Regressionskoeffizient bewirkt ein $\exp(\beta_j) > 1$, also eine Erhöhung der Chance, ein negativer eine Verminderung der Chance, da $\exp(\beta_j) < 1$ ist.

2.3.3.3 Gütekriterien

Wie es in der einfachen linearen Regression das Bestimmheitsmaß R^2 gibt, das der Anteil der durch das Modell erklärten Varianz an der Gesamtvarianz ist, braucht man auch in der logistischen Regression Gütekriterien, um festzustellen, wie gut das Modell die Daten fittet.
Die Devianz (siehe Collet, 1991, Kapitel 3.8) vergleicht das saturierte Modell mit dem vorliegenden Modell mittels der Likelihood der beiden Modelle. Das saturierte Modell ist jenes Modell, das gleich viele Parameter wie Beobachtungen hat. Es wird auch das optimale Modell genannt, weil es eine exakte Vorhersage erlaubt.
Die Devianz ist definiert durch:

$$D(\beta) = -2\ln\left(\frac{L(Y|\beta)}{L(Y|Y)}\right) = -2(\ln(L(Y|\beta) - L(Y|Y))) \tag{32}$$

Die Größe $L(Y|\beta)$ ist die Likelihood des zu untersuchenden Modells, $L(Y|Y)$ bezeichnet die Likelihood des saturierten Modells.
Die Devianz ist unter der Nullhypothese, dass das geschätzte Modell nicht schlechter als das saturierte Modell ist, annähernd χ^2-verteilt mit $n - k - 1$ Freiheitsgraden, wobei n die Anzahl der Beobachtungen und k die Anzahl der geschätzten Parameter exklusive Intercept im Modell ist. Da der Erwartungswert einer χ^2-verteilten Zufallsvariablen die Anzahl der Freiheitsgrade ist, gilt für die Devianz:

$$E(D(\beta)) = n - k - 1$$

Wenn also $\frac{D(\beta)}{n-k-1} \approx 1$ gilt, dann kann von einer guten Modellanpassung ausgegangen werden.

2.4 Gruppenvergleiche

2.4.1 Varianzanalyse

Wie schon im Kapitel 2.3.1.6 kurz angedeutet, kann man ein Modell, in dem nur nominale Variablen als erklärende Variablen eingebunden werden, als varianzanalytisches Modell ansehen. Dies ist auch der einzige Unterschied zwischen dem linearen Regressionsmodell und dem varianzanalytischen Modell.

Die Varianzanalyse wurde in Anlehnung an die Bücher von Backhaus et al. (1996) und Schlittgen (2003) geschrieben.

Die Varianzanalyse ist ein Verfahren, das die Mittelwerte von zwei oder mehr Gruppen miteinander vergleicht. Sie ist eine Verallgemeinerung des Zwei-Stichproben-T-Tests, der in Kapitel 2.2.3.1 behandelt wurde. Im Falle von zwei Gruppen ist die Teststatistik der Varianzanalyse die quadrierte Teststatistik des T-Tests.

Angenommen, es gibt Beobachtungen Y_{ij}, die k Gruppen zugeordnet werden können. Die Größe der Gruppe i wird mit n_i bezeichnet. Die einfaktorielle Varianzanalyse stellt folgendes Regressionsmodell auf:

$$Y_{ij} = \mu_i + u_{ij} \ , \ i = 1, \ldots, k; \ j = 1, \ldots, n_i,$$

wobei μ_i das Mittel der Gruppe i ist. Die u_{ij} sind unabhängige, identisch verteilte Störgrößen, die um null zentriert sind.
Die Null- und Alternativhypothese der Varianzanalyse lauten:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$
$$H_1 : \mu_i \neq \mu_j \text{ für mindestens ein Paar } i \neq j$$

Kann man die Nullhypothese verwerfen, so bleibt immer noch die Frage offen, welche Gruppen sich unterscheiden. Um diese Frage beantworten zu können, werden alle möglichen Paarvergleiche durchgeführt. Bei k Gruppen entspricht dies einem Rechenaufwand von $\binom{k}{2} = \frac{k(k-1)}{2}$ T-Tests. Dies führt zu einem multiplen Testproblem, bei dem das Signifikanzniveau korrigiert werden muss. Dieses Problem wird in Kapitel 2.4.3 behandelt.

Die Varianzanalyse stellt drei Bedingungen an die Variable Y und darf streng genommen nur bei Erfüllung der folgenden Voraussetzungen durchgeführt werden:

- Die Beobachtungen sind unabhängig
- Die Fehler sind normalverteilt
- Die Varianzen in den einzelnen Gruppen sind gleich

Bei einer Verletzung der Annahme der Unabhängigkeit führt eine positive Korrelation zu einer Verschlechterung, eine negative Korrelation zu einer Verbesserung der Schätzung. Man sollte also schon im Vorfeld sicherstellen,

dass es sich um randomisierte Daten handelt.
Sind die Daten nicht normalverteilt, so greift man auf einen nichtparametrischen Test zurück, der keine Verteilungsannahme macht. Im Allgemeinen ist die Teststatistik der Varianzanalyse aber relativ robust gegen Nichtnormalität.
Bei Nichthomogenität der Varianzen kann man versuchen, durch eine varianzstabilisierende Transformation (siehe Kapitel 2.5) Homogenität zu erzeugen. Gelingt dies nicht, so sollte man einen nichtparametrischen Test rechnen. Solch ein nichtparametrischer Test, nämlich der Kruskal-Wallis-Test, wird anschließend genauer beschrieben.

Sind alle Voraussetzungen geprüft, kann man die Teststatistik betrachten. Sie geht von der Idee der Quadratsummenzerlegung aus.

$$\begin{array}{ccccc} \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{Y}_{ij}-\bar{Y}_{..})^2 & = & \sum_{i=1}^{k} n_i(Y_{i.}-\bar{Y}_{..})^2 & + & \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\bar{Y}_{i.})^2 \\ \text{SST (total sum of squares)} & = & \text{SSB (between sum of squares)} & + & \text{SSW (within sum of squares)} \end{array}$$

In dieser Gleichung bezeichnet der Ausdruck $\bar{Y}_{i.}$ den Mittelwert der Gruppe i und der Ausdruck $\bar{Y}_{..}$ das Gesamtmittel.
SST bezeichnet die gesamte Varianz in den Daten, SSB ist die Quadratsumme zwischen den Gruppen, genauer die Abweichung der Gruppenmittel vom Gesamtmittel. SSW berechnet die Quadratsumme innerhalb der Gruppen und beschreibt, wie stark die Daten um das Gesamtmittel streuen. Unter der Nullhypothese sollten SSW und SSB ungefähr gleich sein und ihr Quotient daher nahe bei eins liegen. Ist die Varianz zwischen den Gruppen allerdings größer als die Varianz innerhalb der Gruppen, so weist dies auf systematische Unterschiede hin und sollte bei der Varianzanalyse erkannt werden.
Da laut Voraussetzung die Y_{ij} normalverteilt sind, ist die Quadratsumme χ^2-verteilt. Die Anzahl der Freiheitsgrade hängt von der Anzahl der Summanden und der Anzahl der Parameter ab, die geschätzt werden müssen. SSW ist χ^2-verteilt mit $n-k$ Freiheitsgraden und SSB mit $k-1$ Freiheitsgraden. Der Quotient von zwei χ^2-verteilten Zufallsvariablen dividiert durch deren Freiheitgrade ist laut Definition F-verteilt. Deshalb ist die Teststatistik der Varianzanalyse, nämlich

$$F = \frac{\frac{1}{k-1}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\bar{Y}_{i.})^2}{\frac{1}{n-k}\sum_{i=1}^{k} n_i(Y_{i.}-\bar{Y}_{..})^2} \tag{33}$$

unter der Nullhypothese F-verteilt mit $k-1$ und $n-k$ Freiheitsgraden.

Tab. 2.3: ANOVA-Tabelle

Varianz-quelle	*Freiheits-grade*	*Quadrat-summe*	*mittlere Quadratsumme*	*F-Statisik*	*Signi-fikanz*
Faktor	$k-1$	SSW	$\text{MSW} = \text{SSW}/(k-1)$	$F = \frac{\text{MSW}}{\text{MSB}}$	$\text{P}(> F)$
Fehler	$n-k$	SSB	$\text{MSB} = \text{SSB}/(n-k)$		

Die Ergebnisse der Quadratsummenzerlegung werden in der ANOVA-Tabelle zusammengefasst. Die typische Struktur der ANOVA-Tabelle sieht man in Tabelle 2.3.
Natürlich gibt es auch ein Modell für eine mehrfaktorielle Varianzanalyse. Es gibt dann nicht nur einen Faktor, sondern mindestens zwei. In solch einem Modell kann man auch Wechselwirkungen zwischen den Faktoren berücksichtigen. Die Quadratsumme zwischen den Gruppen wird dann aufgeteilt in die Quadratsummen, die von den einzelnen Faktoren stammen, und in jene, die auf die Wechselwirkungen zurückgehen. Es können auch Kovariablen, also stetige Variablen, und nicht nur nominale Variablen im Modell sein. Man spricht dann von einer *Kovarianzanalyse*.

2.4.2 Kruskal-Wallis-Test

Wie schon im vorigen Kapitel erwähnt, müssen für die Varianzanalyse einige Kriterien, wie zum Beispiel Normalverteilung und Varianzhomogenität, gegeben sein. Sind diese Voraussetzungen nicht erfüllt, so sollte man statt der Varianzanalyse den Kruskal-Wallis-Test (Kruskal und Wallis, 1952, S. 599) einsetzen. Er stellt eine Verallgemeinerung des Wilcoxon-Rangsummentests (Kruskal und Wallis, 1952, S. 602) dar, der bereits in Kapitel 2.2.3.2 beschrieben wurde, so wie die Varianzanalyse eine Verallgemeinerung des T-Tests ist.

Bei der Berechnung der Teststatistik verwendet man nicht die Originaldaten, sondern greift, wie beim Wilcoxon-Rangsummentest, auf deren Ränge zurück. Angenommen, es gibt n Beobachtungen, die sich in k Gruppen zu je $n_i, i = 1, \ldots k$ mit $\sum_{i=1}^{k} n_i = n$ Beobachtungen aufteilen. Die Nullhypothese lautet:

H_0 : Die Beobachtungen kommen aus der gleichen Grundgesamtheit.

Die Alternative ist:

H_1 : Die Verteilungen der Gruppen unterscheiden sich im Lageparameter.

Nun ordnet man die Beobachtungen aufsteigend und weist jedem Wert seinen Rang $R_{ik}, i = 1, \ldots n$ in der Vereinigung der Gruppen zu. Haben zwei

Beobachtungen den gleichen Originalwert, so werden Rangmittel gebildet. Nun wird in jeder Gruppe $i = 1, \ldots, k$ die Rangsumme ermittelt:

$$S_i = \sum_{j=1}^{n_k} R_{jk}.$$

Die Teststatistik des Kruskal-Wallis-Tests (siehe Kruskal, 1952) wird dann folgendermaßen berechnet:

$$K = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{S_i^2}{n_i} - 3(n+1) \tag{34}$$

Sie ist für hinreichend große Stichproben annähernd χ^2-verteilt mit $k-1$ Freiheitsgraden.
Wie bei der Varianzanalyse ist man bei Ablehnung der Nullhypothese daran interessiert, welche Gruppen sich hinsichtlich ihres Lageparameters unterscheiden. Man führt im Anschluss also alle möglichen Paarvergleiche mittels Wilcoxon-Rangsummentests, die sogenannten Post-Hoc-Tests, durch und korrigiert die Signifikanzniveaus, wie im folgenden Kapitel 2.4.3 beschrieben wird.

2.4.3 Korrektur nach Bonferroni

Wie schon in den vorigen Kapiteln erwähnt, werden alle möglichen Paarvergleiche durchgeführt, wenn die Nullhypothese der Varianzanalyse oder des Kruskal-Wallis-Tests abgelehnt wird. Bei k Faktorstufen sind dies $\binom{k}{2} = \frac{k(k-1)}{2}$ Vergleiche, bei 4 Faktorstufen wären dies 6, bei 5 Stufen schon 10 und bei 15 Stufen 105 Tests. Da die Paarvergleiche auf der Basis gleicher Daten durchgeführt werden, liegt ein multiples Testproblem vor. Die Schwierigkeit dabei ist, dass es nicht mehr reicht, dass jeder einzelne Test das Signifikanzniveau α hält. Vielmehr muss sichergestellt sein, dass der multiple Test das Signifikanzniveau α hält. Es muss also sichergestellt sein, dass die Wahrscheinlichkeit, irgendeine wahre Nullhypothese abzulehnen, gleich α ist.

Einen sehr einfachen Ausweg bietet die Bonferroni-Ungleichung (vgl. Christensen, 2002). Angenommen, es liegen k Nullhypothesen vor. Hält jeder einzelne Paarvergleich das Signifikanzniveau α/k, so hält der multiple Test das Signifikanzniveau α. Dies kann leicht mittels der Bonferroni-Ungleichung nachgewiesen werden. Sind B_i beliebige Mengen, so gilt:

$$P(\bigcup_{i=1}^{k} B_i) \leq \sum_{i=1}^{k} P(B_i)$$

Notiert man mit A_i den Ablehnbereich der Nullhypothese i, so ist die Wahrscheinlichkeit, mindestens eine Nullhypothese fälschlicherweise abzulehnen:

$$P(\bigcup_{i=1}^{k} A_i) \leq \sum_{i=1}^{k} P(A_i) = \sum_{i=1}^{k} \frac{\alpha}{k} = \alpha$$

Die Korrektur nach Bonferroni ist sehr konservativ, sie nimmt also eher die Nullhypothese an, bevor sie eine Abweichung von der Nullhypothese als signifikant ansieht. Es gibt natürlich auch bessere Methoden, wie zum Beispiel die Bonferroni-Holm-Prozedur, bei der das Signifikanzniveau schrittweise angehoben wird, oder das Abschlussprinzip, bei dem die Menge der zu testenden Nullhypothesen um alle möglichen Durchschnitte erweitert und ein absteigendes Verfahren durchgeführt wird.

2.5 Datentransformation

Datentransformationen werden vor allem dann durchgeführt, wenn die Verteilung einer Variablen sehr schief oder die Varianz nicht homogen ist. Einige statistische Methoden können in so einem Fall nicht angewendet werden und würden zu verzerrten Ergebnissen führen. Manchmal wird eine Transformation auch zur Erhöhung der Interpretierbarkeit durchgeführt, wenn die Daten zum Beispiel in einer Maßeinheit angegeben sind, die einem nicht geläufig ist. Die literarische Grundlage für dieses Kapitel liefern Heiler und Michels (1994, Kapitel 5) und Schlittgen (2003, Kapitel 8.2).

Eine Transformation ist eine Abbildung T: $\boldsymbol{R} \rightarrow \boldsymbol{R}$. Eine weit verbreitete Klasse von Transformationen ist die Potenztransformation, die sowohl eine symmetrisierende als auch eine varianzstabilisierende Wirkung hat.

Die Potenztransformation hat folgende Gestalt:

$$T(x) = \begin{cases} ax^p + b & , p \neq 0 \\ c\log(x) + d & , p = 0 \end{cases} \tag{35}$$

Meist wird das Vorzeichen von a gleich dem Vorzeichen von p gewählt, da eine negative Potenz die Ordnung der Daten umkehren würde und dies durch die Multiplikation mit einem negativen Wert wieder ausgeglichen wird. Sind die Datenwerte x nicht positiv, so sollte zuvor ein konstanter Wert addiert werden, der alle x positiv macht.
Für $p > 1$ ist die Funktion $T(x)$ konvex[9], was bedeutet, dass kleine Werte

[9]Eine Funktion ist konvex, wenn eine beliebige Strecke zwischen zwei Punkten der Kurve über der Kurve liegt (vgl. Abbildung 2.4).

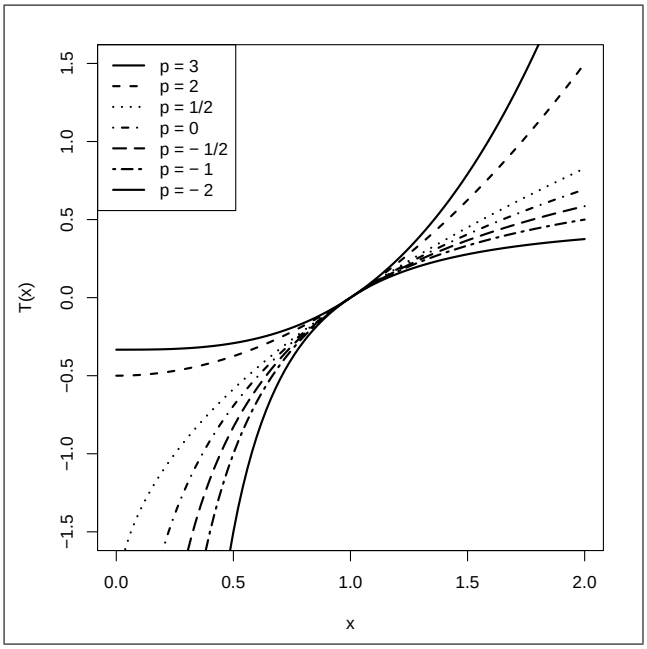


Abb. 2.4: Vergleich der Potenztransformationen für verschiedene Potenzen p und Parameter $a = sign(p)\frac{1}{p}$, $b = sign(p)\left(-\frac{1}{p}\right)$, $c = 1$ und $d = 0$

zusammengezogen werden. Für $p < 1$ ist die Funktion konkav[10], was zu einer Stauchung der größeren Werte führt. Eine Potenz $p > 1$ sollte also bei linksschiefen Verteilungen angewendet werden, für rechtsschiefe Verteilungen ist eine Wahl von $p < 1$ zu empfehlen. Die Potenz $p = 1$ braucht man nicht zu berücksichtigen, sie ist nur eine lineare Transformation, welche die Werte um den Wert b verschiebt und den Wert a verzerrt. Sie ändert aber nichts an der Schiefe der Verteilung.
Einen Vergleich der Transformationen für verschiedene Potenzen p sieht man in Abbildung 2.4.
Welche Transformation im Endeffekt angewendet wird, kann man meist nur durch Versuch und Irrtum herausfinden, indem man die verschiedenen Transformationen durchprobiert. Es genügt im Allgemeinen, Potenzen wie 3, 2, 1, 1/2, 0, -1/2, -1, -2 etc. zu betrachten. Eine graphische Analyse oder die Berechnung der Schiefe mit den transformierten Daten gibt dann die endgültige Entscheidung. Man kann sich auch sogenannter Transformationsplots bedienen, welche die Streuung in Abhängigkeit der Lageparameter betrachten. Liegen die Punkte in etwa auf einer Geraden, so kann die Potenz mittels der Steigung des Transformationsplots ermittelt werden als $p = 1 -$ Steigung.

[10] Gegenteil von konvex

2.6 Fehlende Werte

In vielen Datensätzen sieht man sich mit dem Problem fehlender Werte konfrontiert. Es gibt verschiedene Strategien, mit dieser Angelegenheit umzugehen. Es kommt vor allem darauf an, warum die Werte fehlen. In diesem Abschnitt wird Harrell (2001, Kapitel 3) als Literatur herangezogen.
Man unterscheidet hierbei drei Gründe:

1. Vollständig zufälliges Fehlen (missing completely at random, MCAR): Der Grund für das Fehlen des Beobachtungswertes steht weder im Zusammenhang mit den beobachteten noch mit den fehlenden Werten.

2. Zufälliges Fehlen (missing at random, MAR): Die Werte fehlen nicht völlig zufällig, das Fehlen ist sehr wohl abhängig von den beobachteten Werten.

3. Nichtzufälliges Fehlen (missing not at random, MNAR, bzw. informative missing, IM): Das Fehlen der Beobachtung lässt einen Schluss auf den Wert zu. Dieser Fall tritt ein, wenn der beobachtete Wert systematisch unter bzw. ober den typischen Werten liegt.

Eine einfache und schnell zu realisierende Möglichkeit ist das zeilenweise Löschen, wenn eine Variable nicht erhoben wurde. Es hat jedoch den großen Nachteil, dass die Stichprobengröße reduziert wird. Die fehlenden Werte können auch imputiert werden. Das heißt, es werden Werte geschätzt, die anstatt des fehlenden Wertes eingesetzt werden. Im Falle einer Imputation sollte unbedingt markiert werden, welche Werte geschätzt wurden, damit man später nicht glaubt, dass es sich um wahre Beobachtungen handelt. Die Imputation hat den großen Vorteil – obwohl es im ersten Moment etwas unseriös wirkt –, dass man auf wertvolle Information, die ja für andere Variablen trotzdem vollständig vorliegt, nicht verzichten muss. Es werden nun zwei Methoden zur Imputation vorgestellt.

2.6.1 Singuläre Imputation

Man kann einerseits einen Zufallswert aus den restlichen vorliegenden Daten generieren, andererseits einen konstanten Wert einsetzen, wie etwa den Mittelwert oder den Median für kontinuierliche Variablen oder den Modus für kategoriale Daten. Die Imputation von Mittelwert oder Median hat zur Folge, dass es in der Analyse eine starke Beziehung zwischen der abhängigen und der unabhängigen Variablen gibt, die in Wahrheit nicht besteht. Außerdem wird die Varianz zu klein geschätzt, da die substituierten Werte keine Variation untereinander aufweisen.

Eine verbreitete Methode basiert auf Regressionsmodellen. Angenommen, es gibt in einem Datensatz $X_1, \ldots X_k$ Variablen und man will die fehlenden Werte der Variablen X_j schätzen, so stellt man ein geeignetes Regressionsmodell auf, in dem X_j der Regressand und die restlichen $k-1$ Variablen die Regressoren sind. Dies kann ein einfaches oder multivariates lineares Modell oder bei kategorialen Variablen ein logistisches Modell sein. Die fehlenden Werte werden durch die prognostizierten Werte, zu denen man nach Belieben auch einen zufälligen Fehler addieren kann, ersetzt.

2.6.2 Multiple Imputation

Bei der multiplen Imputation produziert man im Gegensatz zur singulären Imputation für jeden fehlenden Wert nicht nur einen, sondern gleich mehrere Ersatzwerte. Wie viele solcher Ersatzwerte generiert werden sollen, hängt vom Anteil der fehlenden Werte ab. In der Literatur wird meist eine Anzahl zwischen 3 und 10 vorgeschlagen. Im Anschluss mittelt man die vorgeschlagenen Werte und berechnet die Standardfehler, indem man auch die Variation innerhalb der Imputationsvorschläge miteinbezieht.

Es gibt eine Vielzahl neuerer und effizienterer Methoden für die Imputation fehlender Werte, wie die Einbindung des EM-Algorithmus. Bei Interesse findet man nähere Informationen bei Schafer (1997).
Welche Methode der Imputation man bei einem Datensatz verwendet, ist abhängig von der speziellen Verteilung der Daten, dem Anteil an fehlenden Werten und dem Grund des Fehlens, der eingangs des Kapitels beschrieben wurde.

3 Auswertung

3.1 Software

Die Auswertung in dieser Arbeit wurde mit der Open-Source-Software `R` (R Development Core Team (2007)) durchgeführt. Das Programm hat sich weltweit in allen wesentlichen Bereichen der angewandten Statistik durchgesetzt. Die Programmiersprache von `R` orientiert sich stark an der Sprache S und deren Nachfolger S-Plus, weshalb Programme aus `R` im Allgemeinen auch in den Programmen S und S-Plus funktionieren. Es haben sich allerdings feine Unterschiede herauskristallisiert. Ein Team aus freiwilligen Helfern entwickelt `R` ständig weiter und fügt neue Funktionen bzw. Packages hinzu.

Die meisten Funktionen, die man bei der Auswertung des Datensatzes benötigt, findet man in den Packages `stats` und `Design` (Harrell, 2007). `Design` ist ein Package, das aus etwa 180 Funktionen besteht, mit denen man Regressionsmodelle fitten, Hypothesen testen und Schätzungen durchführen kann. Es bietet sowohl Methoden zur Modellvalidation als auch zur Prognose und außerdem einige nützliche graphische Funktionen.

Für die Regressionsanalyse wird für ein logistisches Modell die Funktion `glmD()` aus dem Package `Design` verwendet. Die Funktion dient dazu, um verallgemeinerte lineare Modelle an Daten anzupassen. Sie kann eine Reihe von Verteilungen behandeln, wie zum Beispiel Normal-, Binomial-, Poisson-, Gamma- und Inverse Gauß-Verteilung. Außerdem lässt sie bei der logistischen Regression die drei Link-Funktionen zu, die in Kapitel 2.3.3.1 beschrieben werden. Aus demselben Package wurde die Funktion `ols()` für lineare Modelle eingesetzt. Beide Funktionen entsprechen den Funktionen `glm()` und `lm()` aus dem Package `stats`. Sie speichern aber zusätzlich die Varianz-Kovarianzmatrix und lassen außerdem regularisierte Kleinst-Quadrate-Schätzung (penalized least squares) zu. Außerdem geben sie zusätzlich die Likelihood-Ratio-Teststatistik an, die in der Auswertung verwendet wird, um Hypothesen zu testen. Beide einander entsprechenden Funktionen liefern aber das gleiche Ergebnis für die Parameterschätzungen.
Um eine Variable als restringierten kubischen Spline in die Regression aufzunehmen, wird sie mit der Funktion `rcs()` aus dem Package `Design` transformiert. Diese Funktion berechnet die konstruierten Variablen wie in Gleichung 22 und wählt die Stützstellen, sofern man explizit keine angibt, als fixe Quantile, die Tabelle 2.2 vorschlägt.
Restringierte Wechselwirkungen können mit der Funktion `%ia%` angegeben werden. Dies ist zweckmäßig, wenn zum Beispiel ein Haupteffekt als kubi-

scher Spline im Modell ist, die Wechselwirkung mit einer anderen Variablen aber nur für den linearen Teil geschätzt werden soll.

Um Hypothesen in einem Regressionsmodell zu testen, werden die Funktionen `anova.Design()` und `lrtest()` aus dem Package `Design` verwendet. Die Funktion `anova.Design()` führt den Wald-Test für einige standardmäßige Hypothesen durch. Die Funktion `lrtest()` vergleicht die Likelihood-Ratio-Teststatistiken zweier Modelle, indem sie deren Differenz bildet und mit dem kritischen Wert der χ^2-Verteilung vergleicht.

Die Histogramme bzw. Boxplots werden mit den Funktionen `hist()` bzw. `boxplot()` aus dem Package `graphics` erzeugt. Für die Plots in Kapitel 3.3 gibt es im Package `Design` eine Funktion `plot.Design()`, die direkt aus dem Modellobjekt die gewünschte Abbildung erzeugt. Man kann hierbei angeben, welche Variablen geplottet werden und welche fixiert werden sollen. Gibt man zwei kontinuierliche Variablen an, die gleichzeitig dargestellt werden sollen, so bedient sich `plot.Design()` der Funktion `persp()` aus dem Package `graphics`, um einen 3D-Plot zu erzeugen. Außerdem kann man für ein logistisches Modell einstellen, ob auf der z-Achse die Log-Odds oder der transformierte Prädiktor dargestellt wird. In allen Abbildungen in Abschnitt 3.3 werden wegen der Interpretierbarkeit die Wahrscheinlichkeiten anstatt der Log-Odds aufgetragen.

3.2 Beschreibung des Datensatzes

Der Datensatz stammt aus der Ad-Bench Datenbank, dessen Basis der Sujet-Focus-Test bildet (siehe Sujet-Focus, Focus Institut Marketing Research Ges. m.b.H.). Die Befragung und sofortige Eingabe der Antworten führt ein Interviewer durch. Somit ist sichergestellt, dass alle Fragen allen Personen gestellt, diese auch beantwortet und alle Sujets gezeigt werden. Da die Testpersonen die Sujets in zufälliger Rotation sehen, kann man ausschließen, dass die Reihenfolge einen Einfluss auf die Interpretation der Ergebnisse hat. Die Befragung findet nur in Wien statt, was eine Ausweitung der Interpretation auf die Einwohner Österreichs nicht zulässt.
Im Datensatz befinden sich, wie schon in der Einleitung erwähnt, Werbungen aus sieben verschiedenen Werbeträgern, und zwar aus den Bereichen Hörfunk, Infoscreen, Online, Plakat, Print, Prospekt und TV[11], wobei die

[11] Im folgenden Text werden die Werbeträger häufig abgekürzt, wobei Hörfunk=HF, Infoscreen=IS, Online=OL, Print=PR, Plakat=PL und Prospekt=PROS bedeuten. Außerdem steht in Tabellen und Abbildungen statt Werbeträger oft WT und statt Wirtschaftsbereich WB.

Tab. 3.1: Anzahl der Sujets in den Werbeträgern

Werbeträger	*Anzahl*	*Prozent*
Hörfunk	2.633	18,94 %
Infoscreen	346	2,49 %
Online	96	0,69 %
Plakat	1.788	12,86 %
Print	1.915	13,77 %
Prospekt	848	6,10 %
TV	6.279	45,16 %
Summe	13.905	100,00 %

Prospektsujets bei der Befragung gesondert behandelt werden. Die restlichen Werbeträger werden gleich abgetestet, was einen Intra- und Intermediavergleich möglich macht (siehe Sujet-Focus, Focus Institut Marketing Research Ges.m.b.H.).
In Tabelle 3.1 sieht man die Aufteilung der Sujets auf die einzelnen Werbeträger. Sujets aus der Kategorie „Infoscreen“ wurden erst ab Oktober 2001 in die Befragung aufgenommen. Da diese auch nicht jedes Monat abgefragt wurden, gibt es bei diesem Werbeträger wenige Einträge. In der Kategorie „Online“ finden sich nur 96 Sujets, da diese Mediengattung nur im Juli, September und Oktober 2006 und im ersten Halbjahr 2007 (bis auf Mai) Beobachtungen aufweist. Der Werbeträger „Prospekt“ wurde ab Dezember 2002 in die Erhebung aufgenommen. In den restlichen vier Kategorien „Hörfunk“, „Plakat“, „Print“ und „TV“ gibt es schon seit Anfang des Jahres 1998 Einträge.

Aus dem Datensatz ist nicht ersichtlich, um welches Sujet es sich handelt, das heißt also, von welcher Firma und für welches Produkt geworben wird. Es wurde nur angegeben, in welchen Wirtschaftsbereich sich das beworbene Produkt einordnen lässt. Eine Übersicht, welche Wirtschaftbereiche unterschieden werden und wie sich die Sujets darauf verteilen, gibt Tabelle 3.2. Außerdem werden Warenkorb und Warengruppe angeführt, die eine genauere Zuordnung ermöglichen.

Für alle sieben Werbeträger werden die Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* erhoben. Für alle Werbeträger werden außerdem Imagevariablen erhoben. Die Erhebung unterscheidet sich, wie bereits angesprochen, für den Datensatz „Prospekt“, in dem die Imagevariablen anders eingestuft und erhoben werden.
Die Variable *Bekanntheit* gibt den Prozentsatz der Personen an, die das Su-

Tab. 3.2: Anzahl der Sujets in den Wirtschaftsbereichen

Wirtschaftsbereich	*Anzahl*	*Prozent*
Audio/Video	449	3,23 %
Bau	396	2,85 %
Dienstleistungen	359	2,58 %
Energie	230	1,65 %
Ernährung	1.697	12,20 %
Finanzen	1.342	9,65 %
Freizeit/Sport	202	1,45 %
Getränke	803	5,77 %
Handel	733	5,27 %
Haus&Garten	938	6,75 %
Investitionsgüter	95	0,68 %
KFZ	1.745	12,55 %
Kommunikation/Büro	1.618	11,64 %
Kosmetik	1.358	9,77 %
Medien	492	3,54 %
Persönlicher Bedarf	87	0,63 %
Reinigung	466	3,35 %
Textil	399	2,87 %
Touristik	462	3,32 %
Sonstige	31	0,22 %
NA	3	0,02 %
Summe	13.905	100,00 %

jet vor der Befragung bereits gekannt haben. Nachdem die Testpersonen das Sujet gesehen haben, geben sie auf einer Skala von 1 bis 10 an, wie sehr es sie anspricht, wobei 10 die höchste Bewertung darstellt. Der Mittelwert der angegebenen Punkte wird in der Variablen *Gefälligkeit* gespeichert.
Anschließend werden den befragten Personen die Imagekriterien „ansprechend“, „originell“, „modern“, „informativ“, „aggressiv“, „sympathisch“, „verständlich“, „anregend“, „auffällig“ und „keines davon“ vorgelegt. Daraufhin müssen sie angeben, welches der Imagekriterien das eben gesehene Sujet am besten charakterisiert. In den dazugehörigen Imagevariablen wurde der Prozentsatz der Personen angegeben, die dieses Kriterium genannt haben. Wenn ein Sujet somit in der Imagekategorie „originell“ einen niedrigen Wert aufweist, bedeutet dies nicht, dass es unoriginell ist, sondern nur dass es sich durch Originalität nicht besonders auszeichnet. Die Imagevariablen sind vielmehr ein Polaritätenprofil.
Nachdem die Testpersonen alle Sujets gesehen haben, werden sie gefragt, an

welche sie sich noch erinnern können. Es genügt hierbei aber nicht, einen Oberbegriff, wie zum Beispiel „Handy“, zu nennen, der Person muss die beworbene Marke, zum Beispiel „Nokia“ oder „Sony Ericsson“, einfallen. Den Anteil, wie viele Personen das beworbene Produkt noch im Kopf haben, speichert die Variable *Markenimpact*.

Zu jedem Sujet werden 150 Personen befragt, quotiert nach Geschlecht und Alter der österreichischen Bevölkerung. Die Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* werden außerdem getrennt für die Zielgruppen, die durch Geschlecht und Alter definiert werden, erfasst. Die Alterseinteilung erfolgt in die Gruppen „14 bis 29“, „30 bis 49“ sowie „50 Jahre und älter“. Außerdem werden als Zusatzinformationen sowohl Haushaltsgröße als auch -einkommen erhoben. Auch für diese Kategorien werden die drei Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* separat angeführt.

Weiters werden die Variablen *Länge*, *Aufwand*, *Monat* und *Jahr* in den Datensatz aufgenommen. *Länge* und *Aufwand* kommen sowohl kontinuierlich als auch kategorial vor.

Vor einer explorativen Analyse sollte man die Variablen in Hinblick auf deskriptive Statistiken betrachten. Einige beschreibende Maßzahlen, wie Lage- und Streuungsmaße oder die Schiefe, sowie eine graphische Darstellung sind hilfreich. Als Maßzahlen werden in der Auswertung das Minimum und Maximum, oberes und unteres Quartil, der Median, das arithmetische Mittel, die Standardabweichung und die Schiefe angegeben. Die Definition und die Interpretation dieser Maßzahlen sind in Kapitel 2.1 zu finden.
Für die graphische Darstellung wird entweder ein Histogramm genommen, das die absoluten oder relativen Häufigkeiten veranschaulicht, wenn die empirische Verteilung mit einer theoretischen verglichen wird, oder es wird ein Boxplot gewählt. Am Boxplot kann man fast alle oben genannten Maßzahlen ablesen. Die Box wird durch das erste und dritte Quartil begrenzt, innerhalb der Box liegen also 50 % der Datenpunkte. Der Strich in der Mitte der Box kennzeichnet den Median. Liegt der Median in der Box sehr weit unten bzw. oben, so ist die zugrundeliegende Verteilung rechts- bzw. linksschief. Die Differenz von drittem und erstem Quartil ist ein Maß für die Streuung, der sogenannte Interquartilsabstand. Die *Länge* der vertikalen Linien, die außerhalb der Box liegen, ist maximal das 1,5-fache des Interquartilsabstandes. Daten, die außerhalb dieses Bereichs liegen, werden als Ausreißer deklariert und extra gekennzeichnet.

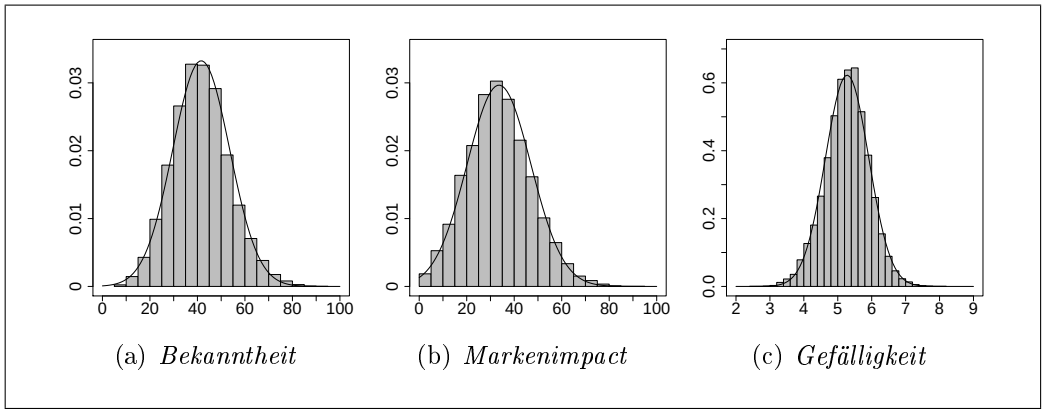
In Tabelle 3.3 sind deskriptive Statistiken für die drei Variablen *Bekannt-*

Tab. 3.3: Deskriptive Maßzahlen der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit*

	Bekanntheit	*Markenimpact*	*Gefälligkeit*
n	13.905	13.901	13.905
NA's	0	4	0
Minimum	6,7416	0,0000	2,5000
1. Quartil	33,3333	24,3421	4,8733
Median	41,0676	33,0043	5,2980
Mittelwert	41,6000	33,5987	5,2753
3. Quartil	49,1520	42,3841	5,6933
Maximum	92,2949	90,1316	8,2000
Stabw	11,9961	13,4465	0,6406
Schiefe	0,2592	0,2416	−0,0899

heit, *Markenimpact* und *Gefälligkeit* zusammengefasst. Die erste Zeile gibt die Anzahl der Einträge und die zweite Zeile die Anzahl der fehlenden Werte an. Als Mittel wird das arithmetische Mittel der Variablen angeführt. Im Schnitt kannten also 41,6 % der Testpersonen das gezeigte Sujet, 33,6 % konnten sich an das vorgeführte Sujet nach der Befragung noch erinnern und es auch nennen. Die durchschnittliche Beurteilung aller Sujets, zusammengefasst auf einer Skala von 1 bis 10, liegt bei 5,28. Die Streuung der Variablen *Bekanntheit* liegt bei zirka 12. Die Variable *Markenimpact* streut mit 13,45 ein bisschen mehr. Die Standardabweichung der Variablen *Gefälligkeit* liegt bei 0,64. Dass diese deutlich kleiner als die Streuung der anderen beiden Variablen ist, liegt daran, dass sie in einem anderen Wertebereich liegt.
Die letzte Zeile enthält die Schiefe. Die Schiefe der Variablen *Bekanntheit* und *Markenimpact* ist positiv, aber relativ klein. Eine positive Schiefe bedeutet, dass die Masse der Daten eher auf der linken Seite liegt und somit der Median kleiner als der Mittelwert ist. Die Schiefe der Variablen *Gefälligkeit* ist negativ, aus der Tabelle ist auch ersichtlich, dass der Mittelwert bei dieser Variablen kleiner als der Median ist.

Histogramme dieser Variablen sieht man in der Abbildung 3.1. Die Linie in den Abbildungen stellt die Dichte der Normalverteilung als Vergleich zur empirischen Verteilung dar. Man kann erkennen, dass die Verteilungen der Variablen *Bekanntheit* und *Markenimpact*, obwohl es sich um Prozentsätze handelt, sehr stark an Normalverteilungen erinnern. Die Verteilung der Variable *Gefälligkeit* ähnelt auch einer Gauß'schen Glockenkurve.



(a) *Bekanntheit* (b) *Markenimpact* (c) *Gefälligkeit*

Abb. 3.1: Histogramme mit Normalverteilungskurve

Die Variable *Länge* liegt als kontinuierliche Variable vor. Betrachtet man die Einträge in der Variablen jedoch genauer, so fällt auf, dass sie nicht sehr viele unterschiedliche Werte annimmt. Außerdem ist auffällig, wenn man die Variablen *Länge* und *Längenkategorien* parallel betrachtet, dass viele Werte genau an den oberen bzw. unteren Bereichsgrenzen liegen. Auch gibt es Wertepaare, die nicht zusammenpassen. So gibt es Werte in der Kategorie „15 bis 20“, die kleiner als 15 sind. Deshalb wurde die Variable für alle Werbeträger in Kategorien eingeteilt und als kategoriale Variable in der Analyse betrachtet. Der Informationsverlust durch Kategorienbildung wird auf jeden Fall wieder wettgemacht, wenn man bedenkt, dass die Einträge falsche Informationen abspeichern.
Es wurden die Kategorien „$\leq$ 15“, „15 bis 20“, „20 bis 25“, „25 bis 30“ und „länger als 30“ gewählt, wobei die Obergrenze eine Inklusion bedeutet. Zur Kategorie „15 bis 20“ werden also jene Sujets gezählt, die länger als 15 Sekunden, aber höchstens 20 Sekunden dauern. Für den Werbeträger „Infoscreen“ ist diese Einteilung nicht vorteilhaft und deswegen wurden für die Analyse eigene Kategorien gewählt. Nähere Informationen dazu findet man in Kapitel 3.3.2. Außerdem gibt es eine Kategorie „Tandem“, die jene Spots zusammenfasst, die hintereinander geschalten werden. Sie werden in einer eigenen Kategorie behandelt, weil sie mit anderen Spots der gleichen *Länge* nicht vergleichbar sind. Innerhalb der Tandem-Spots wurde keine weitere Unterscheidung getroffen.

Bevor mit der Analyse begonnen wird, muss außerdem noch geklärt werden, wie mit fehlenden Werten umgegangen wird. In diesem Datensatz ist auffällig, dass einige Variablen für gewisse Werbeträger nicht erhoben wur-

Tab. 3.4: Anzahl und Anteil fehlender Werte der Variablen *Aufwand* und *Länge*

	Aufwand		*Länge*		*beide fehlend*	
HF	7	(0,27 %)	6	(0,23 %)	1	(0,04 %)
IS	0	(0 %)	0	(0 %)	0	(0 %)
OL	96	(100 %)	96	(100 %)	96	(100 %)
PL	1.788	(100 %)	1.788	(100 %)	1.788	(100 %)
PR	21	(1,1 %)	1.915	(100 %)	21	(1,1 %)
PROS	848	(100 %)	848	(100 %)	848	(100 %)
TV	11	(0,18 %)	36	(0,57 %)	2	(0,03 %)

den (siehe Tabelle 3.4). Zum Beispiel wurde die Variable *Aufwand* für die Werbeträger „Online“, „Plakat“ und „Prospekt“ nicht erhoben, die Variable *Länge* außerdem auch für den Werbeträger „Print“ nicht. Dies ist aber nicht verwunderlich, da Sujets in den vier Mediengattungen keine messbare Länge haben. Die Länge kann also in diesen Kategorien nicht imputiert werden, da dies keinen Sinn ergäbe. Die Variable *Aufwand* in den drei oben genannten Werbeträgern zu imputieren, ist insofern schwierig, als man überhaupt keinen Anhaltspunkt hätte, wie hoch der Aufwand in etwa wäre. Eine Imputation in diesen vier Werbeträgern wird also nicht durchgeführt.

In den restlichen drei Werbeträgern, nämlich „Hörfunk“, „Infoscreen“ und „TV“, wurden die Variablen *Aufwand* und *Länge* sehr wohl erhoben. Die Anzahl und den Anteil der fehlenden Werte sieht man auch in Tabelle 3.4.
Bei den Variablen *Bekanntheit* und *Gefälligkeit* gibt es keine fehlenden Werte. Die Variable *Markenimpact* hat vier fehlende Werte, wobei zwei auf den Werbeträger „Hörfunk“ und zwei auf den Werbeträger „Print“ entfallen.
Da in allen Gruppen und bei allen Variablen der Prozentsatz an fehlenden Werten sehr gering ist, werden auch in diesen Werbeträgern die Werte nicht imputiert. Harrell (2001, Seite 49) merkt an, dass es bei einem Prozentsatz an fehlenden Werten von weniger als 5 % fast keinen Unterschied macht, für welche Art der Imputation man sich entscheidet. Auch eine Analyse auf Basis des vollständigen Datensatzes ist denkbar. Bei dieser Arbeit verwendet man deshalb nur jene Beobachtungen, bei denen alle interessierenden Variablen erhoben wurden.

3.3 Einfluss von Aufwand, Länge und Zeit

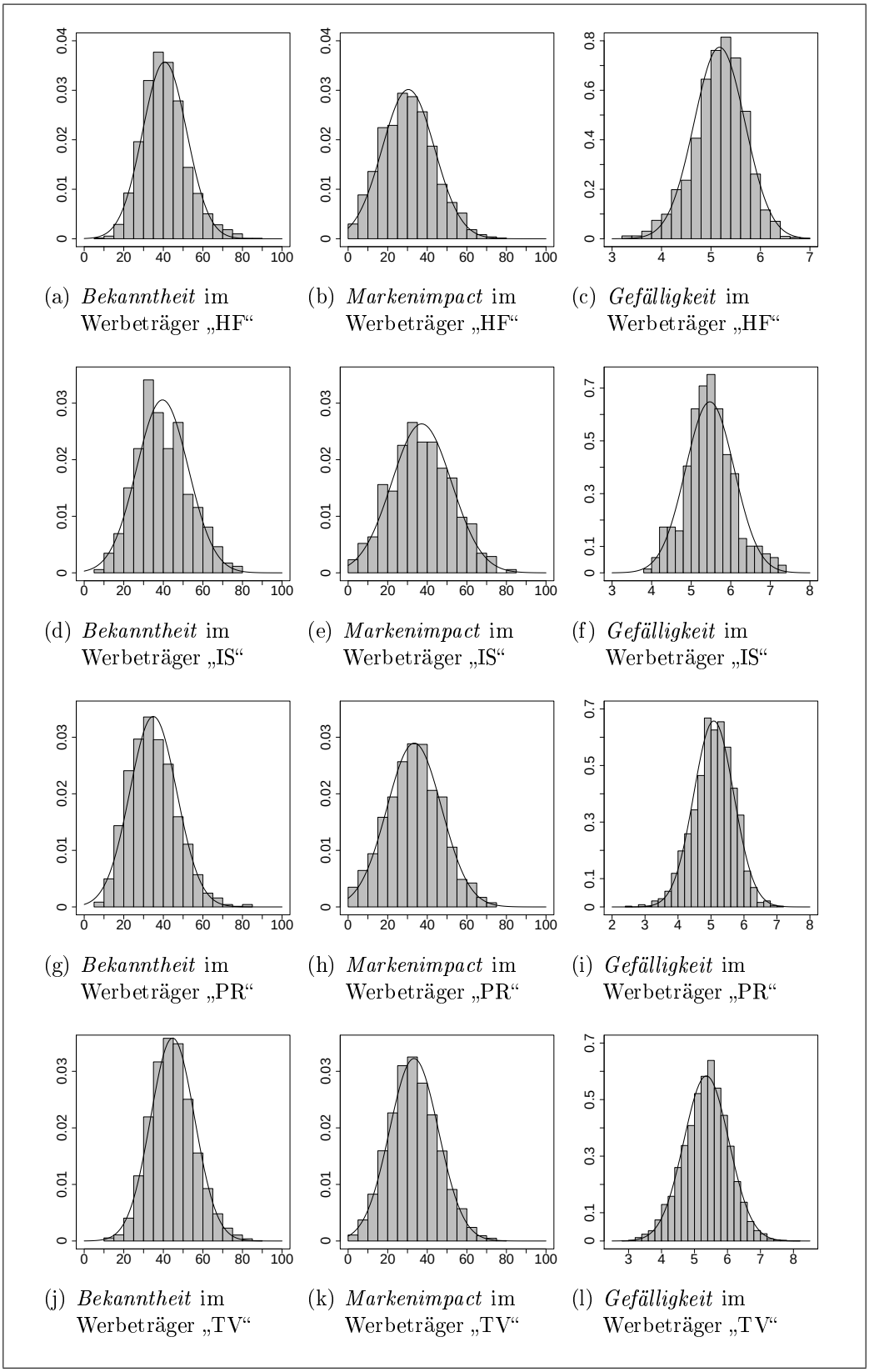
Um zu überprüfen, ob der Aufwand, die Spotlänge oder die Zeit einen Einfluss auf die Variablen *Bekanntheit*, *Gefälligkeit* und *Markenimpact* haben,

werden Regressionsmodelle angepasst. Für die Variablen *Bekanntheit* und *Markenimpact* sollte ein logistisches Modell gefittet werden, da es sich hierbei um Prozentsätze handelt. Die Variable *Gefälligkeit* ist eine kontinuierliche Variable im Bereich 1 bis 10, bei der eine lineare Regression gerechnet wird. Es soll bei der Fragestellung kein Modell zur Prognose aufgestellt werden, es handelt sich lediglich um Hypothesentests. Deshalb wird keine Modellvalidation durchgeführt. Die Hypothesen werden mit der Wald- und der Likelihood-Ratio-Statistik untersucht. Die beiden Tests sollten zum gleichen Ergebnis führen, da die beiden Statistiken trotz unterschiedlicher Berechnung asymptotisch gleich sind (siehe Kapitel 2.2).

Passt man ein Modell mit den Variablen *Aufwand*, *Zeit* und der kategorialen Variablen *Werbeträger* und deren linearen Wechselwirkungen an, so sind die Wechselwirkungen zwischen *Aufwand* und *Werbeträger* bzw. *Zeit* und *Werbeträger* signifikant. Dies und auch inhaltliche Überlegungen sprechen für eine separate Analyse in den Werbeträgern „Hörfunk“, „Infoscreen“, „Print“ und „TV“.
Schätzt man nun getrennt für die Variablen *Bekanntheit* und *Markenimpact* in den vier Werbeträgern sowohl ein logistisches als auch ein lineares Modell und vergleicht den Einfluss des Aufwandes und der Zeit, so ist ersichtlich, dass sich die Oberflächen-Plots sehr ähnlich sind. Dies resultiert daraus, dass beide Variablen sehr symmetrisch um ihren Mittelwert und annähernd normalverteilt sind (siehe Abbildung 3.2). Um aber trotzdem zu vermeiden, dass unmögliche Prozentsätze bzw. Anteile geschätzt werden, werden für die Variablen *Bekanntheit* und *Markenimpact* logistische Modelle mit Logit-Link angepasst.
Die unabhängigen Variablen in diesen Modellen sind *Aufwand*, *Länge* und *Zeit*. Einige beschreibende Maßzahlen werden in den folgenden Tabellen und Abbildungen gegeben.

Die Variable *Aufwand* wird in tausend Euro gemessen. Man sieht sowohl in Abbildung 3.3 als auch in Tabelle 3.5, dass die Verteilung linksschief ist (Schiefe = 2,16). Es gibt also viele „billige“ Werbungen und nur wenige „teure“. Der Median von 104,089 liegt deutlich unter dem Mittelwert von 134,9. Der „Infoscreen“ ist der im Durchschnitt günstigste Werbeträger, gefolgt von „Hörfunk“, viel teurer ist der Werbeträger „TV“, an der Spitze stehen Print-Werbungen. Das teuerste Sujet kostete 1.178.520,8 €.

Die Variable *Länge* wurde, wie schon im vorigen Kapitel beschrieben, in Kategorien eingeteilt. Im Werbeträger „Print“ konnte sie nicht erhoben werden. Die Aufteilung der Variablen auf die einzelnen Werbeträger zeigt Tabelle 3.6.



(a) *Bekanntheit* im Werbeträger „HF“

(b) *Markenimpact* im Werbeträger „HF“

(c) *Gefälligkeit* im Werbeträger „HF“

(d) *Bekanntheit* im Werbeträger „IS“

(e) *Markenimpact* im Werbeträger „IS“

(f) *Gefälligkeit* im Werbeträger „IS“

(g) *Bekanntheit* im Werbeträger „PR“

(h) *Markenimpact* im Werbeträger „PR“

(i) *Gefälligkeit* im Werbeträger „PR“

(j) *Bekanntheit* im Werbeträger „TV“

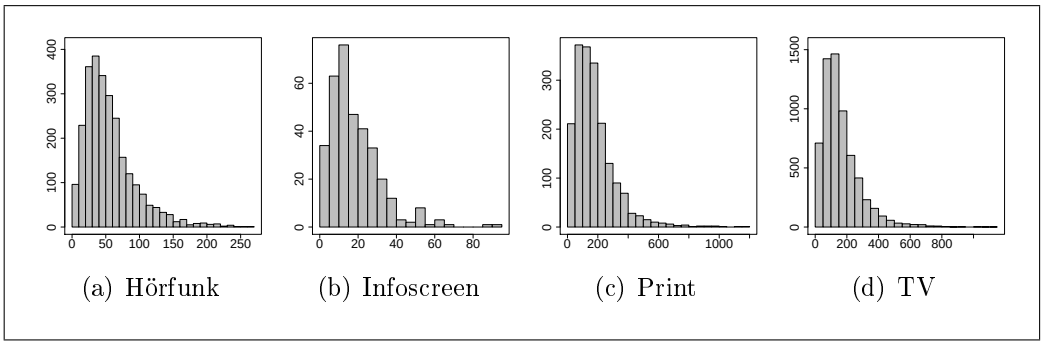
(k) *Markenimpact* im Werbeträger „TV“

(l) *Gefälligkeit* im Werbeträger „TV“

Abb. 3.2: Histogramme der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* mit Normalverteilungskurve in den einzelnen Werbeträgern

Tab. 3.5: Deskriptive Maßzahlen der Variablen *Aufwand* in den Werbeträgern

	Hörfunk	*Infoscreen*	*Print*	*TV*	*Gesamt*
n	2.626	346	1.894	6.268	11.134
NA's	7	0	21	11	39
Minimum	0,1750	0,5725	0,7400	0,6825	0,1750
1. Quartil	29,2230	9,1500	86,2200	80,8776	51,7533
Median	47,3248	14,8950	149,3349	133,0431	104,0893
Mittelwert	55,6677	18,5543	177,7280	161,5771	134,9007
3. Quartil	71,5705	24,6643	231,4066	207,9161	181,4085
Maximum	260,9311	91,6044	1.178,5208	1.146,5380	1.178,5208
Stabw	37,9370	13,4360	136,2383	119,6085	119,0310
Schiefe	1,5809	1,7330	2,1063	1,9703	2,1640



(a) Hörfunk (b) Infoscreen (c) Print (d) TV

Abb. 3.3: Histogramm der Variablen *Aufwand*, getrennt nach Werbeträgern

Es gibt im Werbeträger „Hörfunk“ 6 fehlende, im Werbeträger „Infoscreen“ keinen und im Werbeträger „TV“ 36 fehlende Werte.

Die Variable *Zeit* wird nach der Formel $Zeit = Jahr + Monat/12$ berechnet, wobei das Jahr 1998 mit dem Wert codiert wird. Der Jänner 1998 hat zum Beispiel den Wert $1 + 1/12 = 1{,}08\dot{3}$. Eine genaue Aufschlüsselung, welche Werte welches Jahr und Monat bedeuten, gibt Tabelle 3.7. Die Variable *Zeit* hat bei keinem Werbeträger fehlende Werte.

In allen folgenden Modellen werden die Variablen *Aufwand* und *Zeit* als restringierte kubische Splines modelliert. Die fünf Stützstellen werden als Quantile wie in Tabelle 2.2 gewählt. Die Variable *Länge* wird durch Dummycodierung mit der Referenzkategorie „≤ 15“ ins Modell genommen. Außerdem werden die linearen einfachen Wechselwirkungen zwischen den Varia-

Tab. 3.6: Häufigkeiten der Variablen *Länge*, getrennt nach Werbeträgern

	HF	*IS*	*TV*	*Gesamt*
≤15	103	305	869	1.277
15-20	916	34	1.705	2.655
20-25	1.005	1	1.040	2.046
25-30	462	5	1.943	2.410
30>	98	1	442	541
Tandem	43	0	244	287
NA	6	0	36	42
Summe	2.633	346	6.279	9.258

Tab. 3.7: Bedeutung der Variablen *Zeit*

	1998	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*
Jän.	1,083	2,083	3,083	4,083	5,083	6,083	7,083	8,083	9,083	10,083
Feb.	1,167	2,167	3,167	4,167	5,167	6,167	7,167	8,167	9,167	10,167
März	1,250	2,250	3,250	4,250	5,250	6,250	7,250	8,250	9,250	10,250
Apr.	1,333	2,333	3,333	4,333	5,333	6,333	7,333	8,333	9,333	10,333
Mai	1,417	2,417	3,417	4,417	5,417	6,417	7,417	8,417	9,417	10,417
Juni	1,500	2,500	3,500	4,500	5,500	6,500	7,500	8,500	9,500	10,500
Juli	1,583	2,583	3,583	4,583	5,583	6,583	7,583	8,583	9,583	
Aug.	1,667	2,667	3,667	4,667	5,667	6,667	7,667	8,667	9,667	
Sept.	1,750	2,750	3,750	4,750	5,750	6,750	7,750	8,750	9,750	
Okt.	1,833	2,833	3,833	4,833	5,833	6,833	7,833	8,833	9,833	
Nov.	1,917	2,917	3,917	4,917	5,917	6,917	7,917	8,917	9,917	
Dez.	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	

blen berücksichtigt. Nähme man die Wechselwirkungen zwischen den Spline-Variablen auch hinzu, wäre das Modell zu komplex. Außerdem ändert sich die Form des Zusammenhangs nur gering. In einigen Bereichen ist die Ausprägung ein wenig stärker oder schwächer, dies ist vor allem an den Rändern der Fall.

3.3.1 Einfluss im Werbeträger „Hörfunk“

Im Werbeträger „Hörfunk“ wurden, wie man an Tabelle 3.4 aus Kapitel 3.2 sieht, sowohl *Aufwand* als auch *Länge* erhoben. Bei der Variablen *Aufwand* gibt es 7, bei der Variablen *Länge* 6 fehlende Werte, bei einem Sujet fehlen beide Variablen. Der Datensatz hat insgesamt 2.633 Beobachtungen, wovon für die Analyse $2.621 = 2.633 - 7 - 6 + 1$ übrig bleiben.

Tab. 3.8: Zahlenwerte zu nebenstehender Abbildung

	n	*Bek.*
Aufwand		
[0,175; 29,2)	656	38,63
[29,221; 47,3)	655	40,05
[47,325; 71,6)	655	40,52
[71,640; 260,9]	655	43,47
Länge		
≤ 15	101	40,09
15-20	915	39,87
20-25	1.002	41,54
25-30	462	40,09
>30	98	43,47
Tandem	43	38,41
Zeit		
[1,08; 4,17)	659	40,37
[4,17; 7,00)	677	39,84
[7,00; 8,75)	635	40,49
[8,75; 10,50]	650	42,01
Overall	2.621	40,67

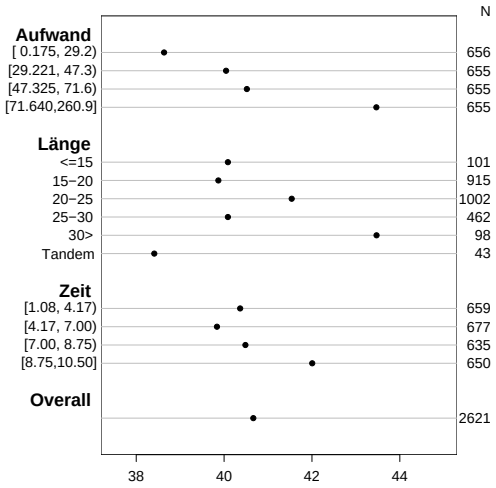


Abb. 3.4: Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“

Bekanntheit

Einen ersten Eindruck, wie sich die Variablen zueinander verhalten, erhält man in Tabelle 3.8 und in Abbildung 3.4.

Die mittlere Bekanntheit steigt bei steigendem Aufwand, und zwar im mittlere Bereich nicht so stark, gegen Ende hin dann deutlicher. Die Länge hat keinen so deutlichen Einfluss. Die Mittelwerte in den Kategorien „≤ 15“, „15 bis 20“ und „25 bis 30“ sind in etwa gleich. Tandem-Spots weisen im Mittel die geringste Bekanntheit auf. Die Variable *Zeit* verhält sich nicht linear, sie sinkt nach dem ersten Quartil und steigt danach.

Die Hypothesen im Modell werden mit dem Wald- und dem Likelihood-Ratio-Test überprüft. Die Ergebnisse der beiden Tests sieht man in Tabelle 3.9. Man erkennt hierbei, dass sich die beiden Teststatistiken nur minimal unterscheiden.

Alle drei im Modell inkludierten Variablen haben einen Einfluss auf die Bekanntheit. Ihre Teststatistiken sind signifikant. Bei den kontinuierlichen Variablen ist der Einfluss nichtlinear. Außerdem sind alle Wechselwirkungen

Tab. 3.9: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Bekanntheit* im Werbeträger „Hörfunk“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	10	749,0746	0,0000	752,3426	0,0000
Aufwand nichtlinear	3	14,5938	0,0022	14,5840	0,0022
Zeit	10	744,9654	0,0000	748,9659	0,0000
Zeit nichtlinear	3	383,4423	0,0000	383,8186	0,0000
Länge	15	294,0579	0,0000	296,3158	0,0000
Aufwand × *Zeit*	1	118,5102	0,0000	118,8907	0,0000
Aufwand × *Länge*	5	55,6895	0,0000	55,7153	0,0000
Zeit × *Länge*	5	87,6341	0,0000	87,8368	0,0000
TOTAL nichtlinear	6	399,7640	0,0000	400,1589	0,0000
TOTAL WW	11	293,1013	0,0000	294,6855	0,0000
TOTAL nichtlinear+WW	17	814,1715	0,0000	818,8878	0,0000
Regression	24	1641,6338	0,0000	1661,3806	0,0000

signifikant. Dies bedeutet, dass sich die Bekanntheit mit dem Aufwand unterschiedlich für die Längenkategorien und über die Zeit hinweg ändert. Analog gilt dies auch für die Zeit.
Das Modell, in dem der *Aufwand* als restringierter kubischer Spline inkludiert ist, hat eine Likelihood-Ratio von 1.661 bei 24 Freiheitsgraden.

Es zeigt sich kein allzu unterschiedliches Bild für die Längenkategorien. Nur „Tandem“-Spots lassen einen differierenden Einfluss erkennen. Dieser wird im Anschluss separat beschrieben. Der Aufwand hat in den Anfangsjahren einen durchgehenden positiven Einfluss. Dieser relativiert sich ab dem Jahr 2002, nur in der Kategorie „$\leq$ 15“ flacht er ab dem Jahr 2002 ab und ist in den letzten Jahren sogar negativ. In den restlichen Kategorien ist der Einfluss stets positiv, wird aber in den letzten Jahren weniger. Die stärkste Steigung in den Anfangsjahren zeigt die Kategorie „länger als 30“.
Der Einfluss der Zeit ist nichtlinear, er ändert also seine Richtung. Bei geringem Aufwand ist er bis zum Jahr 2001 leicht negativ, am größten in der Kategorie „länger als 30“. Denn steigt er bis etwa zum Jahr 2005, woraufhin er abflacht, in der Kategorie „länger als 30“ sogar leicht sinkt. Ist der Aufwand hoch, so ist das anfängliche Sinken der Bekanntheit bis 2001 stärker, dann steigt er etwa ein bis eineinhalb Jahre an und sinkt letztlich wieder.
Die Bekanntheit von „Tandem“-Spots wird bei fortschreitender Zeit größer. Nur bei einem Aufwand von knapp über 200.000 € sinkt sie in den ersten drei Jahren ein wenig. Bei geringem Aufwand steigt sie stetig. Die Bekanntheit steigt bei steigendem Aufwand von etwa 0,2 bis auf 0,4 in den Anfangsjah-

ren, in den letzten Jahren von etwa 0,39 auf 0,55. Das Minimum nimmt die Funktion bei geringem *Aufwand* in den ersten Jahren an, ihr Maximum in der „gegenüberliegenden Ecke“ bei hohem Aufwand und fortgeschrittener Zeit.

Betrachtet man die Wald- und Likelihood-Ratio-Tests für die einzelnen Zielgruppen, so hat bei den weiblichen Befragten und bei jenen, die zwischen 30 und 49 Jahre alt sind, aber auch bei den über 50-Jährigen der Aufwand nur einen linearen Einfluss. Außerdem ist in der Alterklasse „14 bis 29“ die Wechselwirkung zwischen *Aufwand* und *Zeit* nicht signifikant.
Bei den weiblichen Befragten sind die Hörfunk-Spots im ersten Jahr bei hohem Aufwand in allen Längenkategorien bekannter als bei den männlichen Befragten, oder wenn man die Befragten gesamt betrachtet. In den anderen Bereichen gibt es kaum Unterschiede.
Bei der Alterklasse der 14- bis 29-Jährigen sinkt sogar die Bekanntheit mit steigendem Aufwand bei Hörfunk-Spots, die kürzer als 15 Sekunden sind. Dies ist aber nur in dieser einen Längenkategorie zu beobachten. In den restlichen Kategorien steigt die Bekanntheit aber sehr wohl bei steigendem Aufwand. Vor allem im ersten Jahr sind die Bekanntheitswerte bei hohem Aufwand deutlich größer.
Die Altersgruppe der 30- bis 49-Jährigen weist die geringsten Unterschiede zu den Gesamtbefragten auf. Der Verlauf bei wachsendem Aufwand ist aber, wie schon die Teststatistiken gezeigt haben, annähernd linear. Die Gruppe der über 50-Jährigen weist bis zum Jahr 2000 und einem Aufwand von 50.000 € niedrigere Bekanntheitswerte auf. Ab dem Jahr 2005 weisen die beiden zuletzt beschriebenen Altersklassen, über den gesamten Aufwand verteilt, kaum Unterschiede auf.

Markenimpact

Die Variable *Markenimpact* hat im Datensatz „Hörfunk“ zwei fehlende Werte. In diesem Fall verringert sich die Anzahl an vollständigen Beobachtungen von 2.621 auf 2.619.

Sowohl in Abbildung 3.5 als auch in Tabelle 3.10 erkennt man, dass der mittlere Markenimpact bei steigendem Aufwand nach dem ersten Quartil sinkt, dann steigt. Den höchsten Mittelwert hat die Längenkategorie „länger als 30“, den geringsten die Kategorie “15 bis 20“. Der Mittelwert im ersten, zweiten und vierten Quartil der Zeit ist fast konstant, nur im dritten Quartil ist er ein wenig kleiner.
Dass alle drei Variablen einen Einfluss auf den Markenimpact eines Sujets

Tab. 3.10: Zahlenwerte zu nebenstehender Abbildung

	n	*Mar.*
Aufwand		
[0,175; 29,2)	655	29,13
[29,229; 47,3)	655	28,85
[47,325; 71,7)	655	31,05
[71,670;260,9]	654	32,74
Länge		
≤ 15	101	27,79
15-20	914	27,68
20-25	1.001	32,33
25-30	462	31,64
>30	98	36,32
„Tandem“	43	25,12
Zeit		
[1,08; 4,17)	659	31,12
[4,17; 7,00)	677	31,15
[7,00; 8,75)	635	28,37
[8,75;10,50]	648	31,04
Overall	2.619	30,44

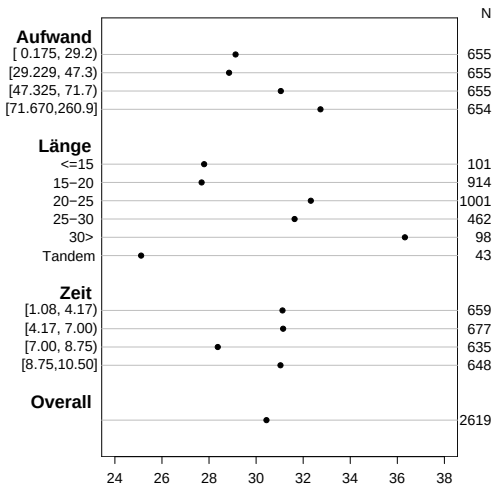


Abb. 3.5: Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“

haben, zeigt Tabelle 3.11. *Aufwand* und *Zeit* haben einen nichtlinearen Einfluss. Die Wechselwirkung des *Aufwandes* mit der *Zeit* hat keinen sehr großen Einfluss, ihr P-Wert liegt nahe bei 0,05. Sie wird aber weiterhin im Modell belassen.

Die Funktion hat in allen Längenkategorien eine ähnliche Form. Nur bei Tandem-Spots scheint der Aufwand eine etwas andere Wirkung zu haben. Wie schon aus Abbildung 3.4 und Tabelle 3.8 hervorgegangen ist, liegen die Kategorien „$\leq$ 15“ und „15 bis 20“ etwa auf gleicher Höhe, dann folgen die Kategorien „20 bis 25“ und „25 bis 30“, die höchsten Werte verzeichnet die Kategorie „länger als 30“.
Der Markenimpact sinkt bis zu einem Aufwand von 40.000 € relativ stark ab, steigt dann bis etwa 70.000 € genauso stark wieder an und bleibt kanpp unter dem Anfangswert. Daraufhin steigt er durchgehend, jedoch nur noch sehr leicht. Den stärksten Anstieg hat die Kategorie „15 bis 20“. Dieser Einfluss ändert sich über die Zeit hinweg kaum, worauf auch schon Tabelle 3.11 schließen lässt.
Der Einfluss der Zeit ist nichtlinear, jedoch annähernd gleich bei variierendem Aufwand. Bis Ende 2000 steigt der Markenimpact, daraufhin sinkt er bis

Tab. 3.11: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „Hörfunk“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	10	529,1209	0,0000	527,9706	0,0000
Aufwand nichtlinear	3	201,4179	0,0000	200,3498	0,0000
Zeit	10	504,2621	0,0000	503,3887	0,0000
Zeit nichtlinear	3	320,7914	0,0000	320,8239	0,0000
Länge	15	1109,0259	0,0000	1116,2610	0,0000
Aufwand × *Zeit*	1	4,0901	0,0431	4,0878	0,0432
Aufwand × *Länge*	5	20,6463	0,0009	20,6484	0,0009
Zeit × *Länge*	5	132,7302	0,0000	132,8783	0,0000
TOTAL nichtlinear	6	513,6415	0,0000	513,1306	0,0000
TOTAL WW	11	170,0759	0,0000	170,4519	0,0000
TOTAL nichtlinear+WW	17	734,7535	0,0000	734,9664	0,0000
Regression	24	2239,6602	0,0000	2263,7301	0,0000

Ende 2004 und steigt dann wieder auf den Wert von 1998. In den Kategorien „20 bis 25“ und „25 bis 30“ steigt er über den Wert von 1998.
„Tandem“-Spots zeigen eine ähnliche Struktur, wenn sich der Aufwand verändert, der Markenimpact sinkt jedoch mit der Zeit. Die ersten zwei bis drei Jahre ist er annähernd konstant, dann sinkt er und flacht gegen Ende hin ab.

Wald- und Likelihood-Ratio-Tests zeigen für die Zielgruppe der Männer, dass die Wechselwirkung zwischen *Aufwand* und *Länge* nicht signifikant ist. Genauso ist bei den Frauen die Wechselwirkung zwischen *Aufwand* und *Zeit* nicht signifikant.
Die männlichen Befragten konnten sich in den letzten Jahren bei Hörfunk-Spots mit hohem Aufwand eher an diese erinnern als Frauen, bei denen genau in diesem Bereich ein geringerer Markenimpact erzielt wurde. Über die Zeit hinweg gibt es kaum Unterschiede. Nur bei Tandem-Spots fällt der Markenimpact über die Zeit stärker als bei Frauen.
In der Zielgruppe der 14- bis 29-Jährigen gibt es keine signifikante Wechselwirkung zwischen *Aufwand* und *Länge*. Dies zeigt auch die graphische Überprüfung. Der Markenimpact steigt hier bei jeder Längenkategorie mit steigendem Aufwand.
Auch in der Klasse der 30- bis 49-Jährigen steht der Aufwand mit den Längenkategorien nicht in Wechselwirkung. Steigender Aufwand bedeutet auch in dieser Gruppe höheren Markenimpact, jedoch steigen die Werte nicht so weit an wie in der vorher beschriebenen Gruppe.

Dass es in der Gruppe der über 50-Jährigen keine gegenseitige Beeinflussung von Aufwand und Zeit gibt, das heißt, dass sich mit steigendem Aufwand der Einfluss über die Zeit ändert, zeigt schon die nicht signifikante Wechselwirkung zwischen diesen beiden Variablen. Der Aufwand wirkt hier nicht so positiv. Der Markenimpact bleibt in den meisten Kategorien bei steigendem Aufwand konstant, fällt in der Kategorie „≤ 15“ sogar leicht ab und steigt in der Kategorie „15 bis 20“ ganz leicht an.
In der Kategorie „Tandem“ unterscheiden sich die drei Altersgruppen am deutlichsten. Während in der mittleren Klasse der Markenimpact entlang des Aufwandes konstant ist und bei fortschreitender Zeit nur leicht abfällt, fällt er in den Gruppe der 14- bis 29-Jährigen bei steigender Zeit von einem Wert im Jahr 1998 zwischen 55 % und knapp über 60 % bis auf knapp über 20 % bzw. 30 %. Die Zielgruppe der über 50-Jährigen liegt im Mittelfeld.
Höhere Längenkategorien weisen in allen Zielgruppen genauso wie in der gesamten Gruppe höhere Werte in der Variablen *Markenimpact* auf.

Gefälligkeit

Bei der Variablen *Gefälligkeit* gibt es keine fehlenden Werte. Es werden also bei der Auswertung nur die zwölf Beobachtungen, die fehlende Werte in den unabhängigen Variablen haben, gelöscht. Demnach gibt es wie bei der Auswertung der Variablen *Bekanntheit* 2.621 vollständige Beobachtungen.

Abbildung 3.6 und Tabelle 3.12 zeigen, dass mit steigendem Aufwand die mittlere Gefälligkeit abnimmt, nur zwischen dem ersten und dem zweiten Quartil gibt es einen Anstieg. Der Rückgang ist aber nicht sehr groß. Auch die *Länge* zeigt ein ambivalentes Bild. Die Mittelwerte in den Kategorien „≤ 15“, „15 bis 20“, „25 bis 30“ und „länger als 30“ sind annähernd gleich. Die restlichen beiden verzeichnen größere Mittelwerte, wobei die Tandem-Spots mit 5,40 den höchsten Gefälligkeitswert aufweisen. In den ersten drei Quartilen der Zeit steigt die mittlere Gefälligkeit von 4,76 auf 5,5, im vierten Quartil ist sie kleiner.

Betrachtet man mehrere Koeffizienten gemeinsam und testet deren Einfluss auf die Gefälligkeit eines Sujets, so erhält man die Ergebnisse in Tabelle 3.13. In dieser Tabelle sieht man, dass die Variable *Aufwand* keinen Einfluss (P-Wert: 0,4995) darauf hat, ob jemandem ein Sujet gefällt. Die Wechselwirkungen, einzeln (P-Werte: 0,5089; 0,8599 und 0,4124) sowie auch gemeinsam betrachtet (P-Wert: 0,7388), sind nicht signifikant. Zeit und Länge haben aber eine Wirkung auf die Gefälligkeit. Um die Art des Einflusses darzustel-

Tab. 3.12: Zahlenwerte zu nebenstehender Abbildung

	n	*Gef.*
Aufwand		
[0,175; 29,2)	656	5,18
[29,221; 47,3)	655	5,21
[47,325; 71,6)	655	5,18
[71,640;260,9]	655	5,13
Länge		
≤ 15	101	5,09
15-20	915	5,13
20-25	1.002	5,24
25-30	462	5,13
>30	98	5,13
„Tandem“	43	5,40
Zeit		
[1,08; 4,17)	659	4,76
[4,17; 7,00)	677	5,13
[7,00; 8,75)	635	5,50
[8,75;10,50]	650	5,32
Overall	2.621	5,17

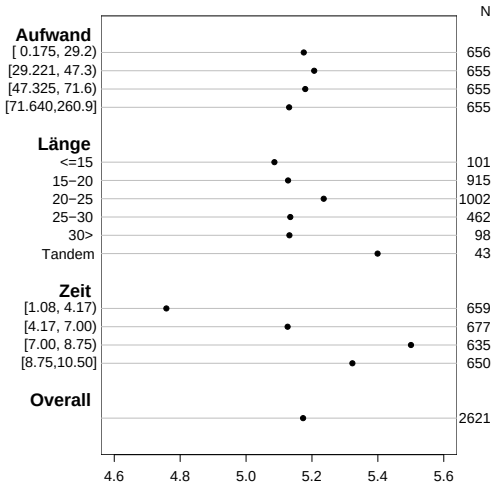


Abb. 3.6: Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“

len, bleiben sowohl die Variable *Aufwand* als auch die Wechselwirkungen im Modell. Man betrachtet die Funktion aber nur für die beiden interessierenden Variablen, nämlich *Zeit* und *Länge*, und setzt für die Variable *Aufwand* ihren Median ein. Das Bild dieses Einflusses unterscheidet sich nicht allzu sehr von jenem, wenn im Modell nur die beiden interessierenden Variablen berücksichtigt werden und deren Einfluss graphisch dargestellt wird. Man kommt in solch einem Modell aber zu verzerrten Schätzungen und schätzt typischerweise zu kleine Konfidenzintervalle und P-Werte.

Abbildung 3.7 zeigt die Veränderung der Gefälligkeit über die Zeit für alle Längenkategorien. Dabei steht jede Kurve für eine Längenkategorie, auf der x-Achse ist die *Zeit* eingetragen. Alle fünf Kurven zeigen ein ähnliches Bild.

Bis zum Jahr 2000 steigen sie ganz leicht, ab dem Jahr 2000 bis Anfang des Jahres 2004 ist die Steigung größer, dann fällt die Gefälligkeit. In den ersten vier Jahren trennt die Kurven noch ein Abstand von höchstens 0,4; mit fortschreitender Zeit rücken sie immer näher aneinander, ab dem Jahr 2006 unterscheiden sie sich nur noch etwa um den Wert 0,1.

Tab. 3.13: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Gefälligkeit* im Werbeträger „Hörfunk“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	10	9,2743	0,5063	9,3469	0,4995
Aufwand nichtlinear	3	4,1014	0,2507	4,1376	0,2470
Zeit	10	1018,1640	0,0000	867,2595	0,0000
Zeit nichtlinear	3	225,9807	0,0000	218,7676	0,0000
Länge	15	43,5659	0,0001	43,6205	0,0001
Aufwand × *Zeit*	1	0,4322	0,5109	0,4363	0,5089
Aufwand × *Länge*	5	1,9036	0,8623	1,9213	0,8599
Zeit × *Länge*	5	4,9851	0,4177	5,0283	0,4124
TOTAL nichtlinear	6	230,2823	0,0000	222,7598	0,0000
TOTAL WW	11	7,6507	0,7442	7,7130	0,7388
TOTAL nichtlinear+WW	17	238,5872	0,0000	230,4502	0,0000
Regression	24	1086,7659	0,0000	916,5434	0,0000

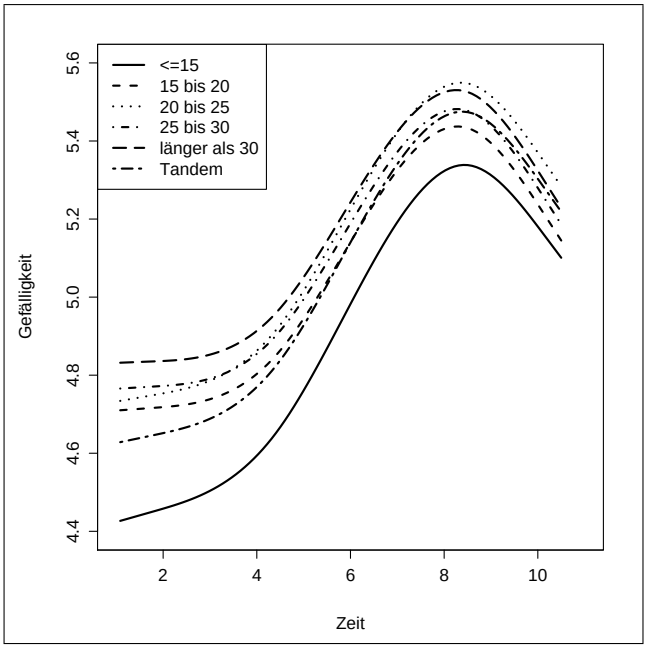


Abb. 3.7: Einfluss von *Zeit* und *Länge* auf *Gefälligkeit* im Werbeträger „Hörfunk“

Die geringsten Werte in der Gefälligkeit haben die Sujets mit der kürzesten Länge, dann folgen die „Tandem"-Spots, die aber etwa ab dem Jahr 2003 über der Kategorie „15 bis 20" liegen, die anfangs noch höher eingestuft wurden. In den ersten Jahren weisen Sujets der Längenkategorie „20 bis 25" die vierthöchsten Werte auf, ab dem Jahr 2004 weisen diese Spots allerdings die höchsten Werte auf. Die längsten Hörfunk-Werbungen gefallen den Befragten durchschnittlich am besten, dieser Trend ändert sich ab dem Jahr 2004 und die durchschnittliche Bewertung dieser Kategorie fällt unter die Bewertung der Spots mit der Länge „20 bis 25".

In allen Zielgruppen hat die Variable *Aufwand* keinen Einfluss, auch die drei Wechselwirkungen sind nicht signifikant.
Ab dem Jahr 2003 gibt es kaum Unterschiede zwischen der gesamten Stichprobe und der Zielgruppe der Männer. In den Jahren davor hat bei den Männern die Längenkategorie „$\leq$ 15" geringere Gefälligkeitswerte, die Kategorie „länger als 30" höhere Werte. Außerdem weist die „Tandem"-Kategorie bis zum Jahr 2002 höhere Werte als die Kategorien „15 bis 20", „20 bis 25" und „25 bis 30" auf.
Bei den Frauen zeigt sich ein anderes Bild. Hier haben die Kategorien „$\leq$ 15" und „Tandem" die geringsten Werte und liegen nah beieinander, ab dem Jahr 2004 entfernen sie sich voneinander. Die restlichen vier Längenkategorien weisen bis zum Jahr 2002 kaum Unterschiede auf. Dann steigen die Gefälligkeitswerte in den Kategorien „20 bis 25" und „länger als 30" schneller an und liegen bis 2007 über den Werten der restlichen Kategorien, während die Kategorien „15 bis 20" und „25 bis 30" nicht so schnell wachsen und im Mittelfeld bleiben.
In der Altersklasse „14 bis 29" liegt die Kategorie „länger als 30" in der Bewertung bis zum Jahr 2004 deutlich über den restlichen Kategorien, dann weist sie ähnliche Werte wie die Kategorien „15 bis 20", „25 bis 30" und „Tandem" auf. Die Längenkategorie „$\leq$ 15" hat über den gesamten Zeitverlauf eine niedrigere Bewertung.
Die Zielgruppe „30 bis 49" unterscheidet sich von der Gesamtbewertung nur in der Kategorie „Tandem", die in den ersten fünf Jahren eine niedrigere Bewertung hat. In der restlichen Zeit unterscheiden sich die Gruppen und Kategorien nur minimal.
Bei den 50-jährigen Befragten liegt die Längenkategorie „$\leq$ 15" deutlich unter den längeren Spots. Die restlichen Kategorien sind in den ersten zwei Jahren ident, dann bilden sich zwei Gruppen, die sich jeweils gleich verhalten. Zu der oberen Gruppe zählen die Kategorien „20 bis 25" und „länger als 30", zur unteren die Kategorien „15 bis 20", „25 bis 30" und „Tandem".
Alle Kurven haben aber über die Zeit hinweg einen ähnlichen Verlauf.

3.3.2 Einfluss im Werbeträger „Infoscreen“

Beim Werbeträger „Infoscreen“ gibt es keine fehlenden Werte, weder bei einer unabhängigen Variablen noch bei den drei abhängigen Variablen. Allerdings ist die Aufteilung in den Längenkategorien, wie Tabelle 3.6 zeigt, sehr ungünstig. In den Kategorien „20 bis 25“, „länger als 30“ und „Tandem“ können keine Parameter geschätzt werden, für die Analyse bleiben dann noch die drei restlichen Kategorien, bei denen die Aufteilung 89 %, 10 % und 1 % ist. Passt man mit diesen Kategorien das Modell an, so bemerkt man, dass sich die Schätzungen in der Längenkategorie „25 bis 30“, in der es ja nur 5 Beobachtungen gibt, sehr stark von den anderen beiden Kategorien unterscheiden. Der Einfluss von Aufwand und Zeit ist vor allem auf die Variable *Bekanntheit* in dieser Kategorie vollkommen anders als in den restlichen Kategorien. Ob dieser Einfluss allerdings aussagekräftig ist, sei dahingestellt.
Betrachtet man die kontinuierliche Variable *Länge*, so fällt auf, dass 37 % der Einträge auf den Wert 10 und 33 % auf den Wert 15 fallen. Aus diesen Gründen werden für diesen Werbeträger eigene Kategorien gebildet. Da eine Trennung in drei oder mehr Bereiche wieder zu einer sehr ungleichen Aufteilung führen würde, teilt man die Spots in zwei Klassen, und zwar in jene, die maximal 10, und in jene, welche länger als 10 Sekunden dauern. In der ersten Kategorie befinden sich dann 181, dies sind in etwa 52 %, und in der längeren demnach 165 Beobachtungen oder knappe 48 %.

Bekanntheit

Einen ersten Einblick, wie sich die Variable *Bekanntheit* im Werbeträger „Infoscreen“ verhält, vermitteln Tabelle 3.14 und Abbildung 3.8. An beiden erkennt man, dass sowohl bei steigendem Aufwand als auch bei steigender Länge die Sujets im Mittel besser bekannt sind.
Zwischen dem ersten und dem zweiten Quartil steigt die mittlere Bekanntheit, ab Juni 2004 fällt sie und hat ihren niedrigsten Wert im letzten Quartil. Das Gesamtmittel im Werbeträger „Infoscreen“ liegt bei 39,58.

In der logistischen Regression ist bei den Haupteffekten nur die erste Variable des Aufwands nicht signifikant. Gesamt gesehen hat der Aufwand aber sehr wohl einen Einfluss (siehe Tabelle 3.15). Auch die Längenkategorie „länger als 10“ unterscheidet sich nicht signifikant von der Referenzkategorie. Ihre Bekanntheitswerte liegen aber wegen des positiven Vorzeichens des Koeffizienten über den Werten der Kategorie „≤ 10“ (vgl. Abbildung 3.9). Außerdem sind die Wechselwirkungen von *Aufwand* und *Zeit* mit der *Länge* nicht signifikant. Die Signifikanzen stimmen mit den P-Werten in Tabelle 3.15 überein.

Tab. 3.14: Zahlenwerte zu nebenstehender Abbildung

	n	*Bek.*
Aufwand		
[0,573; 9,54)	88	34,59
[9,537;15,03)	85	36,96
[15,030;24,88)	87	40,62
[24,882;91,60]	86	46,21
Länge		
≤ 10	181	36,51
>10	165	42,94
Zeit		
[4,83; 5,92)	95	40,40
[5,92; 7,42)	84	43,31
[7,42; 8,92)	83	41,09
[8,92;10,33]	84	33,42
Overall	346	39,58

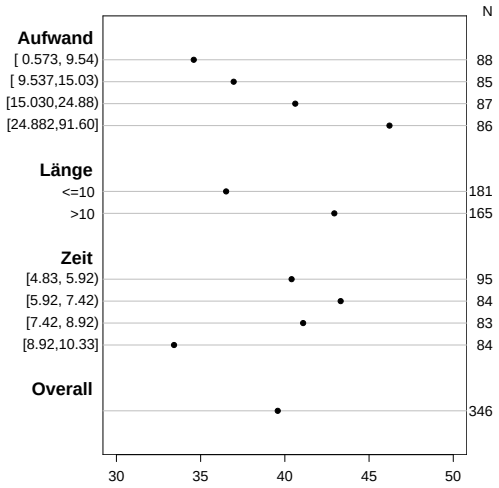


Abb. 3.8: Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“

Das Ergebnis der logistischen Regression schaut folgendermaßen aus:

```
General Linear Model
glmD(formula = cbind(floor(R_TOTAL/100 * 150), 150 - floor(R_TOTAL/100 *
    150)) ~ rcs(Aufwand, 5) + rcs(zeit, 5) + Laenge.kat + Aufwand %ia%
    zeit + Aufwand %ia% Laenge.kat + zeit %ia% Laenge.kat,
    family = binomial, data = IS.reg)
Model L.R.       d.f.          P
       932         12          0
                              Coef      S.E.    Wald Z     P
Intercept                 -1.165802  0.2748412   -4.2   2.218e-05
Aufwand                   -0.014641  0.0089212   -1.6   1.008e-01
Aufwand'                   0.220155  0.1232029    1.8   7.395e-02
Aufwand''                 -0.762542  0.3499985   -2.2   2.935e-02
Aufwand'''                 0.749532  0.2881154    2.6   9.282e-03
zeit                       0.116208  0.0497733    2.3   1.956e-02
zeit'                     -1.870091  0.3948828   -4.7   2.182e-06
zeit''                     4.735487  0.8955335    5.3   1.237e-07
zeit'''                   -6.039628  0.8431117   -7.2   7.865e-13
Laenge.kat>10              0.112247  0.0980384    1.1   2.522e-01
Aufwand %ia% zeit          0.003167  0.0004577    6.9   4.520e-12
Aufwand %ia% Laenge.kat   -0.002228  0.0018712   -1.2   2.338e-01
zeit %ia% Laenge.kat       0.010856  0.0119888    0.9   3.652e-01
```

In Tabelle 3.15 sieht man, dass alle Variablen den Bekanntheitsgrad beeinflussen, die kontinuierlichen Variablen *Aufwand* und *Zeit* sogar nichtlinear.

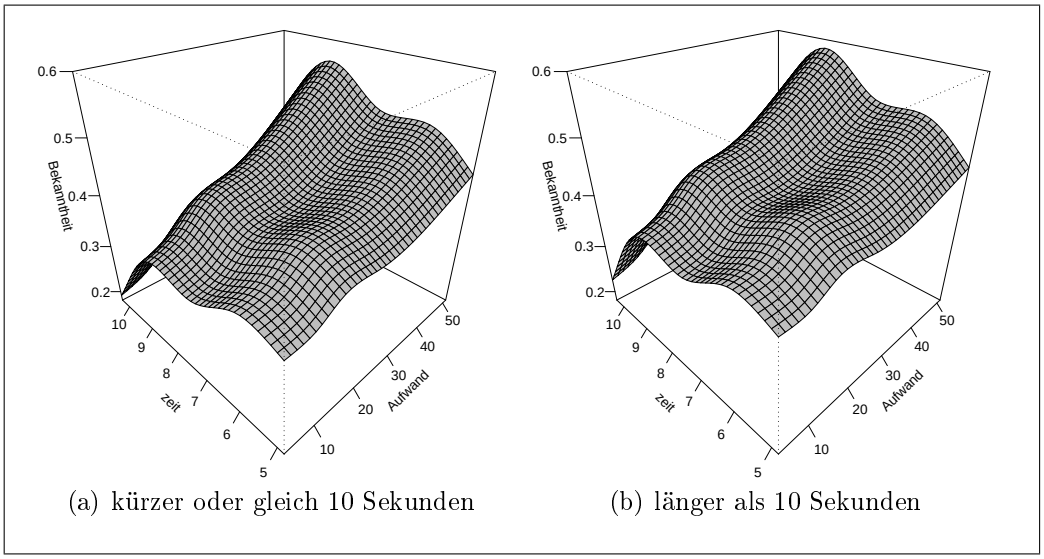
Tab. 3.15: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Bekanntheit* im Werbeträger „Infoscreen“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	6	347,3565	0,0000	350,5998	0,0000
Aufwand nichtlinear	3	13,4585	0,0037	13,4646	0,0037
Zeit	6	386,6191	0,0000	405,6951	0,0000
Zeit nichtlinear	3	212,2543	0,0000	215,9557	0,0000
Länge	3	61,0095	0,0000	60,8845	0,0000
Aufwand × *Zeit*	1	47,8849	0,0000	47,9091	0,0000
Aufwand × *Länge*	1	1,4176	0,2338	1,4182	0,2337
Zeit × *Länge*	1	0,8200	0,3652	0,8201	0,3651
TOTAL nichtlinear	6	219,9796	0,0000	223,7167	0,0000
TOTAL WW	3	63,3921	0,0000	63,4772	0,0000
TOTAL nichtlinear+WW	9	266,4018	0,0000	272,0549	0,0000
Regression	12	893,2965	0,0000	932,2344	0,0000

Die Wechselwirkungen mit der Variablen *Länge* sind allerdings nicht einflussreich. Sie bleiben aber bei der Schätzung dennoch im Modell.

Die Form des Einflusses unterscheidet sich in den beiden Längenkategorien nur leicht. Darauf lassen auch schon die nicht signifikanten Wechselwirkungen in Tabelle 3.15 schließen. Die Werte in der längeren Kategorie liegen knapp über den Werten der Referenzkategorie. Dies sieht man auch in Abbildung 3.9.
Im Laufe des Jahres 2002 steigt die Bekanntheit sowohl bei geringem als auch bei hohem Aufwand, im darauf folgenden Jahr bleibt sie bei hohem Aufwand nahezu unverändert, bei niedrigem Aufwand sinkt sie leicht. Daraufhin steigt die Bekanntheit bei geringem Aufwand sehr wenig, bei hohem Aufwand innerhalb von zwei Jahren von knapp unter 0,5 auf 0,56. Im letzten Jahr sinkt sie auf den Wert des Jahres 2002 ab, bei geringem Aufwand sogar unter 0,2. Die Bekanntheit steigt bei steigendem Aufwand stetig an, in den Anfangsjahren eher schwach, gegen Ende 2005 wird der Anstieg steiler. Die Bekanntheit nimmt ihren maximalen Wert im September/Oktober 2005 bei hohem Aufwand an, den tiefsten Punkt hingegen bei kleinem Aufwand im Mai 2007.

Bei der Zielgruppe der Männer sind ebenfalls die Wechselwirkungen des *Aufwandes* und der *Zeit* mit der *Länge* nicht signifikant, bei den restlichen Gruppen ist der Einfluss von *Aufwand* nur linear. Bei den Frauen ändert sich allerdings die Bekanntheit mit der Zeit anders für jede Längenkategorie.



(a) kürzer oder gleich 10 Sekunden (b) länger als 10 Sekunden

Abb. 3.9: Veränderung der Bekanntheit über die Zeit bei steigendem Aufwand und steigender Dauer des Infoscreen-Sujets

Die Bekanntheit ändert sich zwischen den beiden Längenkategorien in der Zielgruppe der Frauen nur geringfügig. Bei den Männern unterscheidet sie sich dahingehend, dass in der längeren Kategorie im ersten Jahr höhere Bekanntheitswerte auftreten.
In der Gruppe der 14- bis 29-Jährigen steigt die Bekanntheit mit wachsendem Aufwand nur linear. Außerdem zeigt sich in den ersten fünf Jahren in beiden Längenkategorien über die Zeit keine allzu starke Veränderung. Dies ist bis zum Jahr 2003, wenn auch nicht ganz so linear, bei der mittleren Altersklasse ebenfalls so. Für die jüngere Gruppe liegen außerdem die Werte der Bekanntheit um die Mitte des Jahres 2005 für teure Infoscreen-Spots höher. Die Veränderung über die Zeit ist bei der ältesten Gruppe am deutlichsten ausgeprägt. Dies ändert sich nicht mit dem Aufwand der Sujets. Im ersten Jahr, also bis 2003, gibt es einen starken Anstieg, der in den darauffolgenden eineinhalb Jahren bis zum Anfangswert wieder abklingt. Dieser Verlauf wiederholt sich mit dem Unterschied, dass in den letzten eineinhalb Jahren dann die Bekanntheit stärker sinkt.
Längere Spots führen auch bei diesem Werbeträger zu einem höheren Bekanntheitsgrad.

Markenimpact

Es gibt in keiner der Variablen fehlende Werte, es wurden natürlich wieder-

Tab. 3.16: Zahlenwerte zu nebenstehender Abbildung

	n	*Mar.*
Aufwand		
[0,573; 9,54)	88	32,13
[9,537;15,03)	85	37,77
[15,030;24,88)	87	38,67
[24,882;91,60]	86	40,16
Länge		
≤ 10	181	34,18
>10	165	40,42
Zeit		
[4,83; 5,92)	95	39,10
[5,92; 7,42)	84	36,56
[7,42; 8,92)	83	37,57
[8,92;10,33]	84	35,14
Overall	346	37,15

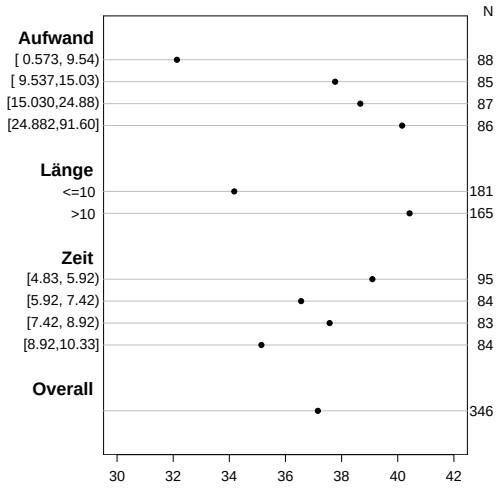


Abb. 3.10: Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“

um die Längenkategorien „≤ 10“ und „länger als 10“ gebildet.

Eine graphische Darstellung der Tabelle 3.16 bietet Abbildung 3.10. Man sieht, dass mit steigendem Aufwand und steigender Länge auch der mittlere Markenimpact größer wird. Bei der Variablen *Zeit* ist das Bild nicht monoton. Nach dem ersten Quartil sinkt das Mittel, im dritten steigt es und liegt zwischen erstem und zweitem Quartil. Das vierte Quartil weist den geringsten mittleren Wert auf.

Alle Variablen bis auf die Längenkategorie „länger als 10“ haben im logistischen Regressionsmodell signifikante Parameterschätzer. Diese Kategorie hat also keinen signifikant höheren Markenimpact als die kürzeren Sujets. Der Einfluss ist aber positiv. Alle drei Wechselwirkungen sind signifikant. Die P-Werte stimmen mit den P-Werten in Tabelle 3.17 überein. Tabelle 3.17 zeigt außerdem, dass alle Variablen, auch die Wechselwirkungen zwischen den Variablen, den Markenimpact beeinflussen.

Tab. 3.17: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „Infoscreen“

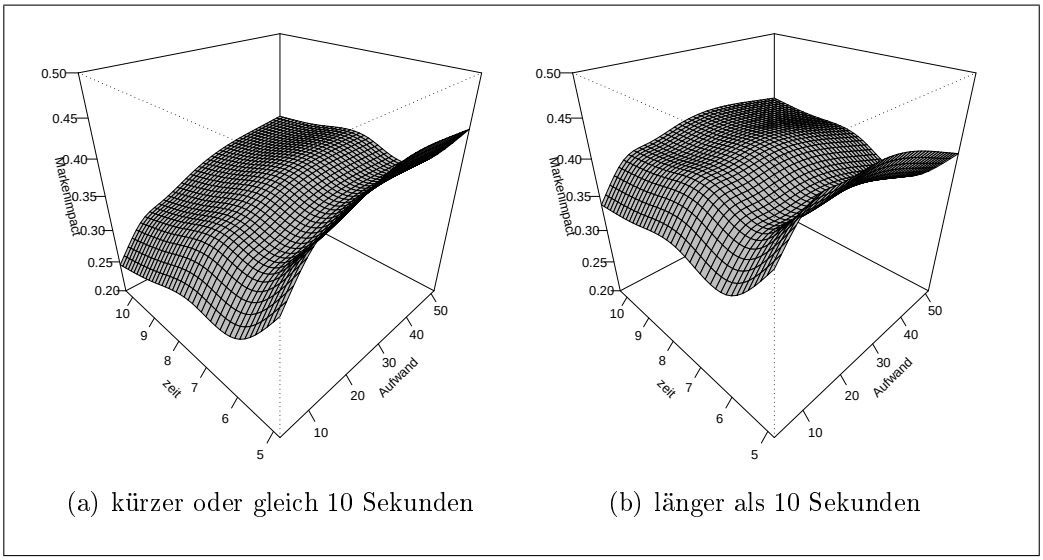
	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	6	132,5718	0,0000	134,7224	0,0000
Aufwand nichtlinear	3	43,2293	0,0000	43,7314	0,0000
Zeit	6	110,9105	0,0000	110,3788	0,0000
Zeit nichtlinear	3	36,3961	0,0000	36,3614	0,0000
Länge	3	153,4263	0,0000	152,8104	0,0000
Aufwand × *Zeit*	1	5,8529	0,0156	5,8493	0,0156
Aufwand × *Länge*	1	12,9738	0,0003	12,8352	0,0003
Zeit × *Länge*	1	14,5193	0,0001	14,5335	0,0001
TOTAL nichtlinear	6	74,9214	0,0000	75,4866	0,0000
TOTAL WW	3	43,6561	0,0000	43,6477	0,0000
TOTAL nichtlinear+WW	9	181,3652	0,0000	184,7675	0,0000
Regression	12	443,7866	0,0000	454,9818	0,0000

Das Ergebnis der logistischen Regression ist:

```
General Linear Model
glmD(formula = cbind(floor(M_TOTAL/100 * 150), 150 - floor(M_TOTAL/100 *
    150)) ~ rcs(Aufwand, 5) + rcs(zeit, 5) + Laenge.kat + Aufwand %ia%
    zeit + Aufwand %ia% Laenge.kat + zeit %ia% Laenge.kat,
    family = binomial, data = IS.reg)

Model L.R.       d.f.          P
       455         12          0
                           Coef      S.E.      Wald Z      P
Intercept                  1.056346  0.2756755   3.832   1.272e-04
Aufwand                    0.034297  0.0090467   3.791   1.500e-04
Aufwand'                  -0.321023  0.1245855  -2.577   9.974e-03
Aufwand''                  0.770417  0.3536060   2.179   2.935e-02
Aufwand'''                -0.518296  0.2908753  -1.782   7.477e-02
zeit                      -0.363811  0.0499909  -7.278   3.400e-13
zeit'                      1.969813  0.3985231   4.943   7.702e-07
zeit''                    -4.008563  0.9041154  -4.434   9.264e-06
zeit'''                    2.668643  0.8501581   3.139   1.695e-03
Laenge.kat> 10             0.009565  0.0976300   0.098   9.220e-01
Aufwand %ia% zeit          0.001112  0.0004596   2.419   1.555e-02
Aufwand %ia% Laenge.kat   -0.006829  0.0018959  -3.602   3.159e-04
zeit %ia% Laenge.kat       0.045089  0.0118332   3.810   1.387e-04
```

Die beiden Längenkategorien zeigen ein etwas unterschiedliches Bild, wie man in Abbildung 3.11 erkennen kann. Dies liegt an den signifikanten Wechselwirkungen zwischen *Aufwand* und *Zeit* mit der *Länge*. Die Werte liegen vor allem bei geringem Aufwand für die längeren Spots höher, bei hohem Aufwand ist der Unterschied nicht allzu groß. Die Funktion für die Kategorie „länger als 10“ hat also, von der Aufwand-Achse aus betrachtet, eine etwas flachere Form, die gegen Ende hin sogar leicht sinkt.



(a) kürzer oder gleich 10 Sekunden (b) länger als 10 Sekunden

Abb. 3.11: Veränderung des Markenimpacts über die Zeit bei steigendem Aufwand und steigender Dauer des Infoscreen-Sujets

In beiden Kategorien steigt der Markenimpact bis zu einem Aufwand von 10.000 € am stärksten an. Ab diesem Zeitpunkt flacht er in der Kategorie „≤ 10“ ab, in der zweiten sinkt er sogar. Dies ändert sich über die Zeit, sodass er in der kürzeren Kategorie sogar leicht steigt und in der längeren abflacht.
Im Laufe des Jahres 2002 sinkt der Markenimpact, daraufhin steigt er etwa eineinhalb Jahre und bleibt annähernd konstant. Den tiefsten Punkt haben beide Funktionen Anfang des Jahres 2003 bei einem minimalen Aufwand, den höchsten Punkt die Funktion der Kategorie „≤ 10“ bei einem Aufwand von etwa 30.000 €, die Kategorie „länger als 10“ bei einem Aufwand von 10.000 €, und zwar beide gegen Ende 2001.
Alle Variablen haben auch in der Gruppe der Männer einen signifikanten Einfluss. Bei den Frauen ändert sich der Einfluss über die Zeit nicht für unterschiedliche Werte des Aufwandes.
Der Markenimpact in der Gruppe der Frauen unterscheidet sich nur minimal von dem der Gesamtgruppe. Im ersten Jahr ist er über den Aufwandsverlauf geradliniger als für die Gesamtgruppe. Somit ändert sich dieser Verlauf über die Zeit hinweg nicht sehr stark, darauf verweist bereits die nicht signifikante Wechselwirkung im Modell. Für die Männer ist der Verlauf über die Zeit sehr ähnlich zu jener der Gesamtgruppe. Mit steigendem Aufwand hat in dieser Gruppe der Markenimpact einen nichtlinearen Verlauf. Es gibt zwischen einem Aufwand von 15.000 € und 20.000 € ein kurzes, flaches Stück.

In der Gruppe der 14- bis 29-Jährigen ist die Wechselwirkung zwischen *Zeit* und *Länge* nicht signifikant, bei den 30- bis 49-Jährigen hat alles einen signifikanten Einfluss. Sowohl die Wechselwirkung zwischen *Aufwand* und *Zeit* als auch *Aufwand* und *Länge* haben bei den älteren befragten Personen kaum einen Einfluss.
Ein Unterschied zu der mittleren Altersklasse und der Gesamtgruppe ist kaum auszumachen. Auch in der Alterklasse „14 bis 29“ ist die bereits in der männlichen Zielgruppe genannte Abflachung bemerkbar. Die Markenimpact-Werte in der Gruppe der über 50-Jährigen verhalten sich auch ähnlich wie in der Gesamtgruppe, vor allem in den ersten drei Jahren. Dann werden bei steigendem Aufwand niedrigere Markenimpact-Werte beobachtet, der Trend ist in der Zielgruppe sogar leicht negativ.

Gefälligkeit

Einen Einblick in das Verhalten der unabhängigen Variablen bezüglich der Gefälligkeit sieht man in Tabelle 3.18 und in Abbildung 3.12.
Anfangs steigt die mittlere Gefälligkeit, wenn der Aufwand steigt, nur im vierten Quartil ist der Mittelwert knapp unter dem Wert des dritten Quartils. Bei der Zeit ist die Änderung bis auf das dritte Quartil, das den größten mittleren Wert aufweist, negativ. Eine längere Dauer bedeutet auch eine höhere Bewertung.

Im linearen Regressionsmodell ist keiner der Parameterschätzer signifikant. Tabelle 3.19 zeigt, dass weder der Aufwand noch die Wechselwirkungen einen Einfluss darauf haben, wie einem Befragten ein Infoscreen-Spot gefällt.
In Abbildung 3.13 kennzeichnen die zwei Kurven die unterschiedlichen Längenkategorien. Je länger der Infoscreen-Spot dauert, desto besser fällt auch die Beurteilung bezüglich der Gefälligkeit aus. Die beiden Kurven sind nicht exakt parallel, weil die Wechselwirkungen bei der Berechnung im Modell bleiben. Außerdem wurde auch die Variable *Aufwand* in linearer Form im Modell behalten. Sie wurde dann für die Abbildung auf ihren Median gesetzt, damit man nur die Veränderung über die Zeit für die beiden Längenkategorien betrachten kann.
In der Anfangszeit der Befragung, das heißt etwa bis zum Jahr 2003, haben die Kurven eine negative Steigung, dann nimmt die Gefälligkeit für knapp zwei Jahre zu, woraufhin sie bis Mitte 2007 fällt. Je fortgeschrittener die Zeit ist, desto größer wird der Unterschied zwischen den beiden Kategorien. Liegt er Ende 2001 noch bei etwa 0,1, so ist die Differenz der beiden Kurven im Juni 2007 bereits 0,4.

Tab. 3.18: Zahlenwerte zu nebenstehender Abbildung

	n	*Gef.*
Aufwand		
[0,573; 9,54)	88	5,38
[9,537;15,03)	85	5,40
[15,030;24,88)	87	5,57
[24,882;91,60]	86	5,54
Länge		
≤ 10	181	5,33
>10	165	5,63
Zeit		
[4,83; 5,92)	95	5,50
[5,92; 7,42)	84	5,49
[7,42; 8,92)	83	5,56
[8,92;10,33]	84	5,33
Overall	346	5,47

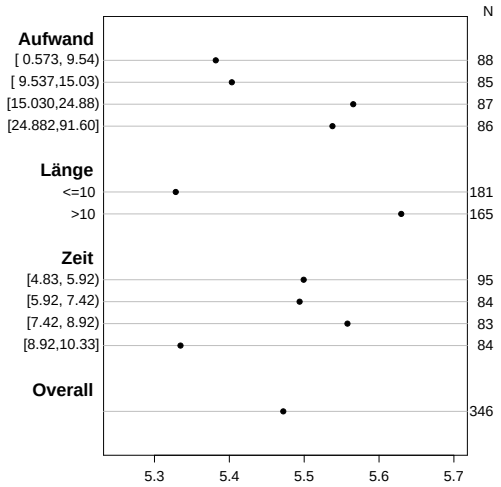


Abb. 3.12: Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“

Tab. 3.19: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Gefälligkeit* im Werbeträger „Infoscreen“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	6	8,6646	0,1933	8,8877	0,1800
Aufwand nichtlinear	3	3,6099	0,3068	3,7306	0,2921
Zeit	6	16,5670	0,0110	16,7993	0,0100
Zeit nichtlinear	3	13,0982	0,0044	13,3487	0,0039
Länge	3	17,5045	0,0006	17,7259	0,0005
Aufwand $\times$ *Zeit*	1	0,4704	0,4928	0,4884	0,4846
Aufwand $\times$ *Länge*	10	0,9358	0,3334	0,9709	0,3244
Zeit $\times$ *Länge*	1	2,6316	0,1048	2,7236	0,0989
TOTAL nichtlinear	6	16,7012	0,0104	16,9320	0,0095
TOTAL WW	3	3,2978	0,3479	3,4097	0,3327
TOTAL nichtlinear+WW	9	19,8209	0,0191	20,0050	0,0179
Regression	12	47,4981	0,0000	46,1352	0,0000

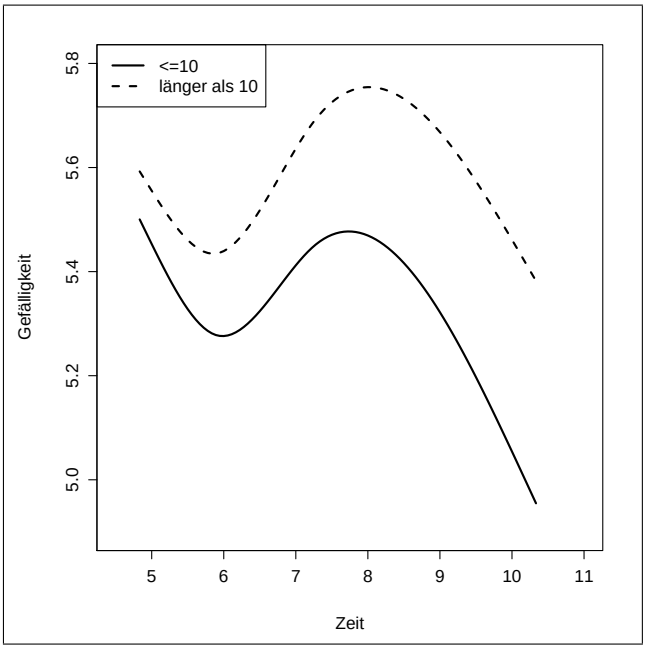


Abb. 3.13: Veränderung der *Gefälligkeit* bezüglich der *Zeit*, getrennt nach Längenkategorien

In den einzelnen Zielgruppen hat der Aufwand ebenfalls nie einen Einfluss. Lediglich die Wechselwirkungen zwischen der *Länge* und der *Zeit* haben in den Zielgruppen „weiblich“ und „30 bis 49“ einen Einfluss. In diesen beiden Gruppen liegen die Kurven im Jänner 1998 sehr nahe beieinander, driften dann aber voneinander weg. Mitte 2007 liegen die Gefälligkeitswerte zirka 0,5 Punkte auseinander. Die Bewertungen in der Längenkategorie „$\leq$ 10“ stimmen aber im Großen und Ganzen mit jenen der Gesamtgruppe überein. Bei der männlichen Zielgruppe ist der Verlauf ähnlich wie in der Gesamtgruppe. Hier sind alle Wechselwirkungen nicht signifikant, die Funktionen der beiden Längenkategorien sind hier also annähernd parallel.
In der Altersgruppe der 14- bis 29-Jährigen unterscheiden sich die beiden Längenkategorien auch nur durch eine Parallelverschiebung. Allerdings liegen die Bewertungen für die Spots über 10 Sekunden deutlich über der Bewertung in der Gesamtgruppe.
Die Gefälligkeitswerte der über 50-Jährigen in den beiden Kategorien sind auch annähernd parallel, sie liegen jedoch sehr nahe zusammen, unterscheiden sich in den ersten Jahren um etwa 0,1 und gegen Mitte 2007 dann um etwa 0,2 Gefälligkeitspunkte.

Tab. 3.20: Zahlenwerte zu nebenstehender Abbildung

	n	*Bek.*
Aufwand		
[0,74; 86,2)	474	33,40
[86,25; 149,4)	473	33,37
[149,41; 231,5)	474	34,59
[231,54;1178,5]	473	38,33
Zeit		
[1,08; 3,58)	475	31,05
[3,58; 5,83)	473	32,45
[5,83; 8,17)	480	37,41
[8,17;10,50]	466	38,81
Overall	1.894	34,92

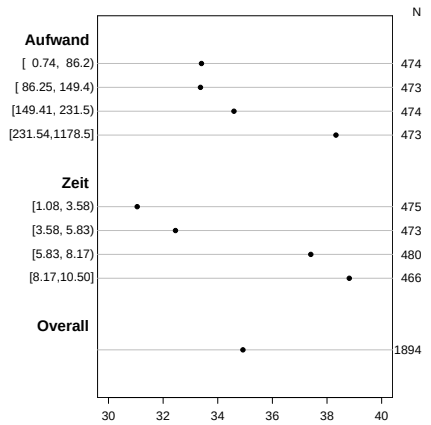


Abb. 3.14: Mittelwerte der Variablen *Bekanntheit*, getrennt nach Quartilen im Werbeträger „Print“

3.3.3 Einfluss im Werbeträger „Print“

Im Werbeträger „Print“ gibt es die Variable *Länge* nicht. Deshalb werden die Modelle nur mit den Variablen *Aufwand* und *Zeit* geschätzt. Die Variable *Aufwand* hat 21 fehlende Werte, wie man Tabelle 3.4 in Kapitel 3.2 entnehmen kann. Die Größe reduziert sich demnach von 1.915 auf $1.915-21 = 1.894$.

Bekanntheit

Wie sich die Bekanntheit in den Quartilen der unabhängigen Variablen verändert, sieht man in Tabelle 3.20 und Abbildung 3.14.
Sowohl in den Quartilen der Variablen *Aufwand* als auch der Variablen *Zeit* steigt die Bekanntheit, wobei die Variable *Zeit* einen größeren Einfluss zu haben scheint.

Im logistischen Regressionsmodell hat nur die Variable *Aufwand* keinen signifikanten Koeffizienten. In Tabelle 3.21 sieht man aber, dass alle Variablen, also auch der Aufwand, einen Einfluss auf die Bekanntheit einer Print-Werbung haben.

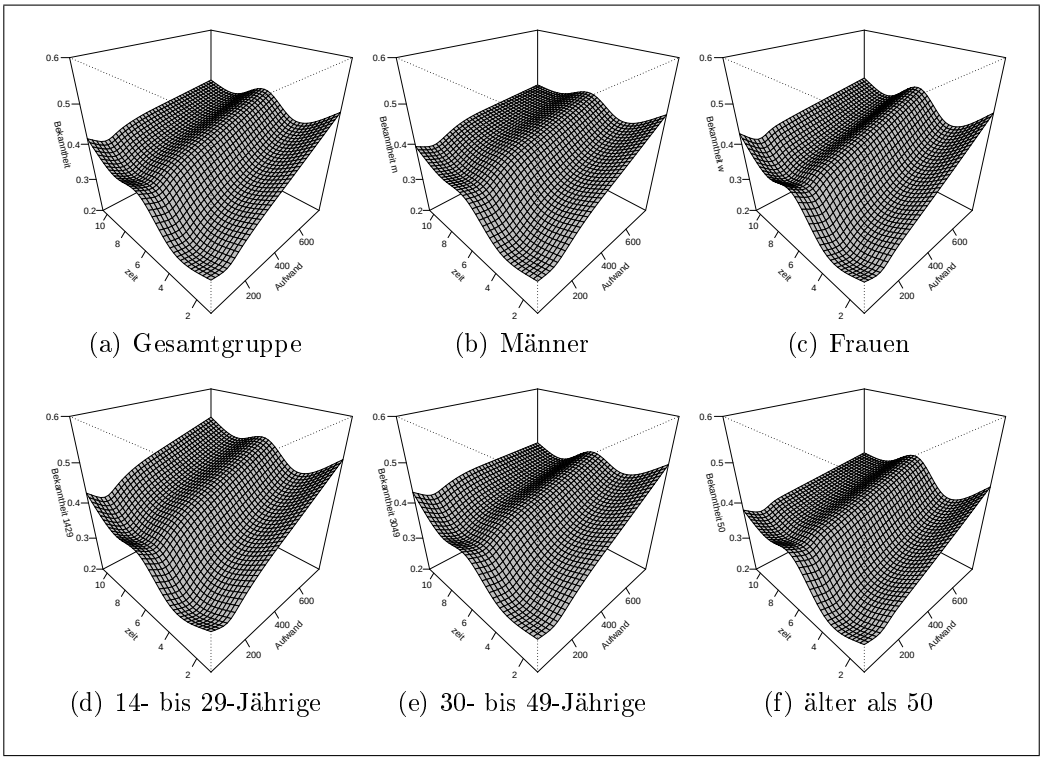
Abbildung 3.15(a) zeigt, dass der Aufwand anfangs einen positiven Einfluss

Tab. 3.21: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Bekanntheit* im Werbeträger „Print“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	5	860,4054	0,0000	850,6235	0,0000
Aufwand nichtlinear	3	45,2332	0,0000	45,1426	0,0000
Zeit	5	1571,0099	0,0000	1589,2358	0,0000
Zeit nichtlinear	3	298,0025	0,0000	298,0749	0,0000
Aufwand × *Zeit*	1	56,4403	0,0000	56,4368	0,0000
TOTAL nichtlinear	6	348,4235	0,0000	348,3956	0,0000
TOTAL nichtlinear+WW	7	410,2007	0,0000	410,4639	0,0000
Regression	9	2288,3882	0,0000	2320,6650	0,0000

auf die Bekanntheit eines Print-Sujets hat. Dieser nimmt etwa ab dem Jahr 2003 ab und ist in den letzten beiden Jahren jeweils bis zu einem Aufwand von 50.000 € sogar leicht negativ. Über die Zeit hinweg steigt die Bekanntheit bei geringem Aufwand, bis Ende 1999 ändert sie sich kaum, dann steigt sie bis Anfang 2003 ziemlich stark, gefolgt von einem flachen Stück und anschließender Steigung. Dieser Verlauf ändert sich bei steigendem Aufwand. Bis Ende 2000 sinkt die Bekanntheit von teuren Print-Sujets, es folgt eine Steigung bis Ende 2002, dann ist der Einfluss wiederum negativ und steigt ab Anfang 2005 wieder auf den Ausgangswert von 1998 von 0,48 bzw. 48 %.

Wald- und Likelihood-Ratio-Tests liefern auch in allen Zielgruppen signifikante Ergebnisse. Die Änderung der Bekanntheit unterscheidet sich bei steigendem Aufwand in den beiden Zielgruppen männlich und weiblich kaum (vgl. 3.15(b) und 3.15(c)). Es ändert sich allerdings der Verlauf über die Zeit. Bei der Gruppe der Männer ist der Verlauf glatter, bei den Frauen stärker ausgeprägt. Das heißt, dass bei einem Abfall der Bekanntheit dieses Absinken bei den Frauen steiler ist.
Zwischen der mittleren Altersgruppe und der Gesamtheit der Befragten gibt es im Verlauf keine Unterschiede. Die Zunahme der Bekanntheit mit steigendem Aufwand ist in der Klasse der 14- bis 29-Jährigen etwas schneller. Über die Zeit verändert sich der Einfluss nicht. In der Gruppe der Personen, die über 50 Jahre alt sind, ist die Veränderung über die Zeit stärker ausgeprägt, außerdem bleibt die Bekanntheit bei steigendem Aufwand konstant. Diese Veränderungen sieht man in den Abbildungen 3.15(d) bis 3.15(f)



(a) Gesamtgruppe (b) Männer (c) Frauen

(d) 14- bis 29-Jährige (e) 30- bis 49-Jährige (f) älter als 50

Abb. 3.15: Veränderung der Bekanntheit über die Zeit bei steigendem Aufwand im Werbeträger „Print“, getrennt für die einzelnen Zielgruppen

Markenimpact

Die Variable *Markenimpact* hat im Datensatz „Print“ zwei fehlende Werte, die Variable *Aufwand* 21 (vgl. Tabelle 3.4). Bei einer Beobachtung fehlen beide Variablen, somit reduziert sich die Dimension für dieses Modell auf $1.915 - 21 - 2 + 1 = 1.893$.
Tabelle 3.22 und Abbildung 3.16 zeigen, dass der Markenimpact im Mittel steigt, wenn mehr für das Sujet aufgewendet wird. Über die Zeit hinweg geschieht dies nicht linear, sondern steigt zwischen erstem und zweitem Quartil, im dritten Quartil ist er niedriger und im vierten Quartil steigt er, bleibt aber im Wert unter den ersten beiden Quartilen.

Es haben alle Variablen einen signifikanten Einfluss auf den Markenimpact eines Print-Sujets, wie die Wald- und Likelihood-Ratio-Teststatistiken in Tabelle 3.23 zeigen.

Tab. 3.22: Zahlenwerte zu nebenstehender Abbildung

	n	*Mar.*
Aufwand		
[0,74; 86,2)	474	30,48
[86,25; 149,6)	473	33,22
[149,62; 231,5)	473	33,53
[231,54;1178,5]	473	36,51
Zeit		
[1,08; 3,58)	475	34,53
[3,58; 5,83)	472	34,71
[5,83; 8,17)	480	31,11
[8,17;10,50]	466	33,43
Overall	1.893	33,44

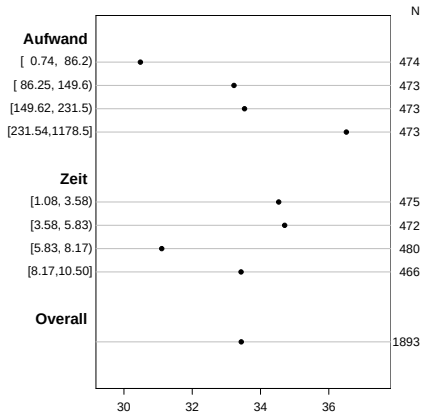


Abb. 3.16: Mittelwerte der Variablen *Markenimpact*, getrennt nach Quartilen im Werbeträger „Print“

Tab. 3.23: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „Print“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	5	642,6615	0,0000	639,9830	0,0000
Aufwand nichtlinear	3	39,4222	0,0000	39,4870	0,0000
Zeit	5	206,6170	0,0000	207,6029	0,0000
Zeit nichtlinear	3	129,3458	0,0000	129,2798	0,0000
Aufwand × *Zeit*	1	34,3600	0,0000	34,3634	0,0000
TOTAL nichtlinear	6	172,8522	0,0000	172,9179	0,0000
TOTAL nichtlinear+WW	7	213,7345	0,0000	214,3055	0,0000
Regression	9	840,5626	0,0000	843,7541	0,0000

Abbildung 3.17(a) zeigt, dass mit steigendem Aufwand auch die Werte der Variablen *Markenimpact* wachsen, in den Jahren 1998 bis Anfang 2005 ändert sich im Verlauf wenig, ab Anfang 2005 ist der Anstieg größer.
Print-Werbungen mit einem geringen Aufwand haben bis Mitte 2002 einen konstanten Markenimpact, der dann bis Ende 2004 leicht sinkt und anschließend auf den Wert von 1998 steigt. Wird für die Sujets mehr Geld aufgewendet, so ändert sich diese Entwicklung ein wenig. Der Markenimpact hat fast durchgehend einen positiven Verlauf. Bis Mitte 2002 steigt er nur wenig, dann stagniert er bis Ende 2004 und es folgt eine Steigung bis zu einem Wert knapp unter 0,5.
Den höchsten Markenimpact weisen teure Sujets aus dem Jahr 2007 auf, den geringsten günstigere gegen Ende 2004 bzw. Anfang 2005.

Sowohl in der Zielgruppe der Frauen als auch in jener der Männer sind alle Variablen signifikant. Der Verlauf für die Gruppe der Frauen unterscheidet sich kaum vom Verlauf, wenn man alle Befragten betrachtet (vgl. Abbildung 3.17(a) und 3.17(c)). In Abbildung 3.17(b) erkennt man, dass sich bei den Männern der Markenimpact über die Zeit kaum verändert. Bis zu einem Aufwand von etwa 220.000 € steigt er an, dann verläuft die Steigung schwächer. Dieser Knick ist in allen Jahren präsent.
In der jüngsten Altersklasse ist der Einfluss des Aufwands nur linear, außerdem ist die Wechselwirkung zwischen Aufwand und Zeit nicht signifikant. Diese beiden Eigenschaften verändern die Funktion deutlich, dies veranschaulicht Abbildung 3.17(d). Zur Gruppe der 30- bis 49-Jährigen gibt es wieder kaum Unterschiede (siehe Abbildung 3.17(e)). In der Klasse der über 50-jährigen Befragten gibt es eine Welle über den gesamten Zeitverlauf zwischen einem Aufwand von 160.000 € bis 200.000 €. Dieses Phänomen kann man auch in der Zielgruppe der Frauen erkennen.

Gefälligkeit

Für die Variable *Gefälligkeit* gibt es wieder 1.894 vollständige Beobachtungen, da die abhängige Variable *Gefälligkeit* keine fehlenden Werte hat.
In Tabelle 3.24 und Abbildung 3.18 erkennt man, dass bei steigendem Aufwand die mittlere Beurteilung der Sujets nicht steigt, sie bleibt ziemlich gleich und variiert nur zwischen 5,068 und 5,107. Ein anderes Bild zeigt sich über die Zeit. Je jünger ein Sujet ist, desto besser wird es bewertet, der Mittelwert reicht von 4,616 bis zu 5,428.

Im linearen Regressionsmodell zeigen die Koeffizienten der Variablen *Auf-*

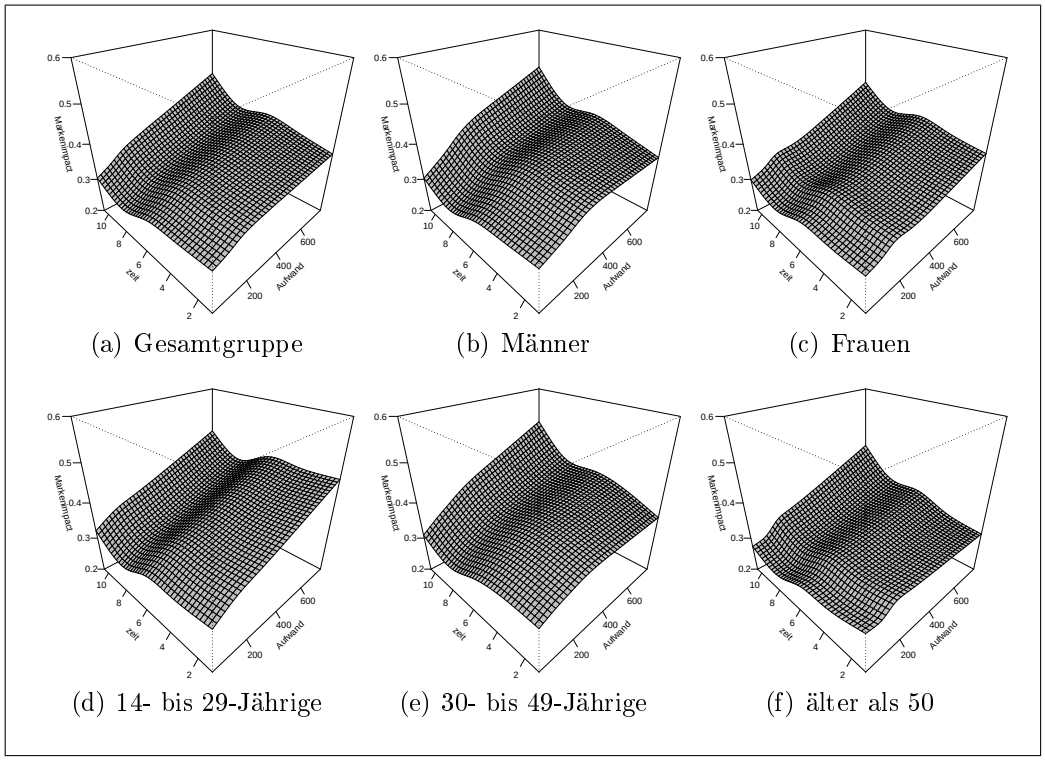


Abb. 3.17: Veränderung des Markenimpacts über die Zeit bei steigendem Aufwand im Werbeträger „Print“, getrennt für die einzelnen Zielgruppen

wand keinen signifikanten Beitrag, ebenso die Wechselwirkung mit der *Zeit*. Die *Zeit* hat im Gegensatz dazu ausschließlich signifikante Parameter.

Tabelle 3.25 zeigt, dass die Bewertung eines Sujets vom Aufwand nicht abhängt. Über die Zeit hinweg ändert sich die Gefälligkeit aber sehr wohl.

Für die Parameterschätzung in Abbildung 3.19 ist sowohl der Aufwand als lineare Einflussgröße als auch die Wechselwirkung zwischen Aufwand und Zeit im Modell geblieben. Der Aufwand wurde bei seinem Median festgehalten und nur die Variable Zeit variiert. Betrachtet man das Modell nur mit der Variablen *Zeit* als restringierten Spline, so ändert sich das Bild nur minimal. Die Bewertung der Sujets für die Gesamtgruppe (dicke durchgehende Linie in Abbildung 3.19) steigt mit der Variablen *Zeit* bis Mitte 2003 am stärksten, flacht dann ab und bleibt auf dem Wert 5,4.

Auch ist in jeder Zielgruppe für die Bewertung des Sujets der Einfluss des

Tab. 3.24: Zahlenwerte zu nebenstehender Abbildung

	n	*Gef.*
Aufwand		
[0,74; 86,2)	474	5,11
[86,25; 149,4)	473	5,08
[149,41; 231,5)	474	5,07
[231,54;1178,5]	473	5,08
Zeit		
[1,08; 3,58)	475	4,62
[3,58; 5,83)	473	5,00
[5,83; 8,17)	480	5,29
[8,17;10,50]	466	5,43
Overall	1.894	5,08

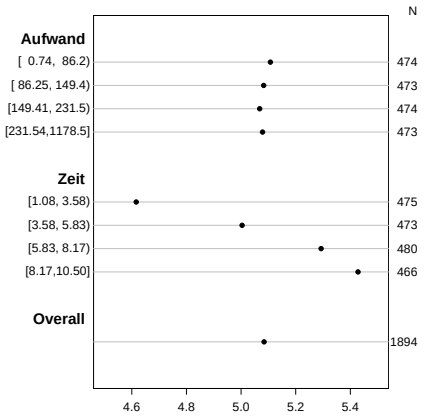


Abb. 3.18: Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Quartile im Werbeträger „Print“

Tab. 3.25: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Gefälligkeit* im Werbeträger „Print“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	5	8,2528	0,1428	8,2785	0,1415
Aufwand nichtlinear	3	6,2102	0,1018	6,2329	0,1008
Zeit	5	825,3265	0,0000	688,0957	0,0000
Zeit nichtlinear	3	83,8281	0,0000	82,4520	0,0000
Aufwand × *Zeit*	1	1,9947	0,1578	2,0043	0,1569
TOTAL nichtlinear	6	87,3135	0,0000	85,8036	0,0000
TOTAL nichtlinear+WW	7	88,5032	0,0000	86,9464	0,0000
Regression	9	831,6310	0,0000	692,4979	0,0000

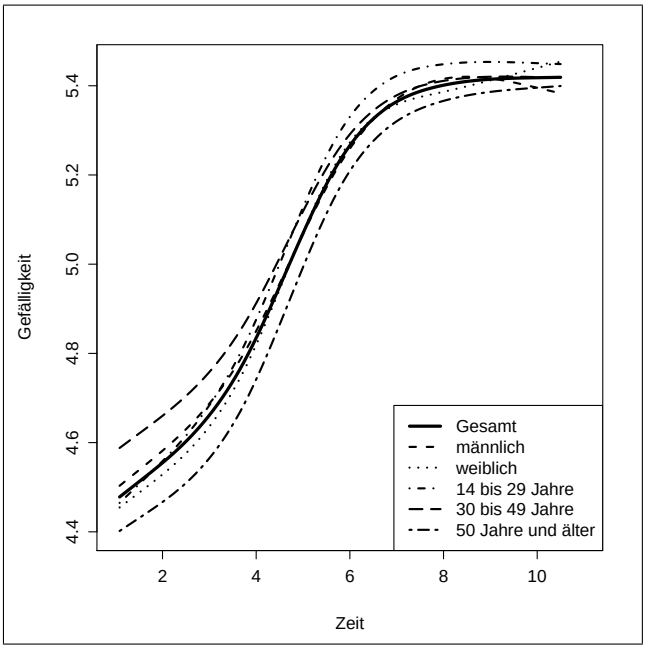


Abb. 3.19: Veränderung der Gefälligkeit von Print-Sujets über die Zeit in den verschiedenen Zielgruppen

Aufwands nicht ausschlaggebend. Ebenfalls sind die Wechselwirkungen zwischen *Aufwand* und *Zeit* in keiner Zielgruppe signifikant.
Zwischen den Gesamtbefragten und der Zielgruppe der Männer und Frauen gibt es kaum einen Unterschied. Im letzten Jahr allerdings ist die Tendenz bei den Männer eher fallend und jene bei den Frauen steigend.
Die Gruppe der 14- bis 29-Jährigen unterscheidet sich in den ersten drei Jahren kaum von der gesamten Stichprobe. Ab dem Jahr 2001 steigt sie steiler und flacht dann auf einen Wert von knapp über 5,4 ab. Die mittlere Altersgruppe liegt bis zum Jahr 2005 höher als die Vergleichsgruppe und fällt dann mit deren Wert zusammen. Die Gruppe der Personen, die 50 Jahre und älter sind, liegen fast parallel etwa 0,1 Gefälligkeitspunkte unter der gesamten Stichprobe. All diese Veränderungen zeigt Abbildung 3.19.

3.3.4 Einfluss im Werbeträger „TV“

Der Datensatz für den Werbeträger „TV“ hat 6.279 Beobachtungen, 36 fehlende Werte weist die Variable *Aufwand* und 11 die Variable *Zeit* auf. Zwei Werte fehlen bei beiden Variablen. Deshalb vermindert sich in diesem Datensatz die Anzahl der Beobachtungen auf $6.279 - 36 - 11 + 2 = 6.234$. Die Variable *Markenimpact* hat in diesem Datensatz keine fehlenden Werte,

Tab. 3.26: Zahlenwerte zu nebenstehender Abbildung

	n	*Bek.*
Aufwand		
[0,682; 80,9)	1.559	40,14
[80,856; 133,2)	1.558	43,33
[133,166; 208,1)	1.561	45,27
[208,120;1146,5]	1.556	49,99
Länge		
≤ 15	868	40,65
15-20	1.703	43,04
20-25	1.038	45,69
25-30	1.942	46,29
>30	440	48,62
„Tandem“	243	46,26
Zeit		
[1,08; 3,83)	1.583	46,29
[3,83; 6,17)	1.538	42,58
[6,17; 8,50)	1.614	42,65
[8,50;10,50]	1.499	47,32
Overall	6.234	44,68

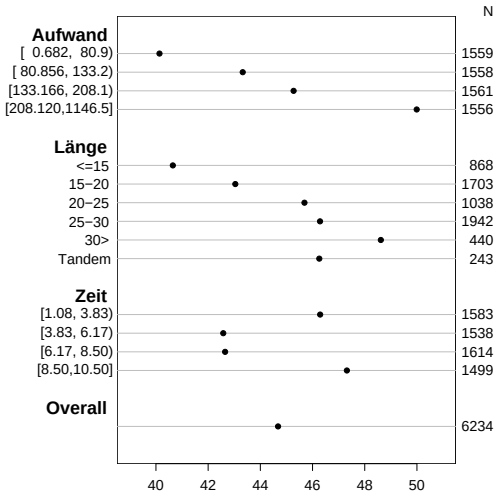


Abb. 3.20: Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“

deshalb gibt es in allen drei Modellen 6.234 Beobachtungen.

Bekanntheit

Einen Eindruck, wie sich die Variable *Bekanntheit* im Werbeträger „TV“ in den einzelnen Quartilen und Kategorien verhält, geben Tabelle 3.26 und Abbildung 3.20. Je höher der Aufwand, desto höher ist auch der Bekanntheitsgrad. Je länger eine Werbung dauert, desto eher kennen die Interviewten diese. Der Mittelwert von „Tandem“-Spots ist in etwa gleich dem Mittel der Kategorie „25 bis 30“, liegt also unter der Kategorie „länger als 30“. Die Mittelwerte im zweiten und dritten Quartil der *Zeit* sind annähernd gleich, jedoch kleiner als der Wert des ersten Quartils. Der Wert im vierten Quartil ist mit 47,32 der größte.
Tabelle 3.27 zeigt, dass alle Variablen einen Einfluss auf die Bekanntheit haben, auch die nichtlinearen Terme und die Wechselwirkungen tragen zum Bekanntheitsgrad bei.
Der Einfluss von Aufwand und Zeit ist für die einzelnen Längenkategorien sehr ähnlich. Die Funktionen unterscheiden sich nur durch unterschiedliche

Tab. 3.27: Wald- und Likelihood-Ratio-Teststatistiken für die Variable *Bekanntheit* im Werbeträger „TV“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	10	4426,5355	0,0000	4466,2868	0,0000
Aufwand nichtlinear	3	75,5437	0,0000	75,4187	0,0000
Zeit	10	1558,7427	0,0000	1563,2386	0,0000
Zeit nichtlinear	3	1362,5936	0,0000	1366,3868	0,0000
Länge	15	473,7068	0,0000	473,4641	0,0000
Aufwand × *Zeit*	1	133,0592	0,0000	133,3891	0,0000
Aufwand × *Länge*	5	140,6511	0,0000	140,7668	0,0000
Zeit × *Länge*	5	119,3817	0,0000	119,4229	0,0000
TOTAL nichtlinear	6	1450,7327	0,0000	1454,6463	0,0000
TOTAL WW	11	344,4654	0,0000	344,7876	0,0000
TOTAL nichtlinear+WW	17	1888,9071	0,0000	1892,7291	0,0000
Regression	24	7642,5556	0,0000	7740,3773	0,0000

Steigungen. Der Aufwand hat einen monoton steigenden Einfluss, der in den ersten Jahren ein bisschen stärker und gegen Ende hin nur noch bis zu einer Bekanntheit von knapp 70 % ansteigt.
Bei geringem Aufwand ändert sich die Bekanntheit bis Ende 2000 kaum, hierauf sinkt sie, steigt ab Anfang 2003 bis 2006 und bleibt dann wieder konstant. Der Verlauf ist sehr symmetrisch. Bei hohem Aufwand ist die Tendenz durchgehend negativ, sie wird nur in den Jahren 2003 und 2004 von einem „geraden Stück“ unterbrochen. In den ersten Jahren sinkt sie steiler, nach dem flachen Stück nur noch sehr wenig. Dieser Trend über die Zeit findet sich in allen Längenkategorien. Mit steigender Länge wird aber, wie oben bereits angesprochen, die Steigung bei wachsendem Aufwand immer geringer, die Funktionen neigen sich also bei hohem Aufwand nach unten.
Die Tandem-Werbungen sind den Werbungen in der Referenzkategorie sehr ähnlich, sie haben bis Ende 2002 allerdings leicht geringere Werte. Das Minimum haben alle Kategorien Ende 2002 bei einem geringen Aufwand, ihr Maximum im Jänner 1998 bei hohem Aufwand.

Auch in den Modellen der Zielgruppen „männlich“ und „weiblich“ und der Alterklassen „14 bis 29“ und „30 bis 49“ haben alle Variablen einen signifikanten Einfluss. Bei der ältesten Altersklasse ist der Einfluss des Aufwands nur noch linear. Dies äußert sich dadurch, dass in den restlichen Altersklassen der Einfluss bei steigendem Aufwand leicht konkav ist, also am Anfang steigt und dann abflacht. Dies ist bei den über 50-Jährigen nicht der Fall, da

Tab. 3.28: Zahlenwerte zu nebenstehender Abbildung

	n	*Mar.*
Aufwand		
[0,682; 80,9)	1.559	30,27
[80,856; 133,2)	1.558	31,84
[133,166; 208,1)	1.561	33,22
[208,120;1146,5]	1.556	37,84
Länge		
≤ 15	868	29,15
15-20	1.703	30,36
20-25	1.038	34,39
25-30	1.942	35,09
>30	440	40,90
„Tandem“	243	35,67
Zeit		
[1,08; 3,83)	1.583	34,61
[3,83; 6,17)	1.538	35,38
[6,17; 8,50)	1.614	29,99
[8,50;10,50]	1.499	33,29
Overall	6.234	33,29

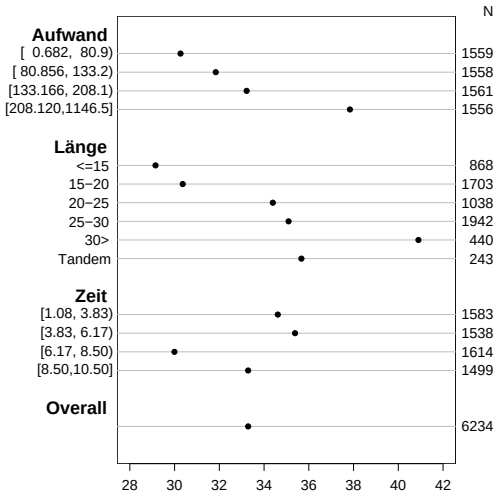


Abb. 3.21: Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“

der Aufwand hier seine Krümmung nicht ändert.
Zwischen den beiden Geschlechtsgruppen gibt es keine nennenswerten Unterschiede.
In der Altersklasse der 14- bis 29-Jährigen liegen im Bereich des Jahres 1998 bei hohem Aufwand die Bekanntheitswerte leicht über den Werten der Gesamtbefragten, in der Klasse der über 50-Jährigen in diesem Bereich leicht darunter. Zur mittleren Alterklasse gibt es keine Differenzen.

Markenimpact

Bei einem Blick auf Tabelle 3.28 und Abbildung 3.21 sieht man, dass der Einfluss von *Aufwand* und *Länge* auf die Wiedererkennung eines TV-Sujets monoton und positiv ist. Die Variable *Zeit* zeigt kein monotones Bild, anfangs steigt sie von 34,61 auf 35,38, dann sinkt sie im dritten Quartil auf 29,99, woraufhin sie im vierten Quartil auf 33,29 steigt.
Im Modell sind alle Haupteffekte signifikant, nicht jedoch die Wechselwirkung zwischen dem *Aufwand* und der *Zeit* bzw. der Längenkategorie „15 bis

Tab. 3.29: Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „TV“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	10	2101,2418	0,0000	2074,7728	0,0000
Aufwand nichtlinear	3	147,4049	0,0000	146,8185	0,0000
Zeit	10	2149,6275	0,0000	2162,3959	0,0000
Zeit nichtlinear	3	1109,2494	0,0000	1108,2326	0,0000
Länge	15	2631,4189	0,0000	2605,7609	0,0000
Aufwand × *Zeit*	1	1,7959	0,1802	1,7962	0,1802
Aufwand × *Länge*	5	326,9974	0,0000	324,9946	0,0000
Zeit × *Länge*	5	180,0319	0,0000	180,0273	0,0000
TOTAL nichtlinear	6	1265,6638	0,0000	1264,5436	0,0000
TOTAL WW	11	505,0721	0,0000	504,0600	0,0000
TOTAL nichtlinear+WW	17	1817,1432	0,0000	1818,1857	0,0000
Regression	24	8161,0632	0,0000	8262,8999	0,0000

20“ und die Wechselwirkung zwischen der *Zeit* und der Tandem-Kategorie.

In Tabelle 3.29 sieht man, dass alle Variablen bis auf die Wechselwirkung der Variablen *Aufwand* und *Zeit* einen signifikanten Beitrag zur Modellierung der Variablen *Markenimpact* haben.
Die Markenimpact-Werte von TV-Spots erhöhen sich bei steigendem Aufwand, bis zu einem Aufwand von 50.000 € allerdings kaum. Dieser Einfluss ändert sich über die Zeit hinweg unwesentlich, was schon am nicht signifikanten P-Wert der Wechselwirkung zwischen *Aufwand* und *Zeit* in Tabelle 3.29 abzulesen ist. Bis Ende 2002 ist der Einfluss auf den Markenimpact konstant, daraufhin sinkt er stetig bis Ende 2004. In den letzten zweieinhalb Jahren ist er wieder gestiegen. Dieses Bild zeigt sich, egal wie hoch der Aufwand ist (vgl. P-Wert von 0,1802 in Tabelle 3.29). Die Längenkategorie „15 bis 20“ unterscheidet sich kaum von der Referenzkategorie. In den darauffolgenden beiden Kategorien ist der Einfluss bei steigendem Aufwand weniger stark, in der Kategorie „länger als 30“ ändert sich der Markenimpact bei wachsendem Aufwand kaum. Für Tandem-Schaltungen zeigt die Abbildung einen ähnlichen Verlauf wie in den beiden Kategorien „20 bis 25“ und „25 bis 30“.

Bei der Überprüfung, welche Variablen in den Zielgruppen einen Einfluss haben, sind einige Wechselwirkungen nicht signifikant. Dies betrifft die Wechselwirkung des *Aufwandes* mit der *Länge* bei den männlichen Befragten und bei den Altersklassen „14 bis 29“ und „30 bis 49“. Der Markenimpact verän-

dert sich mit dem Aufwand also in allen Längenkategorien gleich.
Bei den Frauen und den über 50-Jährigen gibt es keine signifikante Wechselwirkung zwischen *Aufwand* und *Zeit*. Der Verlauf des Markenimpacts für eine gewisse Zeit ändert sich in diesen Gruppen also nicht mit dem Aufwand.
Bei den Männern ist über alle Jahre hinweg der Anstieg des Markenimpacts bei steigendem Aufwand stets höher als bei den Frauen. In der Kategorie „länger als 30“ zeigt sich fast nur noch eine Parallelverschiebung, bei den Männern nach oben und bei den Frauen nach unten. Über die Zeit ändert sich der Markenimpact kaum.
Auch bei den Alterklassen unterscheidet sich der Einfluss über die Zeit kaum. In den Längenkategorien „$\leq$ 15“ und „15 bis 20“ ist der Anstieg für wachsenden Aufwand in der Klasse der 14- bis 29-Jährigen größer, für die 30- bis 49-Jährigen steigt er nochmals, dafür ist er bei den über 50-Jährigen geringer. In den Kategorien „20 bis 25“ und „25 bis 30“ unterscheiden sich die Altersgruppen „14 bis 29“ und „30 bis 49“ kaum. Die Gruppe „50 Jahre und älter“ liegt ein bisschen darunter. In den beiden restlichen Längenkategorien gibt es kaum nennenswerte Differenzen.

Gefälligkeit

Die Mittelwerte der Variablen *Gefälligkeit* steigen mit dem Aufwand, der Zeit und mit der Länge, wobei die Tandem-Werbungen im Mittel bewertet werden wie Werbungen, die zwischen 25 und 30 Sekunden dauern, wie man in Tabelle 3.30 und Abbildung 3.22 sieht.

Die Ergebnisse der linearen Regression zeigen, dass in diesem Modell der Koeffizient der Variablen *Zeit* sowie die dritte konstruierte Spline-Variable des *Aufwands* als auch der *Zeit*, außerdem die Wechselwirkungen der *Zeit* mit den Längenkategorien „15 bis 20“ und „Tandem“ nicht signifikant sind.
Tabelle 3.31 zeigt, dass alle Variablen signifikant zur Erklärung der *Gefälligkeit* beitragen.
Bei den kürzesten TV-Werbungen steigen bis zu einem Aufwand von etwa 50.000 € die Bewertungen, dann flachen sie bis zum Wert 200.000 € ab und fallen dann kontinuierlich. Dieser Einfluss ändert sich über die Zeit hinweg nicht besonders.
In den ersten Jahren ändert sich die Gefälligkeit kaum. Ab 2002 steigen die Gefälligkeitswerte und ab April 2005 fallen sie leicht, bei hohem Aufwand sogar etwas stärker.
Die Längenkategorien zeigen hier ein recht unterschiedliches Bild. Während in der Referenzkategorie die Bewertungen mit steigendem Aufwand fallen,

Tab. 3.30: Zahlenwerte zu nebenstehender Abbildung

	n	*Gef.*
Aufwand		
[0,682; 80,9)	1.559	5,13
[80,856; 133,2)	1.558	5,30
[133,166; 208,1)	1.561	5,43
[208,120;1146,5]	1.556	5,58
Länge		
≤ 15	868	4,97
15-20	1.703	5,27
20-25	1.038	5,39
25-30	1.942	5,52
>30	440	5,59
Tandem	243	5,52
Zeit		
[1,08; 3,83)	1.583	5,06
[3,83; 6,17)	1.538	5,24
[6,17; 8,50)	1.614	5,52
[8,50;10,50]	1.499	5,62
Overall	6.234	5,36

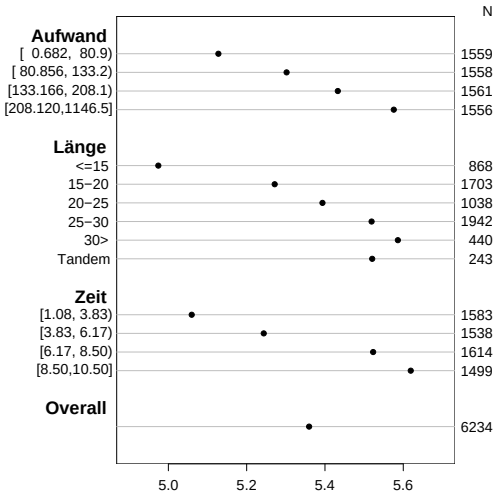


Abb. 3.22: Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“

Tab. 3.31: Wald- und Likelihood-Ratio-Teststatistiken für die Variable *Gefälligkeit* im Werbeträger „TV“

	df	*Wald* χ^2	*P-Wert*	*LR* χ^2	*P-Wert*
Aufwand	10	176,3521	0,0000	174,5943	0,0000
Aufwand nichtlinear	3	79,3331	0,0000	79,1479	0,0000
Zeit	10	985,4540	0,0000	918,3343	0,0000
Zeit nichtlinear	3	103,9174	0,0000	103,4723	0,0000
Länge	15	354,9881	0,0000	346,6011	0,0000
Aufwand × *Zeit*	1	8,4008	0,0038	8,4289	0,0037
Aufwand × *Länge*	5	15,6185	0,0080	15,6617	0,0079
Zeit × *Länge*	5	23,9006	0,0002	23,9508	0,0002
TOTAL nichtlinear	6	180,6910	0,0000	178,8289	0,0000
TOTAL WW	11	61,9148	0,0000	61,8562	0,0000
TOTAL nichtlinear+WW	17	243,1361	0,0000	239,4567	0,0000
Regression	24	1798,0723	0,0000	1585,4641	0,0000

steigen sie bei den restlichen Kategorien bis zu einem Aufwand von 200.000 € und bleiben danach annähernd konstant. Die Funktionen liegen bei längerer Dauer stets höher.
Wald- und Likelihood-Ratio-Test zeigen für die weiblichen Befragten und für die Altersgruppe „50 und älter“ keine signifikante Wechselwirkung zwischen dem Aufwand und der Zeit. In den restlichen Gruppen haben alle Variablen einen signifikanten Einfluss.
In der Längenkategorie „≤ 15“ gibt es keine Unterschiede. In den restlichen Längenkategorien ist auffällig, dass in der Gruppe der Frauen die Werte bei geringem Aufwand und kleiner Zeit höher liegen. Dies führt dazu, dass der Verlauf in den ersten Jahren dem der letzten Jahre ähnlich ist. Diese Auffälligkeit zeigt auch schon die nicht signifikante Wechselwirkung zwischen Aufwand und Zeit. Genau in diesem Bereich hat die Zielgruppe der Männer leicht geringere Gefälligkeitswerte. Sonst sind die Unterschiede zwischen den Gesamtbefragten und den Männern nur minimal.
In der Altergruppe „50 und älter“ gibt es ein ähnliches Phänomen wie bei der Gruppe der Frauen. Auch hier ist die Wechselwirkung ja nicht signifikant. Allerdings liegen jetzt die Gefälligkeitswerte bei hohem Aufwand in den Jahren 1998 bis 2000 unter denen der restlichen Gruppen. Die Klasse der 14- bis 29-Jährigen hat genau in diesem Bereich die Sujets besser bewertet. Die mittlere Zielgruppe unterscheidet sich von den insgesamt Befragten kaum.

3.4 Bewertungsvergleich der Wirtschaftsbereiche

Ob sich die einzelnen Sujets in ihrer Beurteilung in den einzelnen Wirtschaftsbereichen unterscheiden, kann mittels Varianzanalyse überprüft werden. Einen ersten Einblick bekommt man beim Betrachten der Abbildung 3.23. Diese Abbildung stellt die einzelnen Wirtschaftsbereiche in einem Boxplot dar.
Man kann bereits erkennen, dass es Unterschiede zwischen den Gruppen gibt. Der Bereich „Touristik“ wird im Mittel am besten bewertet, dicht gefolgt vom Bereich „Getränke“. Am schlechtesten wird der Wirtschaftsbereich „Reinigung“ bewertet. Ob und welche Unterschiede aber signifikant sind, sieht man erst an der Varianzanalyse und den nachfolgenden Post-Hoc-Tests.
Bevor eine Varianzanalyse durchgeführt wird, werden die Voraussetzungen überprüft. Der Kolmogorov-Smirnov-Test untersucht, ob die Daten einer Normalverteilung gehorchen. Die Teststatistik ist in allen Gruppen nicht signifikant, man kann also von einer Normalverteilung der Daten ausgehen. Der Levene- oder Bartlett-Test überprüft die Nullhypothese, ob die Varianzen über die Gruppen hinweg gleich sind. Beide Tests führen zu dem Ergebnis, dass die Varianzen nicht homogen sind. Bereits in Tabelle 3.32 erkennt man,

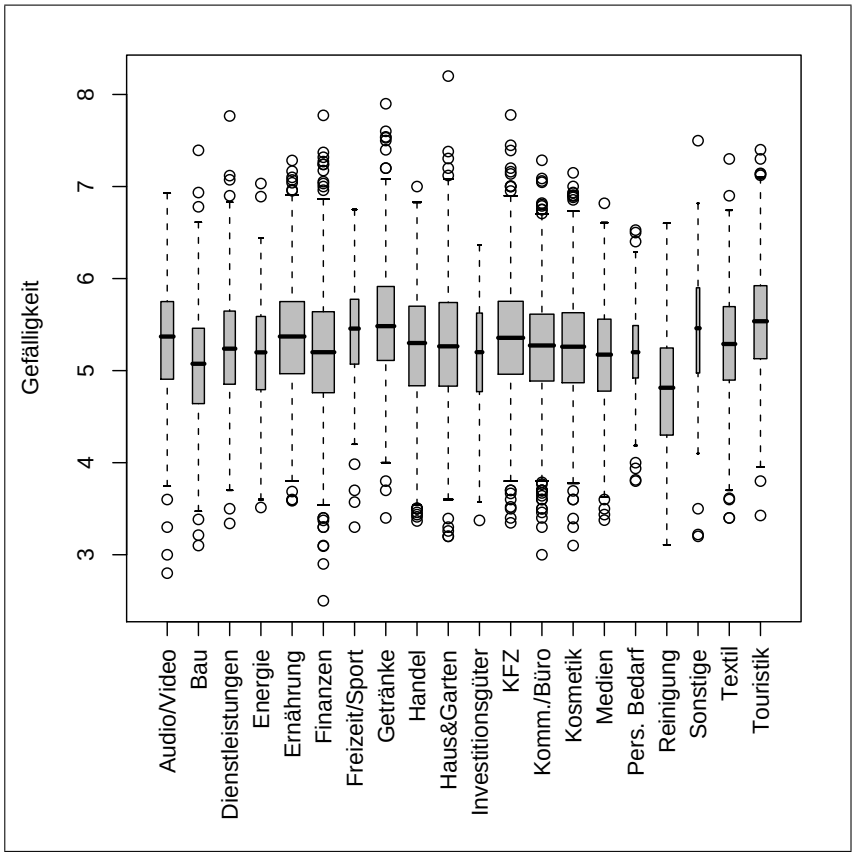


Abb. 3.23: Boxplot der Variablen *Gefälligkeit*, getrennt nach Wirtschaftsbereichen

dass die Stichprobenvarianzen in den Wirtschaftsbereichen unterschiedlich sind. Sie reichen von 0,2846 im Bereich „Persönlicher Bedarf“ bis hin zu 0,9546 im Bereich „Sonstiges“.

Betrachtet man, wie in Kapitel 2.2 besprochen, den Transformationsplot, der die Mittelwerte bzw. Mediane in Beziehung zur Standardabweichung stellt, so erkennt man eine Punktwolke, durch die man keine Gerade legen kann. Auch durch *Versuch und Irrtum* wird keine passende Transformation gefunden, welche die Varianzen stabilisiert. Aus diesem Grunde wird keine Varianzanalyse, sondern der nichtparametrische Kruskal-Wallis-Test durchgeführt.
Die Teststatistik liegt bei 589,12 und ist bei 19 Freiheitsgraden hochsignifikant. Die Nullhypothese, dass es keinen Unterschied zwischen den Gruppen gibt, kann also verworfen werden.

Nun ist man daran interessiert, zwischen welchen Gruppen der Unterschied

Tab. 3.32: Deskriptive Maßzahlen der Variablen *Gefälligkeit* in den einzelnen Wirtschaftsbereichen

	n	*Mittelwert*	*Median*	*Varianz*
Audio/Video	449	5,3053	5,3701	0,4053
Bau	396	5,0339	5,0744	0,4246
Dienstleistungen	359	5,2298	5,2385	0,4205
Energie	230	5,1711	5,1984	0,3996
Ernährung	1.697	5,3567	5,3709	0,3497
Finanzen	1.342	5,1961	5,2000	0,4644
Freizeit/Sport	202	5,4150	5,4567	0,3442
Getränke	803	5,5198	5,4833	0,4060
Handel	733	5,2555	5,3000	0,4521
Haus&Garten	938	5,2812	5,2640	0,4733
Investitionsgüter	95	5,1748	5,2005	0,4109
KFZ	1.745	5,3598	5,3564	0,3773
Komm./Büro	1.618	5,2464	5,2733	0,3270
Kosmetik	1.358	5,2410	5,2600	0,3447
Medien	492	5,1553	5,1737	0,3793
Pers. Bedarf	87	5,1968	5,2000	0,2846
Reinigung	466	4,7707	4,8140	0,4304
Textil	399	5,2920	5,2892	0,3722
Touristik	462	5,5385	5,5364	0,3748
Sonstige	31	5,3425	5,4605	0,9564
NA	3	5,3940	5,7120	0,7011
Overall	13.905	5,2753	5,2980	0,4103

besteht. Mit dem Wilcoxon-Rangsummen-Test werden nun alle Paarvergleiche durchgeführt und deren P-Werte mit der Bonferroni-Methode korrigiert. Dies ist relativ aufwändig, da es 20 Gruppen und somit 190 Vergleiche gibt. In Tabelle 3.33 sieht man die 76 signifikanten Ergebnisse der Wilcoxon-Tests. Die ersten beiden Spalten geben die untere bzw. obere Grenze des Konfidenzintervalls an. Beide Werte sind entweder positiv oder negativ. Würden die Intervalle den Wert 0 einschließen, nähme man die Nullhypothese an. An den Konfidenzintervallen kann man ablesen, welche Gruppe besser bewertet wird. Sind die Werte positiv, so wird die zuerst genannte Gruppe besser bewertet, sind sie negativ, so ist es genau umgekehrt.

Schon in Abbildung 3.23 erkennt man, dass die Boxen der Wirtschaftsbereiche „Getränke“ und „Touristik“ sehr weit oben liegen. Der Bereich „Touristik“ hat den größten Median. Sujets aus diesem Bereich werden signifikant besser

bewertet als Sujets aus den anderen Bereichen, ausgenommen die Bereiche „Freizeit/Sport“, „Getränke“ und „Sonstige“.
Testet man die Unterschiede zum Bereich „Getränke“, so sieht man, dass diese Sujets auch besser beurteilt werden. Kein kennzeichnender Unterschied konnte zu den Bereichen „Freizeit/Sport“, „Sonstige“ und natürlich „Touristik“ festgestellt werden.
Der Bereich „Sonstige“ liegt zwar auch im oberen Bereich, hat aber lediglich 31 Beobachtungen. Dies ist auch ein Grund, warum ein signifikanter Unterschied nur zum Bereich „Reinigung“ aufgedeckt werden kann. Außerdem ist eine Interpretation bezüglich der Gefälligkeit insofern problematisch, als in dieser Kategorie alle möglichen Werbungen zusammengefasst werden können. Sujets aus dem Bereich „Reinigung“ werden schlecht beurteilt. Sie haben signifikant schlechtere Gefälligkeitswerte als alle anderen Wirtschaftsbereiche. Ebenfalls im unteren Bereich liegen die Spots aus der Kategorie „Bau“. Sie weisen keine signifikant schlechteren Werte auf als Spots aus den Kategorien „Reinigung“, „Medien“, „Energie“, „Persönlicher Bedarf“ und „Investitionsgüter“. Im Vergleich mit den restlichen Kategorien ist die Bewertung signifikant schlechter. Die restlichen signifikanten Ergebnisse kann man in Tabelle 3.33 nachlesen. Die Einträge in den ersten beiden Spalten geben die Grenzen des Konfidenzintervalls an. Sind sie positiv, so wird der zuerst genannte Wirtschaftsbereich besser bewertet und umgekehrt.

Tab. 3.33: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Gefälligkeit* in den einzelnen Wirtschaftsbereichen

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
Audio/Video-Bau	0,1242	0,4435	111324	0
Audio/Video-Getränke	-0,3133	-0,0334	152226	0
Audio/Video-Medien	0,0129	0,3151	126939	0,0001
Audio/Video-Reinigung	0,3866	0,7078	151553	0
Audio/Video-Touristik	-0,3604	-0,0574	83749,5	0
Bau-Dienstleistungen	-0,3565	-0,0136	59251	0,0001
Bau-Ernährung	-0,4368	-0,1819	241422	0
Bau-Finanzen	-0,2919	-0,0128	230699,5	0,0001
Bau-Freizeit/Sport	-0,5737	-0,1815	26090	0
Bau-Getränke	-0,6007	-0,3096	95543	0
Bau-Handel	-0,3812	-0,0739	116955,5	0
Bau-Haus&Garten	-0,3762	-0,0786	149818,5	0
Bau-KFZ	-0,4375	-0,1815	249135,5	0

Fortsetzung auf nächster Seite

Tab. 3.33 – Fortsetzung von voriger Seite

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
Bau-Komm./Büro	-0,3279	-0,0781	258885,5	0
Bau-Kosmetik	-0,3247	-0,0682	219506	0
Bau-Reinigung	0,0973	0,4333	113108	0
Bau-Textil	-0,4067	-0,0744	62002	0
Bau-Touristik	-0,6497	-0,3348	51543	0
Dienstleistungen-Getränke	-0,4191	-0,1206	109292	0
Dienstleistungen-Reinigung	0,28	0,6261	115235	0
Dienstleistungen-Touristik	-0,4694	-0,1465	59754	0
Energie-Ernährung	-0,3334	-0,0106	164184	0,0001
Energie-Freizeit/Sport	-0,46	-0,0238	17988,5	0,0001
Energie-Getränke	-0,4976	-0,1429	66217	0
Energie-KFZ	-0,336	-0,0094	169197	1e-04
Energie-Reinigung	0,2053	0,6007	71747,5	0
Energie-Touristik	-0,5393	-0,168	36042,5	0
Ernährung-Finanzen	0,0733	0,2452	1300027,5	0
Ernährung-Getränke	-0,2413	-0,0474	589506,5	0
Ernährung-Komm./Büro	0,0291	0,179	1513319	0
Ernährung-Kosmetik	0,0328	0,192	1276877	0
Ernährung-Medien	0,0745	0,3058	491269	0
Ernährung-Reinigung	0,4533	0,7029	586738	0
Ernährung-Touristik	-0,3	-0,0677	323689,5	0
Finanzen-Freizeit/Sport	-0,4029	-0,0533	107527,5	0
Finanzen-Getränke	-0,4133	-0,199	396450	0
Finanzen-KFZ	-0,2458	-0,0734	1005925	0
Finanzen-Reinigung	0,2859	0,5557	419949	0
Finanzen-Touristik	-0,4691	-0,2155	217420	0
Freizeit/Sport-Komm./Büro	0,0173	0,3316	191969,5	0,0001
Freizeit/Sport-Kosmetik	0,0197	0,3441	161507,5	0
Freizeit/Sport-Medien	0,0776	0,4441	62015,5	0
Freizeit/Sport-Reinigung	0,4488	0,8475	72400	0
Getränke-Handel	0,1068	0,3561	353353,5	0
Getränke-Haus&Garten	0,1161	0,3494	452202,5	0
Getränke-Investitionsgüter	0,0506	0,5718	48530,5	0
Getränke-KFZ	0,0467	0,242	793589	0
Getränke-Komm./Büro	0,1533	0,3427	802403	0
Getränke-Kosmetik	0,1561	0,356	675855,5	0

Fortsetzung auf nächster Seite

Tab. 3.33 – Fortsetzung von voriger Seite

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
Getränke-Medien	0,203	0,47	257331	0
Getränke-Pers. Bedarf	0,0601	0,5361	45289	0
Getränke-Reinigung	0,5865	0,8722	295744,5	0
Getränke-Textil	0,0767	0,3548	191937	0
Handel-Reinigung	0,3443	0,6446	238294	0
Handel-Touristik	-0,4081	-0,1255	129653,5	0
Haus&Garten-Reinigung	0,3531	0,6411	305682	0
Haus&Garten-Touristik	-0,4039	-0,1333	165810,5	0
Investitionsgüter-Reinigung	0,1336	0,6979	29751	0
Investitionsgüter-Touristik	-0,607	-0,0797	15169,5	0
KFZ-Komm./Büro	0,0283	0,1793	1554039,5	0
KFZ-Kosmetik	0,0319	0,1917	1311482,5	0
KFZ-Medien	0,0747	0,3067	504555,5	0
KFZ-Reinigung	0,4528	0,705	601882	0
KFZ-Touristik	-0,3	-0,0666	334033	0
Komm./Büro-Reinigung	0,3495	0,5952	531947	0
Komm./Büro-Touristik	-0,3999	-0,1758	269407,5	0
Kosmetik-Reinigung	0,3381	0,5926	443133	0
Kosmetik-Touristik	-0,4128	-0,1784	225919,5	0
Medien-Reinigung	0,2296	0,5438	152340,5	0
Medien-Touristik	-0,5185	-0,2266	74877	0
Pers. Bedarf-Reinigung	0,1533	0,6878	28030,5	0
Pers. Bedarf-Touristik	-0,5791	-0,1047	13158,5	0
Reinigung-Sonstige	-1,1733	-0,0624	4138,5	0,0001
Reinigung-Textil	-0,6787	-0,3455	52915	0
Reinigung-Touristik	-0,9194	-0,6034	42063,5	0
Textil-Touristik	-0,4073	-0,1	70415	0

3.5 Vergleich der Werbeträger

In diesem Abschnitt soll die Frage beantwortet werden, wie die Werbeträger wahrgenommen werden. Dies wird beantwortet, indem man mittels Gruppenvergleichen überprüft, ob sich die Werbeträger bezüglich der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* unterscheiden.

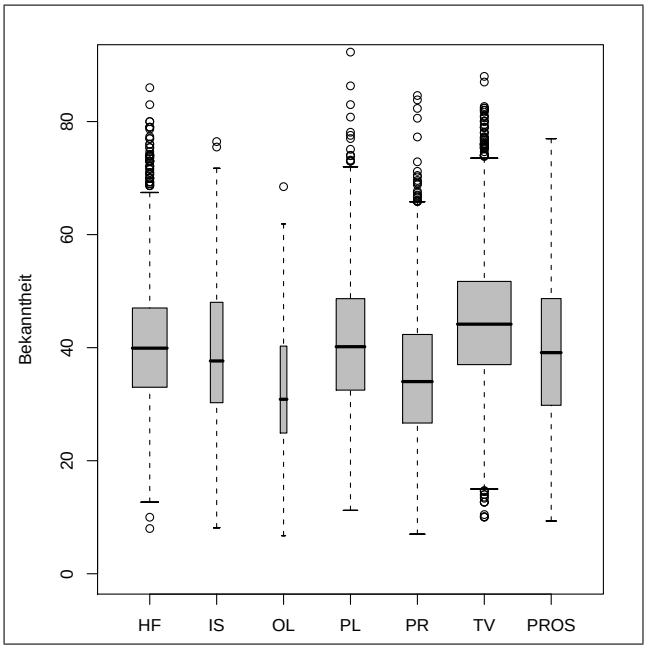


Abb. 3.24: Boxplot der Variablen *Bekanntheit*, getrennt nach Werbeträgern

In Abbildung 3.24 sieht man, wie sich die Variable Bekanntheit in den einzelnen Werbeträgern verhält. Die Breite der Boxen ist ein Indikator dafür, wie viele Beobachtungen in der Gruppe vorhanden sind. Die meisten Beobachtungen gibt es beim Werbeträger „TV“, die wenigsten beim Werbeträger „OL“.
Sowohl in Abbildung 3.24 als auch in Tabelle 3.34 erkennt man, dass der Werbeträger „TV“ die höchsten Werte hat, also am bekanntesten ist. Dann folgen in absteigender Reihenfolge „Plakat“, „Hörfunk“, „Prospekt“, „Infoscreen“, „Print“ und schlussendlich „Online“.

Der Kruskal-Wallis-Test klärt, ob die Unterschiede in zumindest einer Gruppe signifikant sind. Eine Varianzanalyse sollte in diesem Fall nicht gerechnet werden, da die Variablen in den einzelnen Gruppen nicht normalverteilt sind. Dies wäre nicht so fatal, da eine Varianzanalyse gegen Nichtnormalität robust ist, doch die Varianzen sind über die Gruppen hinweg nicht gleich. Weder durch *Versuch und Irrtum* noch durch einen Transformationsplot konnte eine Transformation gefunden werden, die dafür sorgt, dass die Varianzen innerhalb der Gruppen gleich sind.
Der Gruppenvergleich wird also mittels Kruskal-Wallis-Test durchgeführt. Dieser liefert eine Teststatistik von 1.117,8 (P-Wert: 0,0000). Deshalb werden anschließend Post-Hoc-Tests durchgeführt. Es gibt $\binom{7}{2} = 21$ Paarvergleiche,

Tab. 3.34: Deskriptive Maßzahlen der Variablen *Bekanntheit* in den einzelnen Werbeträgern

	n	*Mittelwert*	*Median*	*Varianz*
HF	2.633	40,6624	39,9099	124,5870
IS	346	39,5778	37,6481	170,1676
OL	96	32,3705	30,8654	131,1859
PL	1.788	40,9694	40,1697	139,6469
PR	1.915	34,9889	34,0000	140,1702
PROS	848	39,7980	39,1206	172,2194
TV	6.279	44,6849	44,1571	124,1259
Overall	13.905	41,6000	41,0676	143,9065

deren Ergebnisse in Tabelle 3.35 zusammengefasst sind.
Die Sujets des Werbeträgers „TV“ sind signifikant bekannter als die der restlichen Werbeträger. Plakat-Werbungen und Hörfunk-Spots sind in der Bekanntheit an zweiter bzw. dritter Stelle. Der Unterschied zu den Online-Sujets und Print-Anzeigen ist signifikant. Anschließend folgen in der Bekanntheit Prospekte und Infoscreens. Ihr Bekanntheitsgrad ist ebenfalls höher als der von Online- und Print-Werbungen. Die beiden Mediengattungen „Online“ und „Print“ weisen untereinander keinen Unterschied in der Bekanntheit auf, haben aber einen wesentlich geringeren Bekanntheitsgrad als die restlichen fünf Kategorien.

Für die Variable *Markenimpact* kann der Vergleich auch wegen der fehlenden Varianzhomogenität nicht mit einer Varianzanalyse angestellt werden. Vor dem Vergleich mittels Kruskal-Wallis-Test geben Tabelle 3.36 und Abbildung 3.25 einen Eindruck, wie sich die Variable *Markenimpact* in den einzelnen Werbeträgern verhält. Die Prospekt-Sujets haben den größten Mittelwert und Median, gefolgt von den Online-Werbungen. Ungefähr auf gleicher Höhe liegen Plakat-, Print- und TV-Sujets. Die kleinsten Lageparameter haben Hörfunk-Spots.

Der Kruskal-Wallis-Test hat eine Teststatistik von 563,8038 und bei 6 Freiheitsgraden einen P-Wert nahe 0. Also ist auch hier die Frage interessant, welche Werbeträger unterschiedlich sind.

Tabelle 3.37 zeigt die Ergebnisse der 16 signifikanten Wilcoxon-Tests. Die Hörfunk-Sujets konnten sich die Befragten nach dem Interview am schlechtesten ins Gedächtnis rufen. Die Unterschiede sind zu allen sechs Werbeträgern

Tab. 3.35: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Bekanntheit* in den einzelnen Werbeträgern

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
HF-OL	4,8032	11,8252	178557	0
HF-PR	4,6679	6,791	3230568	0
HF-TV	-4,9756	-3,3858	6479378	0
IS-OL	2,7467	11,3381	22071	0
IS-PR	2,2147	6,6893	398267	0
IS-TV	-7,5514	-3,3628	816601,5	0
OL-PL	-12,3415	-4,9287	50584	0
OL-TV	-16,0186	-8,969	127876	0
OL-PROS	-11,612	-3,338	27279	0
PL-PR	4,8797	7,2843	2199527	0
PL-TV	-4,7807	-2,9095	4552123	0
PR-TV	-10,8465	-8,9999	3253513	0
PR-PROS	-6,4948	-3,2728	638164,5	0
TV-PROS	3,5974	6,4191	3258869,5	0

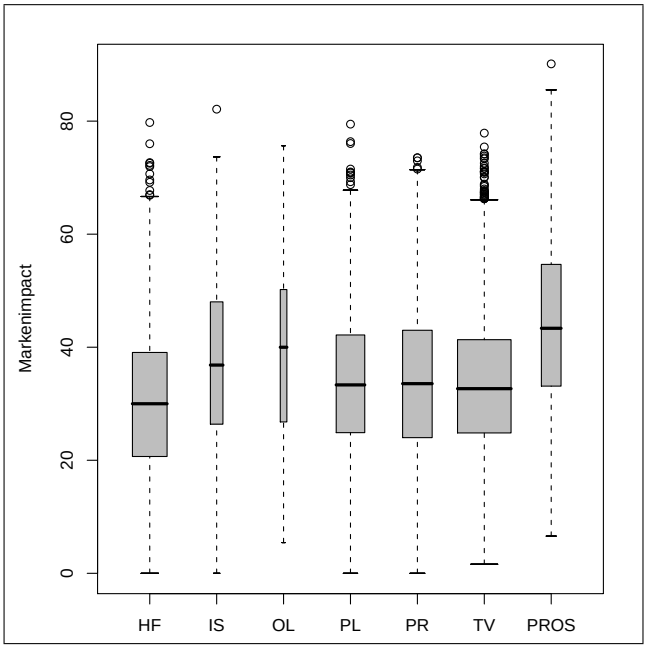


Abb. 3.25: Boxplot der Variablen *Markenimpact*, getrennt nach Werbeträgern

Tab. 3.36: Deskriptive Maßzahlen der Variablen *Markenimpact* in den einzelnen Werbeträgern

	n	*Mittelwert*	*Median*	*Varianz*
HF	2.631	30,4686	30,0000	175,5185
IS	346	37,1547	36,8421	229,3361
OL	96	39,4950	39,9954	266,1194
PL	1.788	33,3346	33,3333	171,1139
PR	1.913	33,4791	33,5526	190,1347
TV	6.279	33,3147	32,6667	152,7895
PROS	848	44,1212	43,3486	225,0658
Overall	13.901	33,5987	33,0043	180,8071

signifikant. Am öftesten können die Befragten Sujets der Kategorie „Prospekte“ nennen, der Durchschnitt liegt bei 44,12 % und ist signifikant höher als die anderen. Dies konnte man auch schon anhand Tabelle 3.36 und Abbildung 3.25 vermuten. Nicht schlechter schneiden die Online-Sujets ab. Zu den Prospekten kann kein signifikanter Unterschied aufgedeckt werden, auch nicht zu den Infoscreen-Spots, jedoch sehr wohl zu den restlichen, nämlich Hörfunk, Plakat, Print und TV. Kein Unterschied konnte zwischen den Werbeträgern „Print“, „Plakat“ und „TV“ festgestellt werden. Werbungen aus diesen drei Kategorien blieben den Befragten eher im Gedächtnis als Hörfunk-Sujets, aber seltener als Prospekt-, Infoscreen- und Online-Werbungen.
Nun werden die Unterschiede in den Werbeträgern bezüglich der Variablen *Gefälligkeit* betrachtet. Einen ersten Eindruck vermittelt Abbildung 3.26, in der ein Boxplot gezeichnet ist, dessen Breite der Boxen wieder auf die Größe der Gruppe schließen lässt. Im unteren Bereich liegen die Werbeträger „Hörfunk“, „Online“, „Plakat“ und „Print“. Die höchste Bewertung erhalten Werbeprospekte, dicht gefolgt von Infoscreen-Spots.

Ein Vergleich mit einer Varianzanalyse ist nicht gerechtfertigt, da die Varianzen nicht homogen sind und dies auch nicht durch eine Transformation erreicht werden kann. Der Kruskal-Wallis-Test hat ein signifikantes Ergebnis, seine Teststatistik ist 502,8196. Deshalb werden anschließend Post-Hoc-Tests ausgeführt, um zu überprüfen, zwischen welchen Gruppen die Unterschiede bestehen. Die signifikanten Ergebnisse sind in Tabelle 3.39 zusammengefasst.

Dass Prospekt-Sujets den höchsten Mittelwert bzw. Median bezüglich der Variablen *Gefälligkeit* besitzen, zeigt schon Tabelle 3.38. Der Unterschied ist nur zu den Infoscreen-Werbungen nicht signifikant.

Tab. 3.37: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Markenimpact* in den einzelnen Werbeträgern

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
HF-IS	-9,3333	-4,1532	338087,5	0
HF-OL	-13,9149	-3,9798	85633	0
HF-PL	-4,4075	-1,902	2039191	0
HF-PR	-4,5191	-1,9868	2183849	0
HF-TV	-3,8702	-2	7211446,5	0
HF-PROS	-15,3528	-11,8327	560241,5	0
IS-PL	0,9892	6,246	353159	0
IS-PR	0,8436	6,2356	375440	0,0001
IS-TV	1,3947	6,236	1252463	0
IS-PROS	-9,9337	-3,8425	110249,5	0
OL-PL	0,9164	10,8995	104444	0,0003
OL-PR	0,7004	10,8969	110941	0,0006
OL-TV	1,2468	10,9108	369114	0,0002
PL-PROS	-12,3333	-8,6318	455230,5	0
PR-PROS	-12,2818	-8,5046	497957	0
TV-PROS	-12,236	-9,064	1562425,5	0

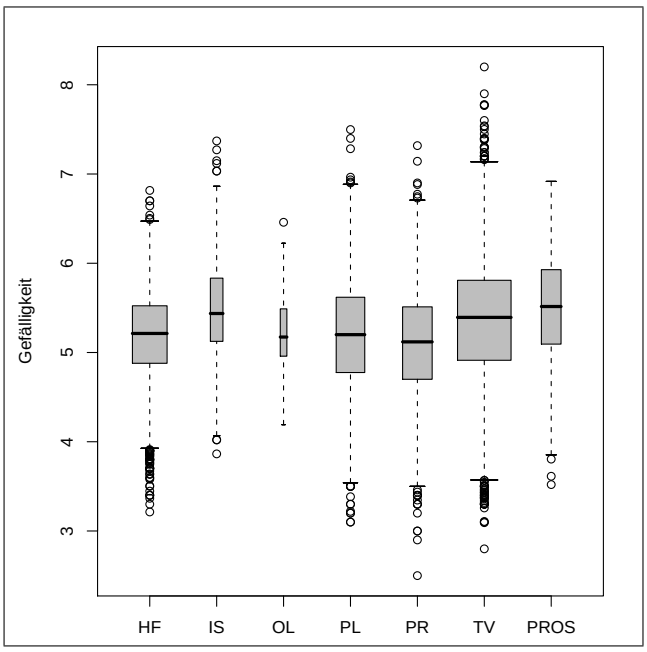


Abb. 3.26: Boxplot der Variablen *Gefälligkeit*, getrennt nach Werbeträgern

Tab. 3.38: Deskriptive Maßzahlen der Variablen *Gefälligkeit* in den einzelnen Werbeträgern

	n	*Mittelwert*	*Median*	*Varianz*
HF	2.633	5,1727	5,2133	0,2660
IS	346	5,4721	5,4370	0,3788
OL	96	5,2550	5,1732	0,1843
PL	1.788	5,1948	5,2000	0,4078
PR	1.915	5,0873	5,1192	0,3679
TV	6.279	5,3582	5,3933	0,4676
PROS	848	5,4953	5,5153	0,3345
Overall	13.905	5,2753	5,2980	0,4103

Ebenso ist der Unterschied zwischen Infoscreen- und TV-Spots nicht signifikant, sehr wohl aber jener zwischen TV-Spots und Prospekten und jener zwischen Infoscreen und den restlichen vier Mediengattungen. Am schlechtesten bewerten die Befragten die Sujets aus dem Print-Bereich, die Differenz ist nur zu Online-Werbungen nicht signifikant. Diese Ergebnisse fasst Tabelle 3.39 zusammen.

3.6 Effizienzvergleich der Werbeträger

In diesem Abschnitt soll die Frage behandelt werden, welcher Werbeträger am effizientesten ist.
Die Effizienz ist der Quotient der Variablen *Aufwand* und *Bekanntheit*. Sie gibt also an, wie viel für einen Prozentpunkt Bekanntheit aufgewendet werden muss. Die Einheiten sind in tausend Euro angegeben. Sie kann nur bei den Werbeträgern „Hörfunk“, „Infoscreen“, „Print“ und „TV“ berechnet werden. Ein Vergleich bezüglich der Online- und Plakat-Sujets ist demnach nicht möglich.

Abbildung 3.27 und Tabelle 3.40 zeigen, dass die Verteilung der Variablen sehr schief ist. Diese Schiefe ist typisch für eine Verhältniszahl. Außerdem ist schon die Verteilung der Variablen *Aufwand* sehr schief, was sich in der *Effizienz* natürlich widerspiegelt. Die meisten Beobachtungen, nämlich 8.904 von insgesamt 11.134, dies sind knappe 80 %, liegen unter dem Wert 5. Bei annähernd 9.000 Sujets werden also pro Prozentpunkt Bekanntheit weniger als 5.000 € aufgewendet.
Diese Tatsache lässt einen Gruppenvergleich mittels Varianzanalyse nicht zu, da nicht von einer Normalverteilung ausgegangen werden kann. Außerdem

Tab. 3.39: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Gefälligkeit* in den einzelnen Werbeträgern

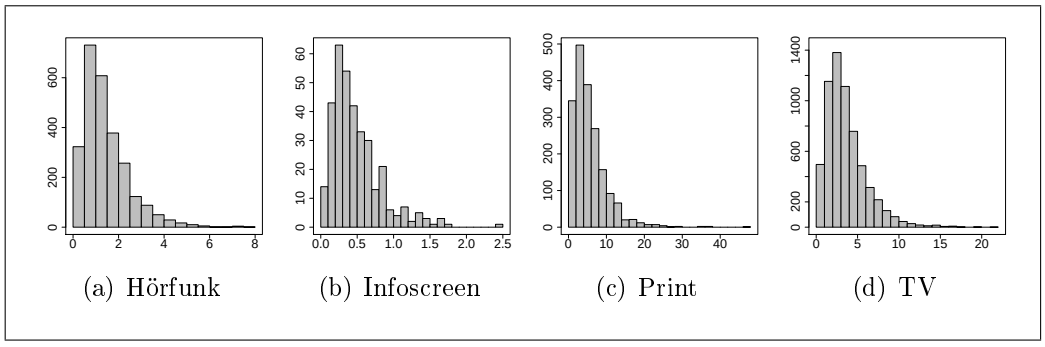
	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
HF-IS	-0,364	-0,1725	328066	0
HF-PR	0,0323	0,1339	2737025,5	0
HF-TV	-0,2271	-0,1422	6815266,5	0
HF-PROS	-0,3867	-0,2508	761311	0
IS-OL	0,0371	0,3935	20622	0,0003
IS-PL	0,1489	0,3725	383152	0
IS-PR	0,2499	0,4649	442426,5	0
OL-PROS	-0,4354	-0,0828	29523	0
PL-PR	0,036	0,162	1867793,5	0
PL-TV	-0,2225	-0,1134	4801061,5	0
PL-PROS	-0,3837	-0,2252	549773	0
PR-TV	-0,32	-0,217	4597478,5	0
PR-PROS	-0,48	-0,3277	511564	0
TV-PROS	-0,2078	-0,0634	2341801,5	0

lehnen sowohl Bartlett- als auch Levene-Test die Nullhypothese ab, dass die Varianzen in den Gruppen homogen sind.

Um die Daten symmetrischer zu machen, wurde versucht, sie zu transformieren. Man erreicht eine symmetrische Verteilung, die einer Normalverteilung sehr ähnlich ist, bei einer Potenztransformation mit der Potenz 0,25. Dies entspricht dem Ziehen der vierten Wurzel. In diesem Fall nimmt der Kolmogorov-Smirnov-Test seine Nullhypothese an. Außerdem zeigt die graphische Betrachtung mittels Histogramms und QQ-Plots, dass von einer Normalverteilung ausgegangen werden kann. Bei dieser Transformation liegt allerdings keine Varianzhomogenität vor. Es wird auch in diesem Fall keine Transformation der Daten gefunden, welche die Anwendung der Varianzanalyse rechtfertigen würde. Man muss also auf einen nichtparametrischen Vergleich zurückgreifen.
Zuvor zeigt Abbildung 3.28, dass Sujets aus dem Werbeträger „Infoscreen“ die geringsten Werte der Variablen *Effizienz* aufweisen. Die gesamte Box des Werbeträgers „Infoscreen“ liegt unter jener des Werbeträgers „Hörfunk“. Diese befindet sich wiederum vollständig unter der Box der TV-Werbungen. Die vier Werbeträger scheinen sich also hinsichtlich der Effizienz zu unterscheiden.

Tab. 3.40: Deskriptive Maßzahlen der Variablen *Effizienz*, getrennt nach Werbeträgern

	Hörfunk	*Infoscreen*	*Print*	*TV*	*Gesamt*
n	2.626	346	1.894	6.268	11.134
NA's	7	0	21	11	39
Minimum	0,0038	0,0121	0,0172	0,0251	0,0038
1. Quartil	0,7448	0,2548	2,5691	1,9434	1,3228
Median	1,2076	0,3984	4,4589	3,0859	2,5561
Mittelwert	1,4369	0,4885	5,5480	3,6163	3,3337
3. Quartil	1,8529	0,6359	7,2807	4,6967	4,4167
Maximum	7,6746	2,4301	46,3712	21,5267	46,3712
Stabw	0,9987	0,3420	4,4629	2,4297	2,9880
Schiefe	1,6433	1,6448	2,3196	1,6055	2,6955

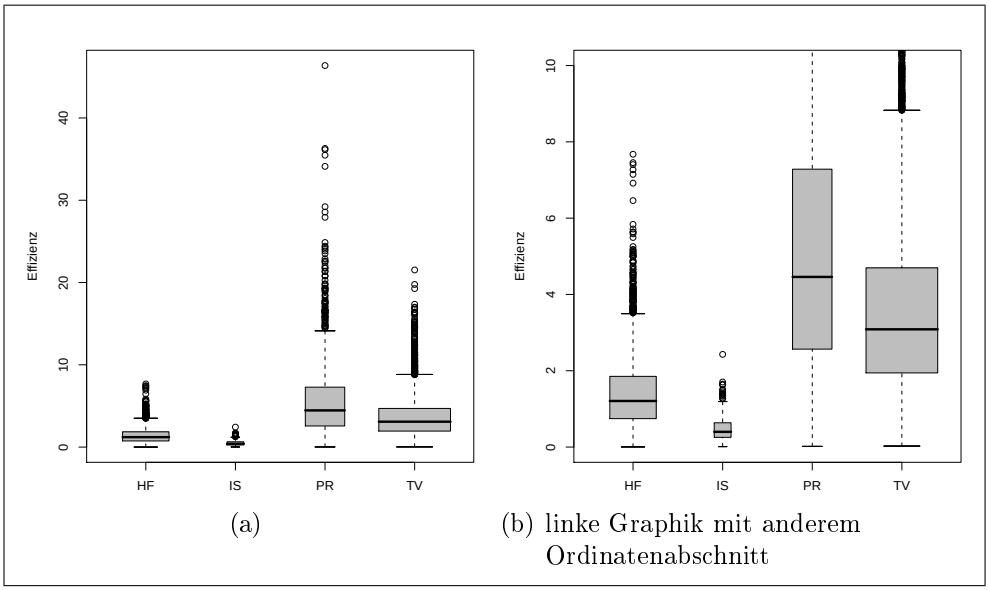


(a) Hörfunk (b) Infoscreen (c) Print (d) TV

Abb. 3.27: Histogramm der Variablen *Effizienz*, getrennt nach Werbeträgern

Dies bestätigt auch der Kruskal-Wallis-Test. Er führt mit einer Teststatistik von 3.649,772 bei 3 Freiheitsgraden zu einer Signifikanz nahe 0. Deshalb werden anschließend alle möglichen Paarvergleiche mittels Wilcoxon-Rangsummentests durchgeführt. Alle sechs Tests sind signifikant, wie Tabelle 3.41 zeigt.

Der effizienteste Werbeträger ist der Infoscreen. Hier muss man für einen Prozentpunkt Bekanntheit am wenigsten Geld investieren. Der Unterschied ist zu allen Werbeträgern signifikant. Hörfunk-Sujets sind effizienter als die Print- und TV-Werbungen. TV-Spots werden hinsichtlich der Effizienz nur besser als die Sujets des Werbeträgers „Print“ eingestuft. Das Print-Medium schneidet bei diesem Vergleich am schlechtesten ab und ist somit der „ineffizienteste“ Werbeträger. Seine Werte sind signifikant schlechter.



(a) (b) linke Graphik mit anderem Ordinatenabschnitt

Abb. 3.28: Boxplot der Variablen *Effizienz*, getrennt nach Werbeträgern

Tab. 3.41: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Effizienz* in den einzelnen Werbeträgern

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
HF-IS	0,6556	0,8428	780.391	0
HF-PR	-3,3266	-2,9245	603.922	0
HF-TV	-1,8654	-1,6768	2.833.281	0
IS-PR	-4,3985	-3,6133	16.706	0
IS-TV	-2,8309	-2,4027	68.176	0
PR-TV	1,087	1,4694	7.576.249,5	0

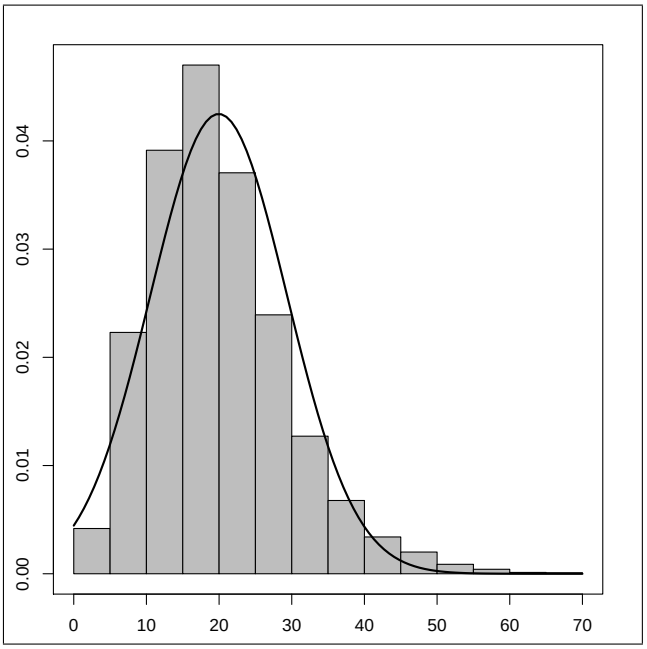


Abb. 3.29: Histogramm der Variablen *Sympathie*

3.7 Sympathievergleich der Werbeträger

In diesem Abschnitt sollen die Werbeträger in Hinblick auf ihre Sympathie untersucht werden. Gibt es einen Unterschied in der Sympathie, werden Sujets bestimmter Werbeträger als sympathischer angesehen bzw. zeichnen sie sich insbesondere durch Sympathie aus? Es ist zu beachten, dass die Variable *Sympathie* ein Polaritäten-Profil angibt. Es wurde nicht die Frage gestellt: „Finden Sie die Werbung sympathisch?“, sondern „Durch welche Kriterien zeichnet sich diese Werbung besonders aus?“. Die Befragten haben danach die zehn Imagevariablen vor sich gesehen und mussten jenes nennen, das für sie am ehesten zutrifft. Dies darf bei der Interpretation nicht außer Acht gelassen werden.

Das Histogramm in Abbildung 3.29 und die Werte in Tabelle 3.42 geben einen ersten Eindruck über die Verteilung der Variablen *Sympathie*. Sie ist leicht rechtsschief und erinnert nicht unbedingt an eine Normalverteilung. Auch die Varianzen bzw. Standardabweichungen in den Gruppen sind nicht homogen, wie Bartlett- und Levene-Test mit einer signifikanten Teststatistik von 396,758 bzw. 50,329 zeigen.

Auch hier kann keine Transformation der Variablen gefunden werden, sodass die Daten nicht mehr rechtsschief, sondern symmetrisch verteilt sind. Be-

Tab. 3.42: Deskriptive Maßzahlen der Variablen *Sympathie* in den einzelnen Werbeträgern

	HF	*IS*	*OL*	*PL*	*PR*	*TV*	*Gesamt*
n	2.633	346	96	1.788	1.915	6.279	13.057
NA's	0	0	0	0	0	0	0
Min	2,667	2,409	3,419	0,000	2,000	0,667	0,000
1.Qu	12,752	14,361	11,412	12,000	12,974	14,000	13,245
Median	17,333	19,427	16,901	17,647	18,000	20,000	18,667
Mittel	18,069	20,278	17,793	18,815	19,115	21,323	19,946
3.Qu	22,555	24,262	21,930	24,000	24,000	27,000	25,102
Max	48,000	69,289	40,503	67,177	69,000	69,000	69,289
Stabw	7,285	8,880	8,191	9,303	8,743	10,185	9,389
Schiefe	0,573	1,216	0,502	0,968	0,894	0,799	0,901

Tab. 3.43: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Sympathie* in den einzelnen Werbeträgern

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
HF-IS	-3,1316	-0,5794	391365	0
HF-PR	-1,3948	0	2386343	0,0021
HF-TV	-3,2097	-2	6810287,5	0
IS-PL	0,1578	3,0663	343561,5	0,0011
OL-TV	-5,8066	-0,3314	242839	0,0011
PL-TV	-3,0675	-1,6209	4791447	0
PR-TV	-2,5996	-1,2035	5288809	0

trachtet man den Transformationsplot, so liegen die Punkte in einer Wolke, durch die man keine Gerade legen kann. Auch durch *Versuch und Irrtum* wird daraufhin ausprobiert, Varianzhomogenität herzustellen. Deshalb kann keine Varianzanalyse gerechnet werden, sondern es muss auf den Kruskal-Wallis-Test ausgewichen werden. Bevor man rechnerisch überprüft, ob ein Unterschied vorliegt, betrachtet man den Boxplot in Abbildung 3.30. Er zeigt, dass sich die Werbeträger nicht allzu sehr bezüglich ihrer Polarität unterscheiden. Den niedrigsten Wert hat der Werbeträger „Online“, am häufigsten zeichnen sich Fernsehwerbungen durch Sympathie aus (vergleiche auch Tabelle 3.42).

Die Kruskal-Wallis-Teststatistik von 231,3553 ist bei 5 Freiheitsgraden signifikant, weshalb im Anschluss die 15 Paarvergleiche durchgeführt werden. Sieben dieser Tests sind signifikant und in Tabelle 3.43 aufgelistet.

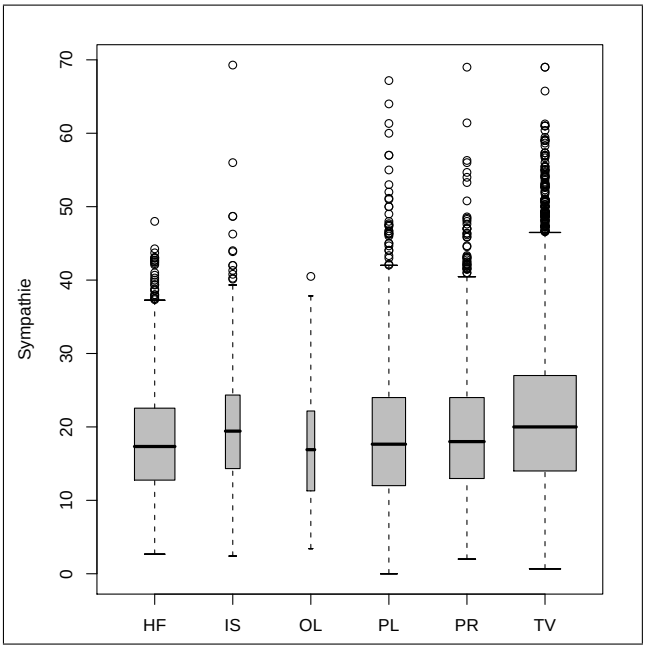


Abb. 3.30: Boxplot der Variablen *Sympathie*, getrennt nach Werbeträgern

Der Werbeträger „TV“ sticht am ehesten durch Sympathie heraus. Er hat sowohl den größten Mittelwert von 21,32 als auch den größten Median von 20 und hat signifikant höhere Werte als alle restlichen Gruppen. Der Unterschied konnte nur zu den Infoscreen-Spots nicht gefunden werden. Diese wiederum stechen den Befragten öfter durch Sympathie ins Auge als Hörfunk-Spots und Plakat-Werbungen. Print-Anzeigen zeichnen sich eher durch Sympathie aus als Hörfunk-Spots, zu anderen Werbeträgern konnte kein signifikanter Unterschied festgestellt werden. Hörfunk-Sujets wurden am seltensten als sympathisch eingestuft. Kennzeichnende Unterschiede gab es zu den Werbeträgern „Infoscreen“, „Print“ und „TV“. Online-Werbungen haben zwar sowohl den kleinsten Mittelwert von 17,793 als auch den kleinsten Median von 16,901 (vgl. Tabelle 3.42), es konnte aber trotzdem nur zu TV-Werbungen ein kennzeichnender Unterschied erkannt werden. Dies könnte vor allem daran liegen, dass es in dieser Gruppe sehr wenige Beobachtungen gibt.

3.8 Effizienzvergleich der Längenkategorien

In diesem Abschnitt soll die Kosten-Nutzen-Relation von Spotlängen eruiert werden. Die Kosten-Nutzen-Relation berechnet die Variable *Effizienz*. Es soll also geklärt werden, ob es einen Unterschied in der Effizienz in den einzelnen Kategorien der Variablen *Länge* gibt. Eine Beschreibung der Variablen *Effi-*

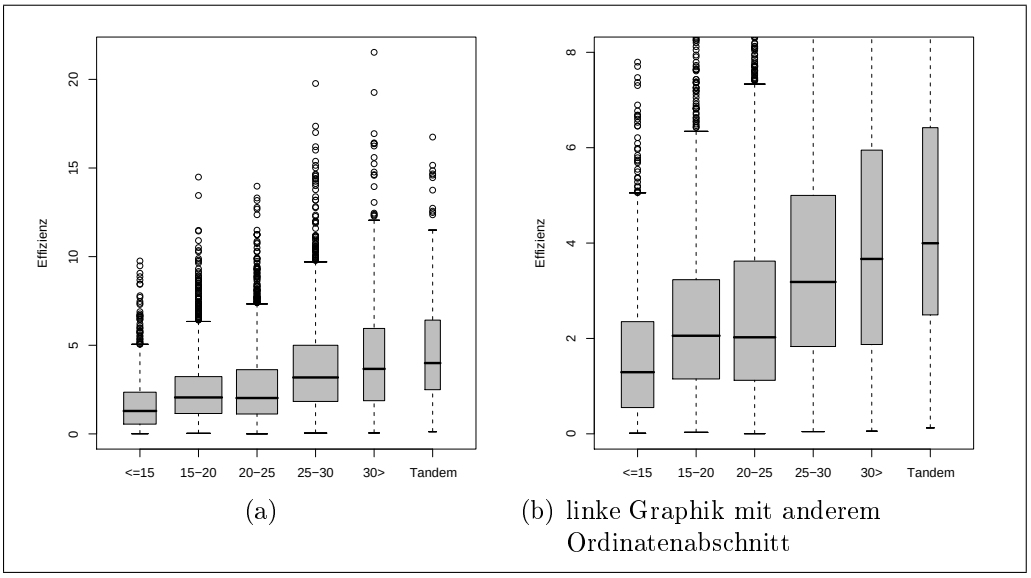
Tab. 3.44: Deskriptive Maßzahlen der Variablen *Effizienz*, getrennt nach Längenkategorien

	≤15	*15-20*	*20-25*	*25-30*	*>30*	*Tandem*	*Gesamt*
n	1.274	2.652	2.041	2.409	539	286	9.240
NA's	3	3	5	1	2	1	18
Min	0,0121	0,0306	0,0038	0,0470	0,0550	0,1255	0,0038
1.Qu	0,5507	1,1507	1,1228	1,8308	1,8730	2,4993	1,1995
Median	1,2926	2,0567	2,0244	3,1841	3,6672	3,9964	2,2923
Mittel	1,6666	2,3891	2,6519	3,6852	4,3832	4,8075	2,8798
3.Qu	2,3521	3,2305	3,6219	4,9995	5,9489	6,3915	3,8795
Max	9,7554	14,4896	13,9736	19,7679	21,5267	16,7459	21,5267
Stabw	1,4637	1,6812	2,1140	2,5370	3,3962	3,1675	2,3380
Schiefe	1,6987	1,5666	1,5232	1,3834	1,4229	1,1474	1,7672

zienz wurde bereits in Abschnitt 3.6 gegeben. Wie sich die Variable innerhalb der Längenkategorien verhält, sieht man in Tabelle 3.44. Es ist zu beachten, dass die Effizienz für diese Fragestellung nur für die Werbeträger „Hörfunk“, „Infoscreen“ und „TV“ berücksichtigt wird. Für die jeweils anderen Werbeträger gibt es keine messbare Länge. Die Effizienz besteht also in diesem Fall aus 9.258 Beobachtungen, wovon 18 fehlen.

Auch in dieser Fragestellung konnte man die Effizienz durch Ziehen der vierten Wurzel in eine symmetrische, normalverteilte Variable transformieren. Die Varianzen werden dadurch auch homogener, doch leider konnte auch diesmal keine Varianzheterogenität ausgeschlossen werden.
Am Boxplot in Abbildung 3.31 kann man schon erkennen, dass es einen Aufwärtstrend gibt. Je länger ein Spot also dauert, umso ineffizienter wird er. Dies ist aber wegen der höheren Kosten, die eine längere Werbung kostet, schlüssig. Laut Graphik ist also die Kategorie „≤ 15“ die effizienteste. Bei ihr braucht man die wenigsten Ausgaben, damit ein Prozent der Befragten die Werbung kennen.

Der Kruskal-Wallis-Test liefert mit einer Teststatistik von 1.172,939 ein signifikantes Ergebnis. Die signifikanten Ergebnisse der anschließend durchgeführten Post-Hoc-Tests sind in Tabelle 3.45 zusammengefasst.
Bei den Paarvergleichen gibt es 13 signifikante Ergebnisse. Es konnte nur zwischen den Kategorien „15 bis 20“ und „20 bis 25“ bzw. „länger als 30“ und „Tandem“ kein Unterschied festgestellt werden.
Die Kategorie „≤ 15“ hat signifikant niedrigere Werte als alle anderen Kate-



(a)

(b) linke Graphik mit anderem Ordinatenabschnitt

Abb. 3.31: Boxplot der Variablen *Effizienz*, getrennt nach Längenkategorien

Tab. 3.45: Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Effizienz* in den Längenkategorien

	untere Grenze	*obere Grenze*	*Teststatistik*	*P-Wert*
≤15/15-20	-0,7824	-0,5302	1182337	0
≤15/20-25	-0,8563	-0,5692	895713	0
≤15/25-30	-1,8992	-1,5257	692050	0
≤15/>30	-2,5107	-1,7894	148611	0
≤15/Tandem	-3,0379	-2,2042	55817	0
15-20/25-30	-1,1965	-0,8825	2155261	0
15-20/>30	-1,7994	-1,1371	446383	0
15-20/Tandem	-2,3595	-1,537	183539	0
20-25/25-30	-1,0935	-0,7452	1785737,5	0
20-25/>30	-1,6694	-0,9727	370433	0
20-25/Tandem	-2,2436	-1,3911	157866	0
25-30/>30	-0,7592	-0,0448	589829	0,0009
25-30/Tandem	-1,3527	-0,4458	271510	0

gorien. Sie ist also die effizienteste Kategorie. Die Kategorie mit Spotlängen zwischen 15 und 20 Sekunden hat signifikant niedrigere Effizienzwerte als Werbungen in den Kategorien „25 bis 30“, „länger als 30“ und „Tandem“. Nur zur nächsthöheren Kategorie konnte der Unterschied nicht als kennzeichnend abgegrenzt werden.
Die Kategorie „20 bis 25“ ist effizienter als alle höher liegenden Kategorien, ebenso gilt dies für die Kategorie „25 bis 30“. Die beiden ineffizientesten Kategorien sind „länger als 30“ und „Tandem“. Zwischen ihnen konnte, wie schon erwähnt, kein Unterschied festgestellt werden. Bei 50 % der Sujets in diesen Kategorien braucht man weniger als 4.400 € bzw. 4.800 €, um die Bekanntheit um einen Prozentpunkt zu steigern, in der Kategorie „$\leq$ 15“ reicht dafür schon ein Mehraufwand von 1.300 €.

4 Interpretation und Zusammenfassung

In der Analyse zur ersten Fragestellung fällt auf, dass der Aufwand auf die Bewertung eines Sujets nur bei den TV-Werbungen einen Einfluss hat (vgl. Tabelle 4.46). Diese Tatsache lassen auch schon die Abbildungen und Tabellen erkennen, welche die Mittelwerte getrennt für die Quartile bzw. die Faktorenlevel vergleichen. Man sieht an diesen, dass die Gefälligkeit im Unterschied zur Bekanntheit und dem Markenimpact nicht steigt, wenn der Aufwand höher wird. Die Gefälligkeit bleibt entweder annähernd konstant oder sinkt sogar bei höheren Quartilen des Aufwandes. Nur TV-Werbungen (siehe Abbildung 3.22 und Tabelle 3.30) zeigen wachsende Gefälligkeitswerte, wenn der Aufwand steigt.
Die Länge eines Sujets hat sowohl auf die Bekanntheit als auch auf den Markenimpact und die Gefälligkeit in allen Werbeträgern und in allen Zielgruppen einen Einfluss. Ein länger dauerndes Sujet hat auch in der interessierenden Größe höhere Werte. Ausnahmen sind die Variablen *Bekanntheit* und *Markenimpact* bei den TV-Werbungen. Hier sinken die Werte bei höheren Längenkategorien nicht, sondern bleiben annähernd konstant.
Über die Zeit hinweg verändern sich die Werte. Diese Veränderung wird durch eine nichtlineare Funktion beschrieben. Wenn die abhängige Variable durch den Aufwand erklärt wird, dann ist diese Veränderung auch nichtlinear. Es gibt aber auch einige Ausnahmen. Die Bekanntheit verändert sich in der Zielgruppe „weiblich“, “30 bis 49“ und „älter als 50“ in den Werbeträgern „Hörfunk“ und „Infoscreen“ nur linear, bei Infoscreen-Sujets außerdem in der jüngsten Altersgruppe. Im Werbeträger „TV“ ist dies in der Gruppe „älter als 50“ der Fall. Bei den Print-Sujets steigt der Markenimpact mit steigendem Aufwand ebenfalls nur linear. Dies ist eindeutig in Abbildung 3.17(d) zu sehen.

Der Verlauf über die Zeit ist weder eindeutig positiv noch negativ. Er ist meist S-förmig, manchmal sogar doppelt S-förmig. Bei steigendem Aufwand jedoch steigt der Wert der abhängigen Variablen, es gibt jedoch einige wenige Ausnahmen. Dies ist bei Hörfunk-Spots, die kürzer als 15 Sekunden dauern, in der Gruppe der 14- bis 29-Jährigen bei der Variablen *Bekanntheit* der Fall. Im ersten Jahr bleibt die Bekanntheit bei steigendem Aufwand konstant, daraufhin sinkt sie kontinuierlich um insgesamt 20 %. Betrachtet man die Werte in dieser Zielgruppe genauer, so sieht man, dass die Bekanntheitswerte in dieser Gruppe unter jenen der Gesamtgruppe liegen. Der Mittelwert der Differenzen der beiden Variablen liegt bei -0,2696 und der Median bei -1, wobei die Bekanntheit der gesamten Gruppe von jener der 14- bis 29-Jährigen abgezogen wird. Die Gesamtgruppe hat also im Mittel höhere Werte. Von den

Tab. 4.46: Zusammenfassung der Ergebnisse der Likelihood-Ratio-Tests („s“ bedeutet signifikant, „ns“ bedeutet nicht signifikant, „-“ bedeutet nicht verfügbar)

	Bekanntheit				*Markenimpact*				*Gefälligkeit*			
	HF	IS	PR	TV	HF	IS	PR	TV	HF	IS	PR	TV
Aufwand	s	s	s	s	s	s	s	s	ns	ns	ns	s
Aufwand NL	s	s	s	s	s	s	s	s	ns	ns	ns	s
Zeit	s	s	s	s	s	s	s	s	s	s	s	s
Zeit NL	s	s	s	s	s	s	s	s	s	s	s	s
Länge	s	s	-	s	s	s	-	s	s	s	-	s
Aufwand × Zeit	s	s	s	s	s	s	s	ns	ns	ns	ns	s
Aufwand × Länge	s	ns	-	s	s	s	-	s	ns	ns	-	s
Zeit × Länge	s	ns	-	s	s	s	-	s	ns	ns	-	s
TOTAL NL	s	s	s	s	s	s	s	s	s	s	s	s
TOTAL WW	s	s	s	s	s	s	s	s	ns	ns	ns	s
TOTAL NL+WW	s	s	s	s	s	s	s	s	s	s	s	s
Regression	s	s	s	s	s	s	s	s	s	s	s	s

101 Beobachtungen in der Längenkategorie „≤ 15“ haben nur 27 Sujets einen Aufwand über 50.000 €, überhaupt nur vier Sujets weisen einen Aufwand höher als 75.000 € auf. Betrachtet man nur diese vier Sujets, so ist ersichtlich, dass zumindest drei dieser vier Sujets in der Gruppe der 14- bis 29-Jährigen weniger bekannt sind als in den übrigen Zielgruppen bzw. in der Gesamtgruppe (siehe Abbildung 4.32). Diese drei Werbungen sind für die Zielgruppe nur bedingt relevant. Es handelt sich hierbei um eine Katzenfutter-Werbung, die gesamt gesehen 51,26 %, von den 14- bis 29-Jährigen allerdings nur 42,88 % kennen. Außerdem betrifft dies eine Kredit-Werbung, bei der die Bekanntheit gesamt gesehen 48,00 %, in der jugendlichen Zielgruppe aber nur 41,67 % beträgt. Personen unter 20 Jahren haben wegen ihrer finanziellen Lage kaum Interesse an Krediten und auch keine Absicht, Werbungen dieser Art bewusst wahrzunehmen. Für die etwas Älteren in dieser Zielgruppe hat dieses Thema eventuell mehr Gewicht, doch auch hier wird es sicherlich wenige geben, die einen Kredit aufzunehmen bereit sind. Der dritte Spot ist eine Werbung aus dem Bereich „KFZ“, bei der sich die Bekanntheit um 14,88 % unterscheidet. Sie beläuft sich bei der Gesamtgruppe auf 40,33 %, in der Gruppe der 14- bis 29-Jährigen auf 25,45 %. Die beworbene Marke ist eher als Familienauto bekannt, darauf lässt auch die hohe Bekanntheit von 49,05 % in der Gruppe der 30- bis 49-Jährigen schließen. Nur die Bekanntheit jenes Sujets mit dem größten Aufwand, nämlich 119.736 €, ist in der Gruppe der 14- bis 29-Jährigen

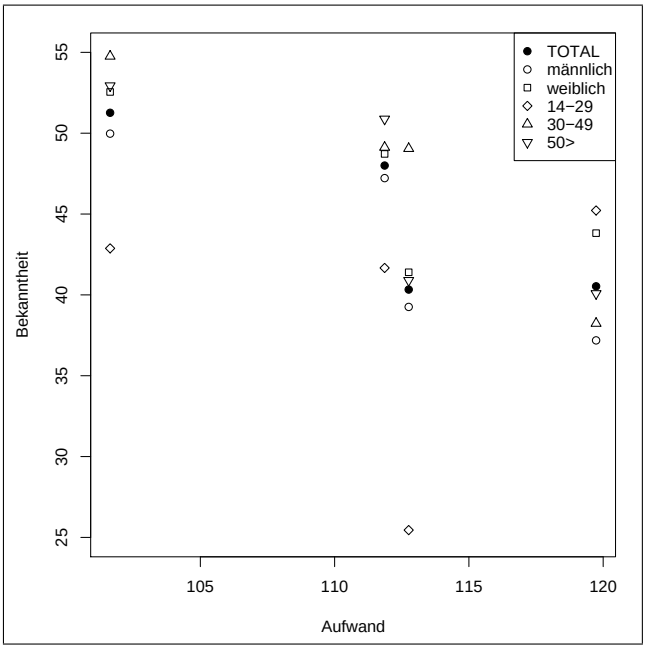


Abb. 4.32: Vergleich der Bekanntheitswerte der einzelnen Zielgruppen der vier höchstens 15 Sekunden dauernden Hörfunk-Spots mit einem Aufwand über 75.000 €

um 4,69 % höher. Diese Marke ist unter anderem wegen ihrer Präsenz in der Formel 1 auch bei den jüngeren Befragten bekannt.
Diese geringen Werte und die Tatsache, dass es im höheren Bereich des Aufwands wenige Beobachtungen gibt, führt zur negativen Veränderung bei steigendem Aufwand in dieser Zielgruppe und Längenkategorie. Auch wenn man weniger Stützstellen für die restringierten kubischen Splines festlegt bzw. den Aufwand nur linear oder mit linearen Splines modelliert, ändert sich an der Struktur wenig. Man sollte deswegen angesichts der vorher angestellten Überlegungen nicht darauf schließen, dass teurere Spots bei diesen Personen weniger bekannt sind als billigere.

Bei den Infoscreen-Werbungen fällt der Wert des Markenimpacts bei den über 50-Jährigen ab einem Aufwand von 20.000 € leicht ab. In den ersten Jahren sinkt der Markenimpact um insgesamt 4 %, in den letzten Jahren um insgesamt 5 %. Diese Verminderung des Wiedererkennunsgwertes scheint nicht signifikant zu sein. Bei näherer Betrachtung der Werte sieht man, dass die Werte der Zielgruppe der über 50-Jährigen entweder niedriger oder nur sehr gering höher sind. Dieser Trend wird aber eher auf Zufallsschwankungen zurückzuführen sein.

Auch der Markenimpact hat in der Zielgruppe der über 50-Jährigen bei den TV-Sujets in allen Längenkategorien bis zu einem Aufwand von 100.000 € fallende Werte. Bis zu diesem Wert ist die Korrelation zwischen dem Markenimpact und dem Aufwand negativ (-0,0444, P-Wert: 0,04097). Dies bedeutet, dass der Markenimpact sinkt, wenn der Aufwand höher wird. Bei Sujets, die einen Aufwand über 100.000 € aufweisen, liegt die Korrelation bei 0,1685 (P-Wert: $< 2{,}2 \cdot 10^{-6}$). Für alle Daten betrachtet ist die Korrelation signifikant positiv und liegt bei 0,1689. Deshalb verschwindet der negative Teil der Funktion auch, wenn man den Aufwand nur linear ins Modell nimmt. Auch wenn die Anzahl der Stützstellen von 5 auf 3 reduziert wird, steigt der Markenimpact selbst für geringe Aufwandswerte. In diesen Modellen werden die Stützstellen als fixe Quantile gewählt, weshalb bei fünf Stützstellen diese bei den Werten 30.367, 86.029, 133.107, 197.915 und 387.327 liegen. Bis zum zweiten Wert sinken die Werte des Markenimpacts, danach steigen sie. Wählt man nur drei Stützstellen, so befinden sich diese bei den Werten 46.201, 133.107 und 307.854. Der zweite Punkt, an dem die Funktion ihr Verhalten ändern kann, befindet sich also erst nach dem Wert 100.000. Die Funktion mit 3 Stützstellen ist jener, die den Aufwand nur linear modelliert, sehr ähnlich.

Die Variable *Gefälligkeit* steigt bei den TV-Spots der Länge „$\leq$ 15“ bis zu einem Aufwand von etwa 200.000 € und fällt dann für alle Zielgruppen stetig. Abnehmende Werte haben ab dem Jahr 2003 alle Zielgruppen auch in den restlichen Längenkategorien, am deutlichsten sieht man dies bei Spots der Dauer „15 bis 20“ und „20 bis 25“. Nur in der Tandem-Kategorie und für die Gruppe der 50-Jährigen bleiben die Werte annähernd gleich. Dies könnte am negativen Zusammenhang zwischen Aufwand und der Beliebtheit der Marken liegen. Bei jüngeren Marken oder Marken, die sich in der Bevölkerung noch nicht durchgesetzt haben, wird oft eine neue und aufwändigere Kampagne gestartet, die Werbungen werden öfter gesendet, was natürlich auch zu höheren Ausgaben führt.[12]
Betrachtet man die Daten, so ist auffällig, dass es in der Längenkategorie „$\leq$ 15“ nur 37 Sujets gibt, die einen Aufwand höher als 200.000 € haben, dies sind lediglich 0,04 %. Man kann in dem Bereich, in dem der Markenimpact bei steigendem Aufwand sinkt, also sehr schwer den Zusammenhang zwischen Aufwand, Zeit und Länge schätzen, da es wenige Beobachtungen gibt. Gleiches gilt auch für die Längenkategorie „15 bis 20“, in der bloß bei 215 Sujets, das heißt also bei knappen 13 %, mehr als 200.000 € aufgewendet wurden. Bei den längeren Spots gibt es in dem Bereich mehr Beobachtungen

[12] mündliche Information aus einem Gespräch mit Personen aus der Werbebranche

Tab. 4.47: Anzahl und Anteil der Sujets in den Längenkategorien

	$\geq$ *100.000*€		$\geq$ *200.000*€		$\geq$ *400.000*€		$\geq$ *600.000*€	
$\leq$*15*	294	(0,339 %)	37	(0,043 %)	1	(0,001 %)	0	(0,000 %)
15-20	1.008	(0,592 %)	215	(0,126 %)	15	(0,009 %)	4	(0,002 %)
20-25	750	(0,723 %)	322	(0,310 %)	44	(0,042 %)	3	(0,003 %)
25-30	1.506	(0,775 %)	750	(0,386 %)	109	(0,056 %)	26	(0,013 %)
$>$*30*	352	(0,800 %)	229	(0,520 %)	71	(0,161 %)	21	(0,048 %)
T	200	(0,823 %)	129	(0,531 %)	38	(0,156 %)	14	(0,058 %)

(siehe Tabelle 4.47). Der Markenimpact sinkt dann auch nur ab dem Jahr 2003. Die Verminderung in der Längenkategorie „20 bis 25“ beträgt höchstens 0,5, jene der Kategorie „25 bis 30“ zwischen 0,1 und 0,3 Bewertungspunkte. In diesem Bereich gibt es auch keine auffällig niedrigen Beobachtungen, sie sind alle eher im Mittelfeld angesiedelt.
Außerdem steigt die Gefälligkeit stets, wenn der Aufwand im Modell nur linear ist. Ähnlich jedoch ist die Art des Zusammenspiels der drei Variablen *Aufwand*, *Zeit* und *Länge*, wenn man anstatt kubischer Splines lineare Splines mit denselben drei, vier oder fünf Stützstellen anpasst.
Modelliert man den Aufwand mit nur einem Knoten beim Wert 100.000€, so steigt bis zu dieser Stützstelle die Bewertung und sinkt danach für die Spots, die kürzer als 15 Sekunden sind, und steigt danach für länger dauernde Werbungen, wobei die Steigerung mit der Länge zunimmt. Verschiebt man die Stützstelle zum Wert 200.000€, wächst die Bewertung bis zu dieser Stützstelle weiterhin, sinkt danach für die Spots der Kategorien „$\leq$ 15“ und „15 bis 20“ und ist konstant für die restlichen Kategorien. Wählt man den Knoten beim Wert 400.000€, so fällt die Gefälligkeit danach für alle Zielgruppen und Längenkategorien. Gleiches passiert, wenn die Stützstelle den Wert 600.000€ hat. Es ist hier also Folgendes zu beobachten: Je höher der Wert der Stützstelle ist, desto steiler ist der Abstieg der Funktion nach diesem Wert.
Man sieht bereits in Tabelle 4.47, dass es in allen Längenkategorien nur sehr wenige Werbungen gibt, die einen Aufwand von mehr als 400.000€ haben. In diesen Bereichen sollten deshalb keine Vorhersagen getroffen werden, denn man kann die immer schlechter werdenden Bewertungen der Sujets bei steigendem Aufwand nicht als aussagekräftig einstufen.

Im Werbeträger „Hörfunk“ sinkt der Wert des Markenimpacts bis zu einem Aufwand von 30.000€ in allen Längenkategorien und in allen Zielgruppen, daraufhin steigt er wieder oder bleibt zumindest konstant. Die Funktion hat

genau in diesem Bereich eine deutlich sichtbare wellenförmige Ausprägung. Betrachtet man die Korrelationen zwischen den Variablen *Aufwand* und *Markenimpact* getrennt für jene Beobachtungen, die einen Aufwand von höchstens 30.000 €, und für jene, die einen höheren Aufwand aufweisen, so ist die Korrelation für die billigeren Sujets immer negativ und für die teureren immer positiv. Dies erklärt die anfangs sinkenden Markenimpactwerte in allen Zielgruppen.
Die Wiedererkennungswerte sinken außerdem in der Gruppe der über 50-Jährigen in der Längenkategorie „$\leq$ 15“ um etwa 0,03 % und in der Kategorie „länger als 30“ um 0,02 %. Vergleicht man die Markenimpactwerte in dieser Gruppe mit jenen der restlichen Gruppen, so ist auffällig, dass diese geringfügig kleiner sind. Außerdem sollte man beachten, dass es nicht allzu viele Werbungen gibt, die in diesen beiden Längenkategorien liegen und einen hohen Aufwand haben.

Beim Vergleich der Werbeträger bezüglich der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* ist global keine Reihung der Mediengattungen möglich. TV-Werbungen sind am besten bekannt, haben aber in der Variablen *Markenimpact* Werte im unteren Bereich. Bezüglich der Variablen *Gefälligkeit* liegen sie an zweiter Stelle nach den Prospekt-Sujets. Diese wiederum zeichnen sich sowohl beim Markenimpact als auch bei der Gefälligkeit durch hohe Werte aus.
Die Werbeträger „Plakat“ und „Print“ zeigen in der Variablen *Markenimpact* keine signifikanten Unterschiede, in den Variablen *Gefälligkeit* und *Bekanntheit* allerdings sehr wohl. Plakat-Werbungen weisen bei diesen beiden höhere Werte auf. Print-Werbungen haben bezüglich Gefälligkeit die geringsten, bezüglich Bekanntheit die zweitkleinsten Werte. Allerdings ist der Unterschied zu den Online-Sujets, welche im Mittel die kleinsten Werte aufweisen, nicht signifikant. Online-Sujets haben bezüglich des Markenimpacts die zweithöchsten Werte, weisen allerdings eine relativ große Streuung auf. Dieser hohe Wert erscheint auf den ersten Blick relativ ungewöhnlich, kann aber vielleicht darauf zurückgeführt werden, dass dieser Werbeträger nur acht Monate lang abgefragt wurde. Ob sich dieser Impact-Wert auf so einem hohen Level halten kann, werden zukünftige Untersuchungen zeigen. Aus der Norm fällt der Wert auch deshalb, weil er in der Bekanntheit signifikant schlechtere Werte als die übrigen Werbeträger hat, beim Markenimpact aber, wie schon erwähnt, relativ hohe Werte aufweist. Das heißt, dass die Befragten durchschnittlich bedeutend mehr Sujets nach dem Interview nennen konnten, als sie am Anfang überhaupt kannten. Dieses Phänomen kann man auch beim Werbeträger „Prospekt“ beobachten, der Unterschied ist allerdings nicht so gravierend.

Hörfunk-Werbungen weisen in der Variablen *Markenimpact* signifikant niedrigere Werte als alle anderen Werbeträger auf. Die Bekanntheit der Hörfunk-Sujets ist im guten Mittelfeld angesiedelt. TV-Werbungen sind im Vergleich zu Hörfunk-Spots signifikant bekannter, weniger bekannt sind Online- und Print-Werbungen.

Im Effizienzvergleich der vier verschiedenen Werbeträger „Hörfunk“, „Infoscreen“, „Print“ und „TV“ hat der Print-Bereich die höchsten Effizienzwerte. Das bedeutet, für 50 % der Print-Werbungen hat ein Prozent Bekanntheit in etwa einen Wert von 4.460 €. Im Vergleich dazu werden beim Werbeträger „Infoscreen“ dafür nur zirka 400 € aufgewendet.

Im Sympathievergleich der Werbeträger „Hörfunk“, „Infoscreen“, „Online“, „Plakat“, „Print“ und „TV“ zeichnen sich TV-Werbungen durch hohe Sympathiewerte aus. Der Bereich Online-Werbung findet dabei die wenigsten Zustimmungen. Im Mittel haben nur 17,79 % der Befragten angegeben, dass sich die Online-Werbung besonders durch Sympathie auszeichnet, während der Mittelwert beim Werbeträger „TV“ bei 21,32 % liegt.

Die effizientesten Spots sind natürlich die kürzesten, da sie auch am wenigsten kosten. Auffällig ist jedoch, dass es zwischen den Kategorien „15 bis 20“ und „20 bis 25“ keinen signifikanten Unterschied gibt. Das heißt, ob eine Werbung nun 15 Sekunden dauert oder 25, ist von der Effizienz her egal. Der Median in der Längenkategorie „20 bis 25“ ist sogar geringfügig kleiner als jener der Kategorie „15 bis 20“ (2,057 zu 2,024, siehe Tabelle 3.44). Bei 50 % der Werbungen in der Kategorie „20 bis 25“ wird für einen Prozentpunkt Bekanntheit weniger als 2.024 € ausgegeben. Wenn man die Extrembeispiele heranzieht, so bedeutet dies, dass man für einen 15 Sekunden dauernden Spot gleich viel bezahlen muss wie für einen 25 Sekunden dauernden Spot, um einen Prozentpunkt Bekanntheit zu erlangen. Das bedeutet, dass ein Spot, der länger als 15 Sekunden dauert, auf 25 Sekunden ausgedehnt werden kann. Natürlich ist ein längerer Spot sowohl in der Produktion als auch bei den Werbeschaltungen teurer, er rentiert sich aber insofern, als man auch mehr Personen damit anspricht.

Diese Arbeit stellt auf keinen Fall eine komplette Bearbeitung des vorliegenden Datensatzes dar. Spannende Fragen für die Zukunft wären zum Beispiel, wie sich der Einfluss des Aufwandes weiterentwickelt oder ob der Aufwand auch in den Werbeträgern „Online“, „Plakat“ und „Prospekt“ einen Einfluss auf die Bekanntheit, die Gefälligkeit oder den Markenimpact hat. Letzteres konnte nicht beantwortet werden, da in diesen drei Werbeträgern der

Aufwand nicht angeführt wurde. Ob es noch weitere Einflussvariablen gibt, konnte wegen der vielen fehlenden Werte nicht geklärt werden.
Außerdem wäre es interessant, wie sich die Prospekte im Vergleich zu den restlichen Werbeträgern bezüglich der Image-Variablen verhalten. Solch ein Vergleich ist nicht möglich, da bei den Prospekten keine Image-Variablen erhoben werden bzw. bestehende Informationen, auch wenn sie ähnlich sind, nicht vergleichbar sind.

Abbildungsverzeichnis

2.1 Lineare Spline-Funktionen 23
2.2 Restringierte kubische Spline-Funktionen 25
2.3 Vergleich der Link-Funktionen des Logit-, Probit- und komplementären Log-Log-Modells 35
2.4 Vergleich der Potenztransformationen 42
3.1 Histogramme der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* mit Normalverteilungskurve 51
3.2 Histogramme der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* mit Normalverteilungskurve in den einzelnen Werbeträgern 54
3.3 Histogramm der Variablen *Aufwand*, getrennt nach Werbeträgern 55
3.4 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“ 57
3.5 Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“ 60
3.6 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“ 63
3.7 Einfluss von *Zeit* und *Länge* auf *Gefälligkeit* im Werbeträger „Hörfunk“ 64
3.8 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“ 67
3.9 Veränderung der Bekanntheit über die Zeit bei steigendem Aufwand und steigender Dauer des Infoscreen-Sujets 69
3.10 Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“ 70
3.11 Veränderung des Markenimpacts über die Zeit bei steigendem Aufwand und steigender Dauer des Infoscreen-Sujets 72
3.12 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“ 74
3.13 Veränderung der *Gefälligkeit* bezüglich der *Zeit*, getrennt nach Längenkategorien 75
3.14 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Quartilen im Werbeträger „Print“ 76
3.15 Veränderung der Bekanntheit über die Zeit bei steigendem Aufwand im Werbeträger „Print“, getrennt für die einzelnen Zielgruppen 78
3.16 Mittelwerte der Variablen *Markenimpact*, getrennt nach Quartilen im Werbeträger „Print“ 79

3.17 Veränderung des Markenimpacts über die Zeit bei steigendem Aufwand im Werbeträger „Print“, getrennt für die einzelnen Zielgruppen 81
3.18 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Quartile im Werbeträger „Print“ 82
3.19 Veränderung der Gefälligkeit von Print-Sujets über die Zeit in den verschiedenen Zielgruppen 83
3.20 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“ 84
3.21 Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“ 86
3.22 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“ 89
3.23 Boxplot der Variablen *Gefälligkeit*, getrennt nach Wirtschaftsbereichen 91
3.24 Boxplot der Variablen *Bekanntheit*, getrennt nach Werbeträgern 96
3.25 Boxplot der Variablen *Markenimpact*, getrennt nach Werbeträgern 98
3.26 Boxplot der Variablen *Gefälligkeit*, getrennt nach Werbeträgern 100
3.27 Histogramm der Variablen *Effizienz*, getrennt nach Werbeträgern 103
3.28 Boxplot der Variablen *Effizienz*, getrennt nach Werbeträgern . 104
3.29 Histogramm der Variablen *Sympathie* 105
3.30 Boxplot der Variablen *Sympathie*, getrennt nach Werbeträgern 107
3.31 Boxplot der Variablen *Effizienz*, getrennt nach Längenkategorien 109
4.32 Vergleich der Bekanntheitswerte der einzelnen Zielgruppen der vier höchstens 15 Sekunden dauernden Hörfunk-Spots mit einem Aufwand über 75.000 € 113

Tabellenverzeichnis

2.1 Kanonischer Link für einige Verteilungen 18
2.2 Empfohlene Quantile bei unterschiedlicher Anzahl an Stützstellen 25
2.3 ANOVA-Tabelle 39
3.1 Anzahl der Sujets in den Werbeträgern 47
3.2 Anzahl der Sujets in den Wirtschaftsbereichen 48
3.3 Deskriptive Maßzahlen der Variablen *Bekanntheit*, *Markenimpact* und *Gefälligkeit* 50
3.4 Anzahl und Anteil fehlender Werte der Variablen *Aufwand* und *Länge* 52
3.5 Deskriptive Maßzahlen der Variablen *Aufwand* in den Werbeträgern 55
3.6 Häufigkeiten der Variablen *Länge*, getrennt nach Werbeträgern 56
3.7 Bedeutung der Variablen *Zeit* 56
3.8 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“ 57
3.9 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Bekanntheit* im Werbeträger „Hörfunk“ 58
3.10 Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“ 60
3.11 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „Hörfunk“ 61
3.12 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Hörfunk“ 63
3.13 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Gefälligkeit* im Werbeträger „Hörfunk“ 64
3.14 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“ 67
3.15 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Bekanntheit* im Werbeträger „Infoscreen“ 68
3.16 Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“ 70
3.17 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „Infoscreen“ 71
3.18 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „Infoscreen“ 74
3.19 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Gefälligkeit* im Werbeträger „Infoscreen“ 74

3.20 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Quartilen im Werbeträger „Print“ 76
3.21 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Bekanntheit* im Werbeträger „Print“ 77
3.22 Mittelwerte der Variablen *Markenimpact*, getrennt nach Quartilen im Werbeträger „Print“ 79
3.23 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „Print“ 79
3.24 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Quartilen im Werbeträger „Print“ 82
3.25 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Gefälligkeit* im Werbeträger „Print“ 82
3.26 Mittelwerte der Variablen *Bekanntheit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“ 84
3.27 Wald- und Likelihood-Ratio-Teststatistiken für die Variable *Bekanntheit* im Werbeträger „TV“ 85
3.28 Mittelwerte der Variablen *Markenimpact*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“ 86
3.29 Wald- und Likelihood-Ratio-Teststatistiken für die abhängige Variable *Markenimpact* im Werbeträger „TV“ 87
3.30 Mittelwerte der Variablen *Gefälligkeit*, getrennt nach Faktorenlevel bzw. Quartilen im Werbeträger „TV“ 89
3.31 Wald- und Likelihood-Ratio-Teststatistiken für die Variable *Gefälligkeit* im Werbeträger „TV“ 89
3.32 Deskriptive Maßzahlen der Variablen *Gefälligkeit* in den einzelnen Wirtschaftsbereichen 92
3.33 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Gefälligkeit* in den einzelnen Wirtschaftsbereichen 93
3.34 Deskriptive Maßzahlen der Variablen *Bekanntheit* in den einzelnen Werbeträgern 97
3.35 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Bekanntheit* in den einzelnen Werbeträgern 98
3.36 Deskriptive Maßzahlen der Variablen *Markenimpact* in den einzelnen Werbeträgern 99
3.37 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Markenimpact* in den einzelnen Werbeträgern 100
3.38 Deskriptive Maßzahlen der Variablen *Gefälligkeit* in den einzelnen Werbeträgern 101
3.39 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Gefälligkeit* in den einzelnen Werbeträgern 102

3.40 Deskriptive Maßzahlen der Variablen *Effizienz*, getrennt nach Werbeträgern 103
3.41 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Effizienz* in den einzelnen Werbeträgern 104
3.42 Deskriptive Maßzahlen der Variablen *Sympathie* in den einzelnen Werbeträgern 106
3.43 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Sympathie* in den einzelnen Werbeträgern 106
3.44 Deskriptive Maßzahlen der Variablen *Effizienz*, getrennt nach Längenkategorien 108
3.45 Signifikante Ergebnisse der Post-Hoc-Tests für die Variable *Effizienz* in den Längenkategorien 109
4.46 Zusammenfassung der Ergebnisse der Likelihood-Ratio-Tests . 112
4.47 Anzahl und Anteil der Sujets mit einem Aufwand von mindestens 100.000 €, 200.000 €, 400.000 €bzw. 600.000 €in den Längenkategorien 115

Literatur

Backhaus, Klaus et al.: Multivariate Analysemethoden. Springer-Verlag, 1996, ISBN 3-540-60917-2

Birnbaum, Z. W.: Numerical Tabulation of the Distribution of Kolmogorov's Statistic for Finite Sample Size. Journal of the American Statistical Association, 47 1952 Nr. 259, 425–441

Bleymüller, Josef, Gehlert, Günther und Gülicher, Herbert: Statistik für Wirtschaftwissenschaftler. Verlag Franz Vahlen München, 2002, ISBN 3-80006-2854-6

Box, G. E. P.: Non-Normality and Tests on Variances. Biometrika, 40 1953, 318–335

Brown, Morton B. und Forsythe, Alan B.: Robust Tests for the Equality of Variances. Journal of the American Statistical Association, 69 1974 Nr. 346, 364–367

Christensen, Ronald: Plane Answers to Complex Questions. Springer-Verlag, 2002, ISBN 0-387-95361-2

Collet, D.: Modelling Binary Data. Chapman & Hall, 1991

Eckey, Hans Friedrich, Kosfeld, Reinhold und Dreger, Christian: Ökonometrie. Gabler Verlag, 2001

Fahrmeir, Ludwig und Tutz, Gerhard: Multivariate Statistical Modeling Based on Generlaized Linear Models. Springer-Verlag, 1994

Feller, W.: On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions. The Annals of Mathematical Statistics, 19 1948 Nr. 2, 177–189

Focus Institut Marketing Research Ges.m.b.H.: Sujet-Focus. Marculangasse 8, A-1220 Wien,, Ad-Bench Datenbank ⟨URL: `http://www.focusmr.com`⟩

Glass, Gene V.: Testing Homogeneity of Variances. American Educational Research Association, 3 1966 Nr. 3, 187–190

Greene, William H.: Econometric Analysis. Prentice Hall, 2000

Harrell, Frank E.: Regression Modeling Strategies. Springer-Verlag, 2001, ISBN 0-387-95232-2

Harrell, Frank E.: Design: Design Package. 2007, R package version 2.1-1

Heiler, Siegfried und Michels, Paul: Deskriptive und explorative Datenanalyse. R. Oldenbourg Verlag GmbH München, 1994, ISBN 3-486-22786-6

Johnston, John: Econometric Methods. McGraw-Hill, 1984

Kruskal, William H.: A Nonparametric Test for the Several Sample Problem. The Annals of Mathematical Statistics, 23 1952 Nr. 4, 525–540

Kruskal, William H. und Wallis, W. Allen: Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association, 47 1952 Nr. 260, 583–621

Lehmann, Erich Leo und D'Abrer, H. J. M.: Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, 1975

Levene, Howard: Robust Tests for Equality of Variances. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, 1960, 278–292

McCullagh, Peter und Nelder, John A.: Generalized Linear Models. CRC Press, 1989, ISBN 0-412-23850-0

Nelder, J. A. und Wedderburn, R. W. M.: Generalized Linear Models. Journal of the Royal Statistical Society. Series A (General), 135 1972 Nr. 3, 370–384, ISSN 00359238

R Development Core Team: R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2007 ⟨URL: `http://www.R-project.org`⟩, ISBN 3-900051-07-0

Rinne, Horst: Tachenbuch der Statistik. Verlag Harri Deutsch, 1995, ISBN 3-8171-1421-4

Schafer, Joseph: Analysis of incomplete multivariate data. CRC Press, 1997

Schlittgen, Rainer: Einführung in die Statistik. Oldenbourg Wissenschaftsverlag GmbH, 2003, ISBN 3-486-27446-5

Stone, C. J. und Koo, C. Y.: Additive Splines in Statistic. Proceedings of the Statistical Computations ASA, 1985, 45–48

Lebenslauf

Persönliche Daten

Name	Nina Huber, Bakk.
Geburtstag	1. August 1984
Staatsbürgerschaft	Österreich
Wohnhaft in	Wien
Familienstand	ledig

Ausbildung

1990 bis 1994	Volksschule Liezen
1994 bis 2002	BG/BRG Stainach
Juni 2002	Matura am BG/BRG Stainach
März 2003 bis Februar 2007	Bakkalaureatsstudium Statistik an der Universität Wien
Februar 2007	Abschluss (mit Auszeichnung) Bakkalaureatsstudium Statistik
seit März 2007	Magisterstudium Statistik an der Universität Wien

Berufserfahrung

WS 05/06	Tutorium „Lineare Modelle“
Februar 2006	Projekt „Entlassungsmanagement“ bei Univ.-Prof. Dr. Wilfried Grossmann
SS 06	Tutorium „Wahrscheinlichkeitsrechnung 2“
WS 06/07	Tutorium „Lineare Modelle“
SS 07	Tutorium „Inferenzstatistik“
WS 07/08	Tutorium „Lineare Modelle“
seit März 2009	Studienassistentin „Grundzüge der Wirtschaftsmathematik“