

Optimale Versuchsplanung für Zählraten mit zufälligen Blockeffekten

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

von Dipl.-Math. Marius Schmidt

geb. am 01.07.1987 in Wolfsburg

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Rainer Schwabe
Prof. Mag. Dr. Werner Müller

eingereicht am: 30.09.2019

Verteidigung am: 27.01.2020

Zusammenfassung

Das Poisson-Gamma-Modell ist eine Verallgemeinerung des Poisson-Modells, das zur Modellierung von Zähldaten verwendet werden kann. Solche Daten treten in Experimenten auf, bei denen die Anzahl von Objekten oder des Auftretens von interessierenden Ereignissen beobachtet wird. Das Poisson-Gamma-Modell ergibt sich aus dem Poisson-Modell, wenn wiederholt Beobachtungen für jede statistische Einheit durchgeführt werden und jeder statistischen Einheit ein Gamma-verteilter Blockeffekt zugewiesen wird. Die Schätzungen der unbekannt Parameter hängen von der Wahl der Versuchseinstellungen ab. Um die Parameter so präzise wie möglich zu schätzen und somit die Qualität der statistischen Analyse zu optimieren, werden optimale Designs bestimmt, welche die optimalen Werte und Häufigkeiten der Versuchseinstellungen angeben.

Die Fisher-Informationsmatrix für das Poisson-Gamma-Modell wird analytisch hergeleitet und in Abhängigkeit von der Fisher-Informationsmatrix für das Poisson-Modell dargestellt. Es werden optimale Designs für das Poisson-Gamma-Modell für verschiedene Optimalitätskriterien bestimmt.

Zunächst wird der Fall bekannter Parameter der Gamma-Verteilung betrachtet. Es wird gezeigt, dass das D -Optimalitätskriterium äquivalent zu einem kombinierten gewichteten Optimalitätskriterium aus D -Optimalität und D_s -Optimalität für die Effektparameter ist. Außerdem werden für das Poisson-Gamma-Modell die D -optimalen Designs für das multiple Regressionsmodell mit einer beliebigen Anzahl von Kovariablen bestimmt. Dabei ergeben sich die D_s -optimalen Designs für das Poisson- und Poisson-Gamma-Modell als Spezialfall. Für lineare Optimalitätskriterien wie L - und c -Optimalität wird gezeigt, dass die optimalen Designs im Poisson- und Poisson-Gamma-Modell übereinstimmen.

Da das Poisson-Gamma-Modell ein nichtlineares Modell ist, hängen die optimalen Designs von den Parametern ab. Um robuste Designs bezüglich Parametermisspezifikation zu erhalten, werden standardisiert Maximin D - und c -optimale Designs hergeleitet. Solche Designs maximieren die minimale Effizienz in einem vorgegebenen Parameterbereich. Sind die Parameter der Gamma-Verteilung bis auf ihr Verhältnis unbekannt, dann ist die Fisher-Informationsmatrix eine Blockdiagonalmatrix. Es wird gezeigt, dass die Resultate zu den optimalen Designs bei bekannten Parametern der Gamma-Verteilung auf diesen Fall übertragen werden können.

Abstract

The Poisson-Gamma model is a generalization of the Poisson model, which can be used for modelling count data. Such data arises in experiments, where the number of objects or occurrences of events of interest is observed. The Poisson-Gamma model results from the Poisson model when repeated observations are carried out for each statistical unit and each of them is assigned a Gamma distributed block effect.

The estimates of the unknown parameters depend on the choice of the covariates. In order to estimate the parameters as precisely as possible and thus to optimize the quality of the statistical analysis, optimal designs are determined, which specify the optimal values and frequencies of the experimental settings.

The Fisher information matrix for the Poisson-Gamma model is derived analytically and represented as a function of the Fisher information matrix for the Poisson model. Optimal designs for the Poisson-Gamma model are determined for different optimality criteria.

First, the case of known parameters of the Gamma distribution is considered. The D -optimality criterion is shown to be equivalent to a combined weighted optimality criterion of D -optimality and D_s -optimality for the effect parameters. Moreover, for the Poisson-Gamma model the D -optimal designs for multiple regression with an arbitrary number of covariates are determined and the D_s -optimal designs for the Poisson and Poisson-Gamma model are obtained as a special case. For linear optimality criteria such as L - and c -optimality, the optimal designs in the Poisson and Poisson-Gamma model are shown to coincide.

Since the Poisson-Gamma model is a nonlinear model, the optimal designs depend on the parameters. To obtain robust designs regarding parameter misspecification standardized maximin D - and c -optimal designs are derived. Such designs maximize the worst efficiency with respect to a prespecified parameter set.

If the parameters of the Gamma distribution are unknown except for their ratio, then the Fisher information matrix is a block diagonal matrix. It is shown that the results for the optimal designs with known parameters of the Gamma distribution can be transferred to this case.

Inhaltsverzeichnis

1	Einleitung	1
2	Poisson-Gamma-Modell	5
2.1	Modell	5
2.2	Dichte und Momente	6
2.3	Fisher-Information bei bekannten Parametern der Gamma-Verteilung . .	8
2.4	Fisher-Information bei unbekanntem Parametern der Gamma-Verteilung .	9
3	Optimale Versuchsplanung	13
3.1	Individuelle Designs und Populationsdesigns	13
3.2	Optimalitätskriterien	14
3.3	Optimale Gewichte	21
3.4	Äquivalenzsätze	22
4	Information und Optimalität für das Poisson-Gamma-Modell	25
5	Lokal optimale Designs	31
5.1	D -optimale Designs	31
5.1.1	Allgemeines Regressionsmodell	32
5.1.2	Regressionsmodell mit einer Kovariablen	35
5.1.3	Multiples Regressionsmodell	41
5.2	Lineare Optimalitätskriterien	50
6	Standardisiert Maximin optimale Designs	53
6.1	Standardisiert Maximin D -optimale Designs	54
6.1.1	Regressionsmodell mit einer binären Kovariablen	54
6.1.2	Regressionsmodell mit einer stetigen Kovariablen	59
6.1.3	Multiples Regressionsmodell	70
6.2	Standardisiert Maximin D_A -optimale Designs	73
6.3	Standardisiert Maximin c -optimale Designs	74
7	Optimale Designs bei unbekanntem Parametern der Gamma-Verteilung	77

8 Diskussion	81
A Anhang	85
A.1 Resultate zu Matrizen	85
A.2 Beweise zu Kapitel 2	86
A.3 Beweise und Hilfsresultat zu Kapitel 4	91
A.4 Beweise und Hilfsresultat zu Kapitel 5	94
A.5 Beweise und Hilfsresultate zu Kapitel 6	97
Abbildungsverzeichnis	111
Tabellenverzeichnis	113
Symbolverzeichnis	114
Literaturverzeichnis	116

1 Einleitung

Zählraten treten in Experimenten auf, bei denen die Anzahl von Objekten oder das Auftreten von interessierenden Ereignissen beobachtet wird. Solche Daten werden häufig mit dem Poisson-Modell beschrieben, bei dem der Erwartungswert der Poisson-verteilten Zielgröße durch eine Linkfunktion mit einem linearen Prädiktor, bestehend aus Kovariablen und unbekanntem Modellparametern, in Verbindung gesetzt wird. In diesen Experimenten können für jede statistische Einheit wiederholt Beobachtungen durchgeführt werden. Wird ein Gamma-verteilter zufälliger Blockeffekt für jede statistische Einheit angenommen, um die Abhängigkeit der Beobachtungen zu modellieren, ergibt sich das Poisson-Gamma-Modell als Verallgemeinerung des Poisson-Modells.

Das Poisson-Gamma-Modell wurde bereits von Arbous und Kerrich (1951) und Bates und Neyman (1952) im Rahmen der Untersuchung der Unfallneigung von Personen betrachtet. Weiterhin wurde es von Waller und Zelterman (1997) und von Solis-Trapala und Farewell (2005) zur Analyse von medizinischen Daten verwendet.

Das Poisson-Gamma-Modell gehört zur Klasse der Verallgemeinerten Linearen Gemischten Modelle, mit denen Experimente mit nichtlinearen Wirkungszusammenhängen beschrieben werden können. Dabei wird unter Verwendung von unbekanntem festen und zufälligen Effekten der Einfluss von Kovariablen auf die Zielgröße des Experimentes modelliert. Es lassen sich deutlich komplexere Experimente modellieren als bei gewöhnlichen Modellen ohne zufällige Effekte, beispielsweise wenn die Beobachtungen nicht unabhängig untereinander sind. Die Analyse solcher Modelle ist jedoch aufwendiger als bei den Modellen ohne zufällige Effekte. Im Allgemeinen kann die Dichte der Zufallsvariablen, welche die Zielgröße des Experimentes modellieren, und somit die Fisher-Informationsmatrix bei Verallgemeinerten Linearen Gemischten Modellen nicht analytisch berechnet werden und muss daher approximiert werden (vgl. Breslow und Clayton (1993), Longford (1994)). Dies ist insbesondere bei der Annahme normalverteilter zufälliger Effekte der Fall.

Das Poisson-Gamma-Modell basiert auf der Poisson- und Gamma-Verteilung, bei denen es sich um konjugierte Verteilungen handelt. Die Dichte und die Fisher-Informationsmatrix lassen sich analytisch berechnen, was die analytische Untersuchung dieses Modells ermöglicht.

Die Schätzungen der unbekanntem Modellparameter hängen von der Wahl der Versuchseinstellungen ab. Die Versuchseinstellungen und die Häufigkeiten ihrer Verwendung werden in einem Versuchsplan zusammengefasst. Die Güte der Versuchspläne wird durch die Verwendung von Optimalitätskriterien beurteilt, welche Funktionen von der Informationsmatrix sind. Um möglichst präzise Parameterschätzungen zu erhalten, sind die Versuchspläne, welche auch Designs genannt werden, bezüglich dieser Kriterien zu optimieren. Mit solchen optimalen Designs kann die Anzahl der experimentellen Einheiten reduziert werden, was zu einer Senkung der experimentellen Kosten führt. Darüber hinaus kann beispielsweise bei Tierversuchen aus ethischen Gründen die Verwendung von optimalen Designs erforderlich sein. Das Ziel ist also die effiziente Planung von Experimenten, um die unbekanntem Parameter bestmöglich schätzen zu können und somit die Qualität der statistischen Analyse zu optimieren.

Für das Poisson-Modell gibt es viele wissenschaftliche Arbeiten zu optimalen Designs. Es wurden von Ford et al. (1992) und Rodríguez-Torreblanca und Rodríguez-Díaz (2007) D - und c -optimale Designs für den Fall einer Kovariablen bestimmt. Dette et al. (2006) betrachteten c - und E -optimale Designs für eine Kovariable. Wang et al. (2006) führten numerische Untersuchungen für zwei Kovariablen mit und ohne einen zusätzlichen Interaktionsterm durch. Für den Fall der multiplen Regression mit einer beliebigen Anzahl von Kovariablen bestimmten Russell et al. (2009) D -optimale Designs und Schmidt (2019) leitete c -, L - und ϕ_p -optimale Designs her. Für multiple binäre Kovariablen berechneten Graßhoff et al. (2013, 2015, 2018) D -optimale Designs. Außerdem bestimmten Konstantinou et al. (2014) standardisiert Maximin c - und D -optimale Designs für den Fall einer Kovariablen.

Im Rahmen der Intelligenztestung betrachteten Graßhoff et al. (2016, 2018) das Poisson-Gamma-Modell bzw. verallgemeinerte Negativ-Binomial-Modell mit einer Beobachtung pro statistischer Einheit und leiteten Resultate zu D -optimalen Designs für die multiple Regression mit einer beliebigen Anzahl von binären Kovariablen her. Rodríguez-Torreblanca und Rodríguez-Díaz (2007) bestimmten für das Negativ-Binomial-Modell D - und c -optimale Designs für eine Kovariable. Schmidt und Schwabe (2017) berechneten D -optimale Designs und Schmidt (2019) bestimmte c -, L - und ϕ_p -optimale Designs für eine beliebige Anzahl von Kovariablen für das Negativ-Binomial-Modell mit unabhängigen Beobachtungen.

Im Fall mehrerer Beobachtungen für jede statistische Einheit und somit abhängigen Beobachtungen existieren nur wenige analytische Resultate zu optimalen Designs für nichtlineare gemischte Modelle. Für das Poisson-Modell mit normalverteilten zufälligen Effekten wurden von Ogungbenro und Aarons (2011) numerisch D -optimale Designs untersucht und von Niaparast (2009) und Niaparast und Schwabe (2013) D -optimale

Designs unter Verwendung der Quasi-Informationsmatrix für den Fall einer Kovariablen berechnet.

Die Arbeit gliedert sich in acht Kapitel und einen Anhang. In Kapitel 2 wird das Poisson-Gamma-Modell eingeführt und sowohl die Dichte als auch die Fisher-Informationsmatrix für den Fall bekannter und unbekannter Parameter der Gamma-Verteilung hergeleitet. Kapitel 3 gibt eine Einführung in die Theorie der optimalen Versuchsplanung. Es werden individuelle Designs sowie Populationsdesigns definiert und verschiedene Optimalitätskriterien eingeführt. Weiterhin werden Äquivalenzsätze zur Bestimmung optimaler Designs angegeben.

In den folgenden drei Kapiteln wird angenommen, dass die Parameter der Gamma-Verteilung bekannt seien. In Kapitel 4 werden Zusammenhänge zwischen den Informationsmatrizen für das Poisson- und Poisson-Gamma-Modell bestimmt. Weiterhin wird die Berechnung optimaler Populationsdesigns auf die Bestimmung optimaler individueller Designs zurückgeführt.

Kapitel 5 beschäftigt sich mit der Herleitung lokal optimaler Designs. Zunächst wird das D -Optimalitätskriterium betrachtet und gezeigt, dass ein Zusammenhang zum D - und D_s -Optimalitätskriterium für das Poisson-Modell besteht. Anschließend werden für das Regressionsmodell mit einer Kovariablen und das multiple Regressionsmodell mit einer beliebigen Anzahl von Kovariablen D -optimale Designs hergeleitet. Als Spezialfall ergeben sich die D_s -optimalen Designs für das Poisson- und Poisson-Gamma-Modell. Außerdem erfolgt die Betrachtung linearer Optimalitätskriterien.

Um die Parameterabhängigkeit der lokal optimalen Designs zu reduzieren, werden in Kapitel 6 standardisierte Maximin Kriterien für D - und c -Optimalität betrachtet. Dabei werden für die Berechnung optimaler Designs analytische und numerische Methoden verwendet.

In Kapitel 7 erfolgt die Berechnung optimaler Designs unter der Annahme unbekannter Parameter der Gamma-Verteilung. Es folgt eine Diskussion der erzielten Ergebnisse in Kapitel 8. Die Arbeit endet mit einem Anhang, in dem sich einige technische Hilfsresultate und ein Teil der Beweise befinden.

2 Poisson-Gamma-Modell

In diesem Kapitel wird das Poisson-Gamma-Modell betrachtet, welches eine Verallgemeinerung des Poisson-Modells ist. Zunächst erfolgt die Beschreibung des Modells und die analytische Berechnung der Dichte. Darauf aufbauend wird die Fisher-Informationsmatrix für das Poisson-Gamma-Modell für die Fälle bekannter und unbekannter Parameter der Gamma-Verteilung hergeleitet.

2.1 Modell

Es werden n statistische Einheiten betrachtet, welche beispielsweise Gruppen oder Individuen sein können, für die m Experimente mit Zielgrößen Y_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m$, durchgeführt werden. Jeder statistischen Einheit wird ein Gamma-verteilter zufälliger Blockeffekt $\Theta_i \sim \gamma(a, b)$ mit Formparameter $a > 0$ und inversem Skalenparameter $b > 0$ zugewiesen, wobei die Blockeffekte untereinander unabhängig sind. Die Dichte der Gamma-Verteilung $\gamma(a, b)$ ist durch

$$f_{\gamma(a,b)}(\theta) = \frac{b^a}{\Gamma(a)} \cdot \theta^{a-1} \cdot e^{-b\theta} \quad (2.1)$$

für $\theta > 0$ gegeben, wobei $\Gamma(a)$ die Gamma-Funktion bezeichnet, welche die Gleichung $\Gamma(a+1) = a \cdot \Gamma(a)$ erfüllt. Weiterhin gilt $E_{a,b}(\Theta_i) = a/b$ und $\text{Var}_{a,b}(\Theta_i) = a/b^2$.

Ist $\Theta_i = \theta_i$ gegeben, so wird angenommen, dass die Zufallsvariablen Y_{ij} unabhängig Poisson-verteilt mit einem von θ_i abhängigen Parameter λ_{ij} sind. Die Poisson-Verteilung $\text{Po}(\lambda)$ mit Parameter λ besitzt die diskrete Dichte

$$f_{\text{Po}(\lambda)}(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad (2.2)$$

für $k \in \mathbb{N}_0$. Der Erwartungswert der Poisson-Verteilung λ_{ij} wird unter Verwendung der kanonischen Linkfunktion, also einer Logarithmus-Transformation, mit dem linearen Prädiktor, welcher aus einem Term mit festen Effekten und einem additiven zufälligen

Effekt $v_i = \ln(\theta_i)$ besteht, in Zusammenhang gesetzt:

$$\ln(\lambda_{ij}) = \mathbf{f}(\mathbf{x}_{ij})^T \boldsymbol{\beta} + v_i. \quad (2.3)$$

Es folgt:

$$\lambda_{ij} = \theta_i \cdot \exp(\mathbf{f}(\mathbf{x}_{ij})^T \boldsymbol{\beta}). \quad (2.4)$$

Dabei ist $\mathbf{x}_{ij} \in \mathbb{R}^k$ der Kovariablenvektor. Der Vektor $\mathbf{f} = (1, f_1, \dots, f_{p-1})^T$ besteht aus bekannten Regressionsfunktionen und der Vektor $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{p-1})^T$ ist der unbekannte Parametervektor.

2.2 Dichte und Momente

Im Folgenden wird eine beliebige statistische Einheit i betrachtet und zur Vereinfachung der Notation der Gruppenindex i weggelassen. Um die Fisher-Informationsmatrix zu berechnen, wird die Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$ benötigt, welche sich aus dem Integral

$$f_{\mathbf{Y}}(\mathbf{y}) = \int_{-\infty}^{\infty} f_{\mathbf{Y}|\Theta=\theta}(\mathbf{y}) \cdot f_{\Theta}(\theta) \, d\theta \quad (2.5)$$

ergibt, bei dem über den zufälligen Effekt integriert wird. Dabei ist $f_{\mathbf{Y}|\Theta=\theta}(\mathbf{y})$ die bedingte Dichte von \mathbf{Y} gegeben $\Theta = \theta$ und $f_{\Theta}(\theta)$ ist die Dichte des zufälligen Effekts Θ . Dieses Integral ist im Allgemeinen für beliebige Verteilungen nicht analytisch berechenbar. Bei der Poisson- und Gamma-Verteilung handelt es sich um konjugierte Verteilungen und das Integral und somit die Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$ lassen sich analytisch herleiten.

Satz 2.1

Die Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$ ist gegeben durch:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma(a + \sum_{j=1}^m y_j)}{\Gamma(a) \cdot \prod_{j=1}^m y_j!} \cdot \frac{e^{\sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j}}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^{\sum_{j=1}^m y_j}} \cdot \left(\frac{b}{b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}\right)^a. \quad (2.6)$$

Beweis:

Es sei $\mathbf{y} = (y_1, \dots, y_m)$. Integration über den zufälligen Effekt liefert:

$$f_{\mathbf{Y}}(\mathbf{y}) = \int_0^{\infty} f_{\mathbf{Y}|\Theta=\theta}(\mathbf{y}) \cdot f_{\Theta}(\theta) \, d\theta$$

$$\begin{aligned}
&= \int_0^\infty \prod_{j=1}^m \left(\theta^{y_j} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j}}{y_j!} \cdot e^{-\theta \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \right) \cdot \frac{b^a}{\Gamma(a)} \cdot \theta^{a-1} \cdot e^{-b \cdot \theta} d\theta \\
&= \int_0^\infty \theta^{\sum_{j=1}^m y_j} \cdot \frac{e^{\sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j}}{\prod_{j=1}^m y_j!} \cdot e^{-\theta \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \cdot \frac{b^a}{\Gamma(a)} \cdot \theta^{a-1} \cdot e^{-b \cdot \theta} d\theta.
\end{aligned}$$

Mit $\tilde{a} = a + \sum_{j=1}^m y_j$ und $\tilde{b} = b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}$ ergibt sich

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{b^a \cdot e^{\sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j} \cdot \Gamma(\tilde{a})}{\Gamma(a) \cdot \prod_{j=1}^m y_j! \cdot \tilde{b}^{\tilde{a}}} \cdot \int_0^\infty \frac{\tilde{b}^{\tilde{a}}}{\Gamma(\tilde{a})} \cdot \theta^{\tilde{a}-1} \cdot e^{-\tilde{b} \cdot \theta} d\theta = \frac{b^a \cdot e^{\sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j} \cdot \Gamma(\tilde{a})}{\Gamma(a) \cdot \prod_{j=1}^m y_j! \cdot \tilde{b}^{\tilde{a}}}$$

als die gemeinsame Dichte von Y_1, \dots, Y_m . \square

Die in Satz 2.1 berechnete Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$ wurde bereits von Bates und Neyman (1952) hergeleitet, wobei der Erwartungswert der Poisson-Verteilung λ_j nicht als $\lambda_j = \theta \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta})$ modelliert wurde. Die Dichte wurde ebenfalls von Waller und Zelterman (1997) und Solis-Trapala und Farewell (2005) für den Fall gleicher Parameter $a = b$ der Gamma-Verteilung angegeben.

Die Randdichte einer Zufallsvariablen Y_j , $j = 1, \dots, m$, ergibt sich als Spezialfall von Satz 2.1 für $m = 1$ und ist durch

$$f_{Y_j}(y) = \frac{b^a \cdot \Gamma(a + y)}{\Gamma(a) \cdot y!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+y}} \quad (2.7)$$

gegeben. Die Zufallsvariablen Y_j sind Poisson-Gamma-verteilt, wobei diese Verteilung auch als verallgemeinerte Negativ-Binomialverteilung bezeichnet wird. Im folgenden Satz werden Erwartungswert und Varianz von Y_j angegeben (vgl. Molenberghs et al., 2010), wobei der Beweis sich im Anhang befindet.

Satz 2.2

Für den Erwartungswert und die Varianz von Y_j gilt:

$$E_{\boldsymbol{\beta}, a, b}(Y_j) = \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}, \quad (2.8)$$

$$\text{Var}_{\boldsymbol{\beta}, a, b}(Y_j) = \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \left(\frac{1}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} + 1 \right). \quad (2.9)$$

Im Gegensatz zur Poisson-Verteilung, bei der Erwartungswert und Varianz gleich sind, ist bei der Poisson-Gamma-Verteilung die Varianz größer als der Erwartungswert. Diese Eigenschaft wird Überdispersion genannt.

Da die Gamma-verteilten Blockeffekte Θ_i , $i = 1, \dots, n$, für die verschiedenen statistischen Einheiten unabhängig sind, sind die Poisson-Gamma-verteilten Zufallsvariablen Y_{i_1, j_1} und Y_{i_2, j_2} mit $i_1 \neq i_2$ für verschiedene statistische Einheiten ebenfalls unabhängig. Die Dichte von $\mathbf{Y}_{\text{Ges}} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ für n statistische Einheiten ist somit das Produkt der Dichten für eine statistische Einheit, welche in Satz 2.1 gegeben sind.

2.3 Fisher-Information bei bekannten Parametern der Gamma-Verteilung

Es sei $P = \{P_{\boldsymbol{\vartheta}}; \boldsymbol{\vartheta} \in \mathcal{T}\}$ eine Familie von Verteilungen mit Dichten $f_{\boldsymbol{\vartheta}}$ bezüglich eines σ -endlichen Maßes μ für alle $\boldsymbol{\vartheta} \in \mathcal{T}$. Dann ist

$$\mathbf{I}(\boldsymbol{\vartheta}) = E_{\boldsymbol{\vartheta}} \left(\left(\frac{\partial}{\partial \boldsymbol{\vartheta}} \ln(f_{\boldsymbol{\vartheta}}(\mathbf{Y})) \right) \left(\frac{\partial}{\partial \boldsymbol{\vartheta}^T} \ln(f_{\boldsymbol{\vartheta}}(\mathbf{Y})) \right) \right) \quad (2.10)$$

bei Existenz des Erwartungswertes die Fisher-Informationsmatrix.

Die Fisher-Informationsmatrix für das Poisson-Modell ist im Fall von m unabhängigen Beobachtungen durch

$$\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) = \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T \quad (2.11)$$

gegeben. Unter Verwendung von Satz 2.1 für die Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$ lässt sich die Fisher-Informationsmatrix für eine einzelne statistische Einheit für das Poisson-Gamma-Modell berechnen. Dabei bezeichnet $\mathbf{e}_{1,p} \in \mathbb{R}^p$ den ersten Standard-Einheitsvektor.

Satz 2.3

Die Fisher-Informationsmatrix für den Parametervektor $\boldsymbol{\beta}$ ist gegeben durch:

$$\begin{aligned} \mathbf{I}(\boldsymbol{\beta}) &= \frac{a}{b} \cdot \left(\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T - \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)^T}{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} + b} \right) \\ &= \frac{a}{b} \cdot \left(\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) - \frac{\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta})}{\mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} + b} \right). \end{aligned} \quad (2.12)$$

Der Beweis wird im Anhang angegeben. Satz 2.3 zeigt, dass sich die Fisher-Informationsmatrix für das Poisson-Gamma-Modell $\mathbf{I}(\boldsymbol{\beta})$ in Abhängigkeit von der Fisher-Informationsmatrix $\mathbf{I}_{\text{Po}}(\boldsymbol{\beta})$ für das Poisson-Modell darstellen lässt.

Bemerkung 2.4

Da die Beobachtungen zwischen den statistischen Einheiten unabhängig sind, berechnet sich die Fisher-Informationsmatrix $\mathbf{I}_{\text{Ges}}(\boldsymbol{\beta})$ für n unabhängige statistische Einheiten als Summe der Fisher-Informationsmatrizen $\mathbf{I}_i(\boldsymbol{\beta})$ für jede einzelne statistische Einheit i :

$$\mathbf{I}_{\text{Ges}}(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{I}_i(\boldsymbol{\beta}). \quad (2.13)$$

Falls bei jeder statistischen Einheit nur eine Beobachtung durchgeführt wird, also für $m = 1$, ergibt sich das verallgemeinerte Negativ-Binomial-Modell (vgl. Graßhoff et al., 2016, 2018), für das die Fisher-Informationsmatrix durch

$$\mathbf{I}_{\text{Ges}}(\boldsymbol{\beta}) = \sum_{i=1}^n \frac{a \cdot \exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta})}{\exp(\mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta}) + b} \cdot \mathbf{f}(\mathbf{x}_i) \mathbf{f}(\mathbf{x}_i)^T \quad (2.14)$$

gegeben ist.

Durch den zufälligen Blockeffekt $v_i = \ln(\theta_i)$ sind die Zufallsvariablen Y_1, \dots, Y_m für eine statistische Einheit nicht unabhängig. Daher lässt sich die Fisher-Informationsmatrix (2.12) für eine statistische Einheit nicht als Summe der Fisher-Informationsmatrizen für jeweils eine Beobachtung darstellen.

2.4 Fisher-Information bei unbekanntem Parametern der Gamma-Verteilung

In diesem Abschnitt wird der Fall betrachtet, dass die Parameter a und b der Gamma-Verteilung unbekannt sind. Die Beweise zu den in diesem Abschnitt angegebenen Sätzen befinden sich im Anhang.

Die Dichte für $\mathbf{Y} = (Y_1, \dots, Y_m)$ in Satz 2.1 lässt sich umformen zu:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma(a + \sum_{j=1}^m y_j)}{\Gamma(a) \cdot \prod_{j=1}^m y_j!} \cdot \frac{\prod_{j=1}^m \left(\frac{1}{b} \cdot e^{\sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^{y_j}}{\left(1 + \frac{1}{b} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^{\sum_{j=1}^m y_j}} \cdot \left(\frac{1}{1 + \frac{1}{b} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \right)^a.$$

Die Parameter β_0 und b treten immer als Produkt $(1/b) \cdot e^{\beta_0}$ auf. Um eine Überparametrisierung des Modells zu vermeiden, wird der Parameter $\tilde{\beta}_0 = -\ln(b) + \beta_0$ eingeführt. Dann ist der Parametervektor durch $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_0, \beta_1, \dots, \beta_{p-1})$ gegeben. Die Parameter β_0 und $\tilde{\beta}_0$ können beliebige Werte in den reellen Zahlen annehmen. Somit kann bei der Berechnung der Fisher-Informationsmatrix statt $\tilde{\boldsymbol{\beta}}$ äquivalent $\boldsymbol{\beta}$ mit bekanntem Para-

meter $b = 1$ verwendet werden. Da neben dem p -dimensionalen Parametervektor $\boldsymbol{\beta}$ auch der Formparameter a der Gamma-Verteilung als unbekannt vorausgesetzt wird, ist die Fisher-Informationsmatrix eine $(p + 1) \times (p + 1)$ -Matrix. Bei der Berechnung der Fisher-Informationsmatrix wird die Digamma-Funktion $\psi(z) = \frac{d}{dz} \ln(\Gamma(z))$ und die Trigamma-Funktion $\psi_1(z) = \frac{d^2}{dz^2} \ln(\Gamma(z))$ verwendet.

Satz 2.5

Es sei $b = 1$. Die Fisher-Informationsmatrix

$$\mathbf{I}(\boldsymbol{\beta}, a) = \begin{pmatrix} \mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\beta}} & \mathbf{I}_{\boldsymbol{\beta},a} \\ \mathbf{I}_{\boldsymbol{\beta},a}^T & I_{a,a} \end{pmatrix} \quad (2.15)$$

für die Parameter $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{p-1})^T$ und a besitzt die folgenden Einträge:

$$\mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\beta}} = a \cdot \left(\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) - \frac{\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta})}{\mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} + 1} \right), \quad (2.16)$$

$$\mathbf{I}_{\boldsymbol{\beta},a} = \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{1 + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}, \quad (2.17)$$

$$I_{a,a} = -E_{\boldsymbol{\beta},a} \left(\psi_1 \left(a + \sum_{j=1}^m Y_j \right) \right) + \psi_1(a). \quad (2.18)$$

Es wird nun der Fall betrachtet, dass die Parameter a und b der Gamma-Verteilung zwar unbekannt sind, aber $a/b = c$ mit bekanntem $c \in \mathbb{R}$ gilt. Das Verhältnis der beiden Parameter a und b ist somit fest vorgegeben und es folgt $E_{a,b}(\Theta_i) = a/b = c$ für den Erwartungswert des Gamma-verteilten Blockeffekts Θ_i . Die Dichte für $\mathbf{Y} = (Y_1, \dots, Y_m)$ in Satz 2.1 ergibt sich für $b = a/c$ zu:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma(a + \sum_{j=1}^m y_j)}{\Gamma(a) \cdot \prod_{j=1}^m y_j!} \cdot \frac{e^{\sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j}}{\left(\frac{a}{c} + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^{\sum_{j=1}^m y_j}} \cdot \left(\frac{\frac{a}{c}}{\frac{a}{c} + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \right)^a. \quad (2.19)$$

Unter der Annahme $a/b = c$ hängt somit die Fisher-Informationsmatrix nur von den Parametern $\boldsymbol{\beta}$ und a ab und ist eine $(p + 1) \times (p + 1)$ -Matrix. Desweiteren hat diese Annahme auch Auswirkungen auf die Komponenten der Fisher-Informationsmatrix, insbesondere auf die Blöcke $\mathbf{I}_{\boldsymbol{\beta},a}$ und $I_{a,a}$, wie der folgende Satz zeigt. Dabei ist $\mathbf{0}_p$ der p -dimensionale Nullvektor.

Satz 2.6

Es sei $a/b = c$ mit bekanntem $c \in \mathbb{R}$. Die Fisher-Informationsmatrix

$$\mathbf{I}(\boldsymbol{\beta}, a) = \begin{pmatrix} \mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\beta}} & \mathbf{I}_{\boldsymbol{\beta},a} \\ \mathbf{I}_{\boldsymbol{\beta},a}^T & I_{a,a} \end{pmatrix} \quad (2.20)$$

für die Parameter $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{p-1})^T$ und a besitzt die folgenden Einträge:

$$\mathbf{I}_{\boldsymbol{\beta},\boldsymbol{\beta}} = c \cdot \left(\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) - \frac{\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta})}{\mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} + \frac{a}{c}} \right), \quad (2.21)$$

$$\mathbf{I}_{\boldsymbol{\beta},a} = \mathbf{0}_p, \quad (2.22)$$

$$I_{a,a} = -E_{\boldsymbol{\beta},a} \left(\psi_1 \left(a + \sum_{j=1}^m Y_j \right) \right) + \psi_1(a) - \frac{c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{a \cdot \left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)}. \quad (2.23)$$

Im Gegensatz zum Fall beliebiger Parameter a und b ist die Fisher-Informationsmatrix für $a/b = c$ eine Blockdiagonalmatrix. Für den Fall $a = b$ wurde die Blockdiagonalstruktur der Informationsmatrix auch von Waller und Zeltermann (1997) gezeigt. Diese Struktur vereinfacht die Berechnung optimaler Designs, wie in Kapitel 7 gezeigt wird.

Der Eintrag $I_{a,a}$ der Fisher-informationsmatrix $\mathbf{I}(\boldsymbol{\beta}, a)$ hängt sowohl für $b = 1$ als auch für $a/b = c$ von $E_{\boldsymbol{\beta},a}(\psi_1(a + \sum_{j=1}^m Y_j))$ ab. Um diesen Erwartungswert zu berechnen, ist der folgende Satz hilfreich, welcher die Dichte der Summe der Poisson-Gamma-verteilten Zufallsvariablen Y_1, \dots, Y_m angibt.

Satz 2.7

Für die Dichte von $Z = \sum_{j=1}^m Y_j$ gilt:

$$f_Z(z) = \frac{b^a \cdot \Gamma(a+z)}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^a \cdot \Gamma(a) \cdot z!} \cdot \left(\frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \right)^z. \quad (2.24)$$

Mit der in Satz 2.7 gegebenen Dichte von $Z = \sum_{j=1}^m Y_j$ folgt

$$\begin{aligned} E_{\boldsymbol{\beta},a} \left(\psi_1 \left(a + \sum_{j=1}^m Y_j \right) \right) &= \sum_{z=0}^{\infty} \psi_1(a+z) \cdot f_Z(z) \\ &= \sum_{z=0}^{\infty} \frac{\psi_1(a+z) \cdot b^a \cdot \Gamma(a+z)}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^a \cdot \Gamma(a) \cdot z!} \cdot \left(\frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \right)^z, \end{aligned}$$

wobei je nach betrachtetem Fall entweder $b = 1$ oder $b = a/c$ gewählt wird.

Sowohl für beliebigen Parameter a und $b = 1$ als auch für $a/b = c$ ergibt sich wie im Fall bekannter Parameter der Gamma-Verteilung die Fisher-Informationsmatrix für n unabhängige statistische Einheiten als $\mathbf{I}_{\text{Ges}}(\boldsymbol{\beta}, a) = \sum_{i=1}^n \mathbf{I}_i(\boldsymbol{\beta}, a)$. Insbesondere ist für $a/b = c$ die Fisher-Informationsmatrix $\mathbf{I}_{\text{Ges}}(\boldsymbol{\beta}, a)$ ebenfalls eine Blockdiagonalmatrix.

3 Optimale Versuchsplanung

Die Qualität der Parameterschätzungen hängt auch von den gewählten Versuchseinstellungen ab. Um die Parameter bestmöglich zu schätzen, müssen die Versuchseinstellungen in einem gewissen Sinne optimal gewählt werden. In diesem Kapitel werden die theoretischen Grundlagen der optimalen Versuchsplanung dargestellt.

3.1 Individuelle Designs und Populationsdesigns

Zunächst wird eine einzelne statistische Einheit betrachtet. Zur Durchführung des Versuchs werden Versuchseinstellungen $\mathbf{x}_1, \dots, \mathbf{x}_l \in \mathcal{X}$ gewählt, wobei $\mathcal{X} \subset \mathbb{R}^k$ der Designraum ist, der entweder kompakt oder gleich \mathbb{R}^k ist. Die Häufigkeiten der Verwendung der Versuchseinstellungen sind durch $r_1, \dots, r_l \in \mathbb{N}$ gegeben, wobei $\sum_{j=1}^l r_j = m$ gilt. Dann ist ein individuelles Design ξ gegeben durch:

$$\xi = \left\{ \begin{array}{ccc} \mathbf{x}_1 & \dots & \mathbf{x}_l \\ r_1 & \dots & r_l \end{array} \right\}. \quad (3.1)$$

Anstatt der absoluten Häufigkeiten werden oft die relativen Häufigkeiten $w_j = r_j/m$, $j = 1, \dots, l$, betrachtet. Sie geben an, wie oft die entsprechende Versuchseinstellung für die statistische Einheit verwendet wird. Somit lässt sich das Design ξ schreiben als:

$$\xi = \left\{ \begin{array}{ccc} \mathbf{x}_1 & \dots & \mathbf{x}_l \\ r_1/m & \dots & r_l/m \end{array} \right\}. \quad (3.2)$$

So ein Versuchsplan ξ wird exaktes individuelles Design genannt, welches sich als diskretes Wahrscheinlichkeitsmaß auffassen lässt. Es soll ein Design ξ gefunden werden, so dass die Parameter möglichst gut geschätzt werden können. Da die Gewichte $w_j = r_j/m$ nur ganzzahlige Vielfache von $1/m$ annehmen können, müsste somit das Design ξ über eine Teilmenge der ganzen Zahlen optimiert werden. Um das Optimierungsproblem zu vereinfachen, wird das Konzept der exakten individuellen Designs auf approximative

individuelle Designs

$$\xi = \begin{Bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_l \\ w_1 & \cdots & w_l \end{Bmatrix} \quad (3.3)$$

verallgemeinert, welche Wahrscheinlichkeitsmaße auf den Borel-Mengen von \mathcal{X} mit endlichem Träger sind (vgl. Silvey, 1980, S. 15). Es werden nicht nur diskrete Gewichte $w_j = r_j/m$ zugelassen, sondern beliebige Gewichte $0 \leq w_1, \dots, w_l \leq 1$ mit $\sum_{j=1}^l w_j = 1$. Die Menge aller approximativen Designs ξ auf \mathcal{X} wird mit Ξ bezeichnet. Die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ für ein Design ξ ergibt sich durch Standardisierung der Fisher-Informationsmatrix mit der Anzahl der Beobachtungen m und der Zulassung von beliebigen Gewichten. Die Menge der Informationsmatrizen für ein gegebenes $\boldsymbol{\beta}$ sei $\mathcal{M}_{\boldsymbol{\beta}} = \{\mathbf{M}(\xi; \boldsymbol{\beta}) : \xi \in \Xi\}$.

In der praktischen Anwendung kann der Fall auftreten, dass ein approximatives Design bei vorgegebener Anzahl m an Beobachtungen nicht realisiert werden kann. Dann müssen die Gewichte geeignet gerundet werden, um Vielfache von $1/m$ zu erhalten (vgl. Pukelsheim und Rieder, 1992).

Für das gesamte Experiment mit n statistischen Einheiten besteht das Populationsdesign

$$\zeta = \begin{Bmatrix} \xi_1 & \cdots & \xi_r \\ q_1 & \cdots & q_r \end{Bmatrix} \quad (3.4)$$

aus den individuellen Designs ξ_1, \dots, ξ_r und zugehörigen Gewichten $0 \leq q_1, \dots, q_r \leq 1$ mit $\sum_{i=1}^r q_i = 1$. Das Gewicht q_i gibt den Anteil der statistischen Einheiten an, welchen das individuelle Design ξ_i zugewiesen wird. Da die Beobachtungen zwischen den statistischen Einheiten unabhängig sind, kann die Informationsmatrix für das Populationsdesign ζ durch $\mathbf{M}(\zeta; \boldsymbol{\beta}) = \sum_{i=1}^r q_i \mathbf{M}(\xi_i; \boldsymbol{\beta})$ erhalten werden.

3.2 Optimalitätskriterien

Die Versuchseinstellungen sind so zu wählen, dass die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ maximal wird. Für einen mehrdimensionalen Parametervektor $\boldsymbol{\beta}$ ist $\mathbf{M}(\xi^*; \boldsymbol{\beta}) \geq \mathbf{M}(\xi; \boldsymbol{\beta})$ im Sinne von $\mathbf{M}(\xi^*; \boldsymbol{\beta}) - \mathbf{M}(\xi; \boldsymbol{\beta})$ positiv semidefinit für alle $\xi \in \Xi$ in der Regel nicht zu erfüllen. Daher erfolgt die Einführung von Kriteriumsfunktionen und somit Optimalitätskriterien, um verschiedene Designs trotzdem miteinander vergleichen zu können. Optimale Designs basieren somit auf der Optimierung einer reellwertigen Funktion ϕ der Informationsmatrix bezüglich des Designs (vgl. Silvey, 1980, S. 10). Die Optimalitätskriterien werden bezüglich des individuellen Designs ξ eingeführt. Für Populationsdesigns

ζ können sie analog definiert werden.

Ein Optimalitätskriterium lässt sich schreiben als $\Phi(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$. Jedes Optimalitätskriterium kann als ein Maximierungsproblem oder ein Minimierungsproblem aufgefasst werden, wobei beide Optimalitätsprobleme beispielsweise durch Vorzeichenwechsel ineinander transformiert werden können. Da die Anzahl der Stützpunkte eines optimalen Designs vor der Optimierung nicht bekannt ist, handelt es sich um Optimierungsprobleme mit unbekannter Dimension. Beim Maximierungsproblem sei $\Phi : \Xi \rightarrow \mathbb{R} \cup \{-\infty\}$. Ein Design $\xi^* \in \Xi$ ist Φ -optimal, falls $\Phi(\xi^*) \geq \Phi(\xi)$ für alle $\xi \in \Xi$ gilt (vgl. Schwabe, 1996, S. 8). Analog lässt sich das Minimierungsproblem mit $\Phi : \Xi \rightarrow \mathbb{R} \cup \{\infty\}$ definieren.

Die Wahl eines geeigneten Optimalitätskriteriums hängt von verschiedenen Faktoren ab. Ein wichtiges Auswahlkriterium liegt darin, ob der gesamte Parametervektor $\boldsymbol{\beta}$ oder nur ein Parametersubsystem geschätzt werden soll. Um den gesamten Parametervektor schätzen zu können, muss die Informationsmatrix regulär sein. Soll hingegen nur ein Parametersystem $\mathbf{A}^T \boldsymbol{\beta}$ mit einer $p \times s$ -Matrix \mathbf{A} mit $\text{Rang}(\mathbf{A}) = s < p$ geschätzt werden, so muss die Informationsmatrix nicht regulär sein. Daher wird nun der Begriff der Identifizierbarkeit eingeführt (vgl. Silvey, 1980, S. 25).

Definition 3.1 (Identifizierbarkeit)

Sei \mathbf{A} eine $p \times s$ -Matrix mit $\text{Rang}(\mathbf{A}) = s < p$. Für gegebenes $\boldsymbol{\beta}$ sind die Linearkombinationen $\mathbf{A}^T \boldsymbol{\beta}$ identifizierbar für ein Design ξ , wenn $\mathbf{A} = \mathbf{M}(\xi; \boldsymbol{\beta}) \mathbf{H}$ für eine Matrix $\mathbf{H} \in \mathbb{R}^{p \times s}$ gilt.

Besteht Interesse die Linearkombinationen $\mathbf{A}^T \boldsymbol{\beta}$ zu schätzen, muss die Identifizierbarkeitsbedingung erfüllt sein. Es werden Optimalitätskriterien betrachtet, bei denen sich die zu maximierende Kriteriumsfunktion als

$$\phi(\mathbf{M}(\xi; \boldsymbol{\beta})) = \begin{cases} \varphi(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^- \mathbf{A}) & \text{falls } \mathbf{M}(\xi; \boldsymbol{\beta}) \in \mathcal{M}_{\boldsymbol{\beta}, \mathbf{A}} \\ -\infty & \text{sonst} \end{cases} \quad (3.5)$$

darstellen lässt. Dabei ist $\mathbf{M}(\xi; \boldsymbol{\beta})^-$ eine verallgemeinerte Inverse der Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$. Mit $\mathcal{M}_{\boldsymbol{\beta}, \mathbf{A}}$ wird die Menge der Informationsmatrizen $\mathbf{M} \in \mathcal{M}_{\boldsymbol{\beta}}$ bezeichnet, für welche $\mathbf{A}^T \boldsymbol{\beta}$ identifizierbar ist. Für $\mathbf{M} \in \mathcal{M}_{\boldsymbol{\beta}, \mathbf{A}}$ hängt $\mathbf{A}^T \mathbf{M}^- \mathbf{A}$ nicht von der Wahl der verallgemeinerten Inverse ab und $\mathbf{A}^T \mathbf{M}^- \mathbf{A}$ ist eine positiv definite $s \times s$ -Matrix (vgl. Silvey, 1980, S. 25).

Bei nichtlinearen Modellen wie dem Poisson-Gamma-Modell hängen die optimalen Designs vom Parametervektor $\boldsymbol{\beta}$ ab. Daher werden sie auch lokal optimale Designs genannt (vgl. Chernoff, 1953).

In der folgenden Definition wird das D -Optimalitätskriterium definiert, bei dem die

Determinante der Informationsmatrix maximiert wird und welches das am häufigsten verwendete Optimalitätskriterium ist.

Definition 3.2 (D -Optimalität)

Ein Design ξ^* mit regulärer Informationsmatrix $\mathbf{M}(\xi^*; \boldsymbol{\beta})$ ist D -optimal, falls die Ungleichung $\det(\mathbf{M}(\xi^*; \boldsymbol{\beta})) \geq \det(\mathbf{M}(\xi; \boldsymbol{\beta}))$ für alle $\xi \in \Xi$ gilt.

Das D -optimale Design minimiert das Volumen des asymptotischen Konfidenzellipsoids für die Parameter (vgl. Silvey, 1980, S. 10). Ein weiterer Vorteil ist, dass das D -optimale Design invariant bezüglich regulärer linearer Transformationen des Parameters ist (vgl. Pronzato und Pázman, 2013, S. 110).

Das D -Optimalitätskriterium lässt sich auch in der Form wie in Gleichung (3.5) schreiben. Für reguläre Informationsmatrizen kann auch der Logarithmus der Determinante betrachtet werden, da es sich um eine streng monoton wachsende Funktion handelt. Ist die Informationsmatrix singulär, so wird ihr der Wert $-\infty$ zugewiesen. Dann gilt für das D -Optimalitätskriterium:

$$\phi(\mathbf{M}(\xi; \boldsymbol{\beta})) = \begin{cases} \ln(\det(\mathbf{M}(\xi; \boldsymbol{\beta}))) & \text{falls } \mathbf{M}(\xi; \boldsymbol{\beta}) \text{ regulär} \\ -\infty & \text{falls } \mathbf{M}(\xi; \boldsymbol{\beta}) \text{ singulär} \end{cases}. \quad (3.6)$$

Die so definierte Kriteriumsfunktion ist streng konkav auf der Menge der positiv definiten Matrizen und somit ist die Informationsmatrix $\mathbf{M}(\xi^*; \boldsymbol{\beta})$ eindeutig (vgl. Silvey, 1980, S. 17, 41). Dies bedeutet im Allgemeinen jedoch nicht, dass das optimale Design ξ^* eindeutig ist.

Das D_A -Optimalitätskriterium kann zur Berechnung von optimalen Designs zur Schätzung von $\mathbf{A}^T \boldsymbol{\beta}$ verwendet werden.

Definition 3.3 (D_A -Optimalität)

Es sei \mathbf{A} eine $p \times s$ -Matrix mit $\text{Rang}(\mathbf{A}) = s < p$. Ein Design ξ^* ist D_A -optimal, falls $\mathbf{A}^T \boldsymbol{\beta}$ identifizierbar ist und $\det(\mathbf{A}^T \mathbf{M}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{A}) \leq \det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A})$ für alle $\xi \in \Xi$ gilt, für die $\mathbf{A}^T \boldsymbol{\beta}$ identifizierbar ist.

Sind nur s einzelne Parameter von $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$ von Interesse, so wird das Kriterium D_s -Optimalität genannt. Für $\mathbf{A}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1})$ mit $(p-1)$ -dimensionalem Nullvektor $\mathbf{0}_{p-1}$ und $(p-1) \times (p-1)$ -Einheitsmatrix \mathbf{I}_{p-1} ergibt sich D_s -Optimalität für die $s = p-1$ Parameter $\beta_1, \dots, \beta_{p-1}$ (vgl. Silvey, 1980, S. 11, 26).

Die symmetrische positiv semidefinite Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ lässt sich als Block-

matrix

$$\mathbf{M}(\xi; \boldsymbol{\beta}) = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{12}^T & \mathbf{M}_{22} \end{pmatrix} \quad (3.7)$$

mit $(p-s) \times (p-s)$ -Teilmatrix $\mathbf{M}_{11} = \mathbf{M}_{11}(\xi; \boldsymbol{\beta})$ und $s \times s$ -Teilmatrix $\mathbf{M}_{22} = \mathbf{M}_{22}(\xi; \boldsymbol{\beta})$ darstellen. Mit Satz A.1 b) folgt für die Kriteriumsfunktion bei D_s -Optimalität mit $\mathbf{A}^T = (\mathbf{0}_{s,p-s}, \mathbf{I}_s)$ (vgl. Pronzato und Pázman, 2013, S. 110):

$$\det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}) = \det(\mathbf{M}_{22} - \mathbf{M}_{12}^T \mathbf{M}_{11}^{-1} \mathbf{M}_{12})^{-1}. \quad (3.8)$$

Dabei ist $\mathbf{S} = \mathbf{M}_{22} - \mathbf{M}_{12}^T \mathbf{M}_{11}^{-1} \mathbf{M}_{12}$ das verallgemeinerte Schurkomplement von \mathbf{M}_{11} in $\mathbf{M}(\xi; \boldsymbol{\beta})$. Die Linearkombinationen $\mathbf{A}^T \boldsymbol{\beta}$ sind genau dann identifizierbar, wenn \mathbf{S} regulär ist (vgl. Pronzato und Pázman, 2013, S. 110).

Satz 3.4

Es sei $\mathbf{A}^T = (\mathbf{0}_{s,p-s}, \mathbf{I}_s)$ und die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ sei eine 2×2 -Blockmatrix wie in Gleichung (3.7). Ist die Teilmatrix $\mathbf{M}_{11}(\xi; \boldsymbol{\beta})$ für $\xi \in \Xi$ regulär, dann ist $\mathbf{A}^T \boldsymbol{\beta}$ genau dann für ξ identifizierbar, wenn $\mathbf{M}(\xi; \boldsymbol{\beta})$ regulär ist. Falls $\mathbf{M}_{11}(\xi; \boldsymbol{\beta})$ für alle $\xi \in \Xi$ regulär ist, muss ein optimales Design eine reguläre Informationsmatrix besitzen.

Beweis:

Die Identifizierbarkeit von $\mathbf{A}^T \boldsymbol{\beta}$ ist äquivalent zur Regularität des Schurkomplements $\mathbf{S} = \mathbf{M}_{22} - \mathbf{M}_{12}^T \mathbf{M}_{11}^{-1} \mathbf{M}_{12}$ von $\mathbf{M}_{11} = \mathbf{M}_{11}(\xi; \boldsymbol{\beta})$ in $\mathbf{M}(\xi; \boldsymbol{\beta})$. Für reguläre Teilmatrix \mathbf{M}_{11} ist \mathbf{S} nach Satz A.1 a) wiederum genau dann regulär, wenn $\mathbf{M}(\xi; \boldsymbol{\beta})$ regulär ist. Somit folgt die Äquivalenz der Identifizierbarkeit und der Regularität der Informationsmatrix. \square

Im folgenden Satz wird die Kriteriumsfunktion für D_s -Optimalität im Fall einer regulären Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ berechnet (vgl. Atkinson et al., 2007, S. 139).

Satz 3.5

Es sei $\mathbf{A}^T = (\mathbf{0}_{s,p-s}, \mathbf{I}_s)$ und $\mathbf{M}(\xi; \boldsymbol{\beta})$ sei regulär mit Darstellung als 2×2 -Blockmatrix wie in Gleichung (3.7). Dann gilt für die Kriteriumsfunktion für D_s -Optimalität:

$$\det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}) = \frac{\det(\mathbf{M}_{11})}{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}. \quad (3.9)$$

Beweis:

Da $\mathbf{M}(\xi; \boldsymbol{\beta})$ regulär ist, ist auch \mathbf{M}_{11} regulär. Somit folgt mit Gleichung (3.8) für die

Kriteriumsfunktion:

$$\det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}) = \det(\mathbf{M}_{22} - \mathbf{M}_{12}^T \mathbf{M}_{11}^{-1} \mathbf{M}_{12})^{-1}.$$

Mit Satz A.2 folgt

$$\det(\mathbf{M}_{22} - \mathbf{M}_{12}^T \mathbf{M}_{11}^{-1} \mathbf{M}_{12})^{-1} = \left(\frac{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}{\det(\mathbf{M}_{11})} \right)^{-1} = \frac{\det(\mathbf{M}_{11})}{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}$$

für die Kriteriumsfunktion. □

Besteht Interesse eine Linearkombination $\mathbf{c}^T \boldsymbol{\beta}$ mit minimaler Varianz zu schätzen, so wird als Optimalitätskriterium c -Optimalität verwendet. Die Informationsmatrix eines c -optimalen Designs ist nicht notwendigerweise regulär. Es wird aber die Identifizierbarkeit von $\mathbf{c}^T \boldsymbol{\beta}$ gefordert, d. h. es muss $\mathbf{c} = \mathbf{M}(\xi; \boldsymbol{\beta}) \mathbf{h}$ für einen Vektor $\mathbf{h} \in \mathbb{R}^p$ gelten.

Definition 3.6 (c -Optimalität)

Ein Design ξ^* ist c -optimal, falls die Linearkombination $\mathbf{c}^T \boldsymbol{\beta}$ identifizierbar ist und falls $\mathbf{c}^T \mathbf{M}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{c} \leq \mathbf{c}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{c}$ für alle $\xi \in \Xi$ gilt, für die $\mathbf{c}^T \boldsymbol{\beta}$ identifizierbar ist.

Während im Linearen Modell beim c -Optimalitätskriterium die Varianz des Schätzers von $\mathbf{c}^T \boldsymbol{\beta}$ minimiert wird, entspricht c -Optimalität bei nichtlinearen Modellen der Minimierung der asymptotischen Varianz (vgl. Atkinson et al., 2007, S. 142–143).

Ein weiteres Kriterium ist L -Optimalität. Es findet Anwendung, wenn Interesse darin besteht Linearkombinationen $\mathbf{A}^T \boldsymbol{\beta}$ zu schätzen.

Definition 3.7 (L -Optimalität)

Es sei $\mathbf{B} = \mathbf{A} \mathbf{A}^T$ eine symmetrische positiv semidefinite Matrix. Ein Design ξ^* ist L -optimal, falls $\mathbf{A}^T \boldsymbol{\beta}$ identifizierbar ist und $\text{Spur}(\mathbf{M}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{B}) \leq \text{Spur}(\mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{B})$ für alle $\xi \in \Xi$ gilt, für die $\mathbf{A}^T \boldsymbol{\beta}$ identifizierbar ist.

Ist $\mathbf{B} = \mathbf{I}_p$ die Einheitsmatrix, so ergibt sich A -Optimalität. Ein A -optimales Design minimiert die Summe der asymptotischen Varianzen der Schätzer der einzelnen Komponenten des Parametervektors.

Das *IMSE*-Optimalitätskriterium, bei dem der Integrated Mean Squared Error minimiert wird, ist ebenfalls ein Spezialfall von L -Optimalität. Beim Poisson-Gamma-Modell folgt für die asymptotische Varianz des Schätzers der Wirkungsfunktion

$$\hat{Y}(\mathbf{x}) = \frac{a}{b} \cdot \exp(\mathbf{f}(\mathbf{x})^T \hat{\boldsymbol{\beta}})$$

mit der Delta-Regel:

$$\text{Var}_{\text{asymp}}(\hat{Y}(\mathbf{x})) = \left(\frac{a}{b}\right)^2 \cdot \exp(2\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}).$$

Beim *IMSE*-Kriterium wird für eine Wahrscheinlichkeitsverteilung μ das Integral

$$\begin{aligned} & \int_{\mathcal{X}} \text{Var}_{\text{asymp}}(\hat{Y}(\mathbf{x})) \mu(d\mathbf{x}) \\ &= \int_{\mathcal{X}} \left(\frac{a}{b}\right)^2 \cdot \exp(2\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \mu(d\mathbf{x}) \\ &= \text{Spur} \left(\mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \cdot \int_{\mathcal{X}} \left(\frac{a}{b}\right)^2 \cdot \exp(2\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T \mu(d\mathbf{x}) \right) \\ &= \text{Spur}(\mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{B}) \end{aligned}$$

mit der Matrix

$$\mathbf{B} = \int_{\mathcal{X}} \left(\frac{a}{b}\right)^2 \cdot \exp(2\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T \mu(d\mathbf{x})$$

bezüglich des Designs ξ minimiert (vgl. Atkinson et al., 2007, S. 143).

Beim Poisson-Modell lautet der Schätzer der Wirkungsfunktion $\hat{Y}_{\text{Po}}(\mathbf{x}) = \exp(\mathbf{f}(\mathbf{x})^T \hat{\boldsymbol{\beta}})$. Es folgt somit $\mathbf{B}_{\text{Po}} = \int_{\mathcal{X}} \exp(2\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T \mu(d\mathbf{x})$. Also sind die Matrizen \mathbf{B} und \mathbf{B}_{Po} bis auf den multiplikativen Faktor $(a/b)^2$ identisch.

Das *IMSE*-Optimalitätskriterium ist wie das *D*-Optimalitätskriterium invariant bezüglich regulärer linearer Transformationen des Parameters.

Für $\mathbf{B} = \mathbf{c}\mathbf{c}^T$ mit $\mathbf{c} \in \mathbb{R}^p$ ergibt sich *c*-Optimalität als Spezialfall von *L*-Optimalität, da für die Kriteriumsfunktion

$$\text{Spur}(\mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{B}) = \text{Spur}(\mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{c}\mathbf{c}^T) = \text{Spur}(\mathbf{c}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{c}) = \mathbf{c}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{c}$$

gilt. Hierbei wurde die Invarianz der Spur unter zyklischen Vertauschungen verwendet. Um die Güte verschiedener Designs miteinander zu vergleichen, kann die Effizienz eines Designs bestimmt werden. Dazu werden positiv homogene Optimalitätskriterien mit zu maximierender Kriteriumsfunktion betrachtet. Ein Optimalitätskriterium ϕ ist positiv homogen, wenn $\phi(\delta \cdot \mathbf{M}) = \delta \cdot \phi(\mathbf{M})$ für alle $\delta > 0$ und für alle symmetrischen positiv semidefiniten Matrizen \mathbf{M} gilt (vgl. Pukelsheim, 1993, S. 115).

Definition 3.8 (Effizienz)

Es sei ϕ ein positiv homogenes Optimalitätskriterium mit zu maximierender Kriteriums-
funktion und ξ^* sei ein ϕ -optimales Design. Dann ist die ϕ -Effizienz eines Designs ξ
gegeben durch:

$$\text{eff}_\phi(\xi; \boldsymbol{\beta}) = \frac{\phi(\mathbf{M}(\xi; \boldsymbol{\beta}))}{\phi(\mathbf{M}(\xi^*; \boldsymbol{\beta}))}. \quad (3.10)$$

Ist das Optimalitätskriterium $\phi > 0$ zu minimieren und $1/\phi$ ein positiv homogenes Op-
timalitätskriterium, dann ist die ϕ -Effizienz eines Designs ξ nach Definition 3.8 gegeben
durch:

$$\text{eff}_\phi(\xi; \boldsymbol{\beta}) = \frac{\phi(\mathbf{M}(\xi^*; \boldsymbol{\beta}))}{\phi(\mathbf{M}(\xi; \boldsymbol{\beta}))}. \quad (3.11)$$

Die Effizienz eines Designs liegt immer im Intervall $[0, 1]$, wobei ein optimales Design eine
Effizienz von eins besitzt. Eine positiv homogene Kriteriums-
funktion für D -Optimalität ist durch $\phi(\mathbf{M}(\xi; \boldsymbol{\beta})) = \det(\mathbf{M}(\xi; \boldsymbol{\beta}))^{1/p}$ gegeben. Für die D -Effizienz eines Designs ξ
gilt dann:

$$\text{eff}_D(\xi; \boldsymbol{\beta}) = \left(\frac{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi^*; \boldsymbol{\beta}))} \right)^{\frac{1}{p}}. \quad (3.12)$$

Durch die p -te Wurzel ist die D -Effizienz antiproportional zur Anzahl der Versuchs-
durchführungen. Würde die Effizienz des Designs ξ bei 0.5 liegen, so wären doppelt so
viele Versuchsdurchführungen wie beim optimalen Design ξ^* notwendig, um die gleiche
Güte zu erhalten (vgl. Atkinson et al., 2007, S. 151).

Da das c -Optimalitätskriterium sich als Maximierungsproblem mit positiv homogener
Kriteriums-
funktion $\phi(\mathbf{M}(\xi; \boldsymbol{\beta})) = (\mathbf{c}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{c})^{-1}$ schreiben lässt, folgt für die c -
Effizienz eines Designs ξ :

$$\text{eff}_c(\xi; \boldsymbol{\beta}) = \frac{\mathbf{c}^T \mathbf{M}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{c}}{\mathbf{c}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{c}}. \quad (3.13)$$

Bei D_A -Optimalität mit einer $p \times s$ -Matrix \mathbf{A} ist eine positiv homogene Kriteriums-
funktion $\phi(\mathbf{M}(\xi; \boldsymbol{\beta})) = \det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A})^{-1/s}$. Die D_A -Effizienz eines Designs ξ ist somit
gegeben durch:

$$\text{eff}_{D_A}(\xi; \boldsymbol{\beta}) = \left(\frac{\det(\mathbf{A}^T \mathbf{M}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{A})}{\det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A})} \right)^{\frac{1}{s}}. \quad (3.14)$$

Bei nichtlinearen Modellen hängen die optimalen Designs vom unbekanntem Parametervektor $\boldsymbol{\beta}$ ab. Um diese Abhängigkeit einzuschränken, kann eine Menge von möglichen Parameterwerten B vorgegeben werden. Dieses Konzept führt zu standardisiert Maximin optimalen Designs (vgl. Müller (1995), Dette (1997)).

Definition 3.9 (Standardisierte Maximin Optimalität)

Es sei $B \subseteq \mathbb{R}^p$ kompakt. Ein Design ξ^* ist standardisiert Maximin ϕ -optimal für $\mathbf{A}^T \boldsymbol{\beta}$, falls $\mathbf{A}^T \boldsymbol{\beta}$ für alle $\boldsymbol{\beta} \in B$ identifizierbar ist und

$$\min_{\boldsymbol{\beta} \in B} \text{eff}_\phi(\xi^*; \boldsymbol{\beta}) \geq \min_{\boldsymbol{\beta} \in B} \text{eff}_\phi(\xi; \boldsymbol{\beta}) \quad (3.15)$$

für alle $\xi \in \Xi$ gilt, für die $\mathbf{A}^T \boldsymbol{\beta}$ für alle $\boldsymbol{\beta} \in B$ identifizierbar ist.

Bei lokal optimalen Designs ist die Anzahl der Stützpunkte nach dem Satz von Carathéodory beschränkt. Dies gilt im Allgemeinen nicht bei standardisiert Maximin optimalen Designs, bei welchen die Anzahl der Stützpunkte sogar beliebig groß werden kann (vgl. Braess und Dette, 2007).

3.3 Optimale Gewichte

In diesem Abschnitt werden Informationsmatrizen von der Form

$$\mathbf{M}(\xi; \boldsymbol{\beta}) = \sum_{j=1}^l w_j \cdot \lambda(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T \quad (3.16)$$

betrachtet. Dabei ist λ die Intensitätsfunktion (vgl. Fedorov, 1972, S. 39). Eine Informationsmatrix von dieser Form besitzen beispielsweise Verallgemeinerte Lineare Modelle (vgl. McCulloch und Searle, 2001, S. 142), worunter auch das Poisson-Modell fällt, bei dem die Intensitätsfunktion $\lambda(z) = e^z$ ist.

Mit Designmatrix $\mathbf{X} = (\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_l))^T$, Gewichtsmatrix $\mathbf{W} = \text{Diag}(w_1, \dots, w_l)$ und Intensitätsmatrix $\boldsymbol{\Lambda} = \text{Diag}(\lambda(\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta}), \dots, \lambda(\mathbf{f}(\mathbf{x}_l)^T \boldsymbol{\beta}))$ folgt die Matrix-Darstellung:

$$\mathbf{M}(\xi; \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \mathbf{X}. \quad (3.17)$$

Handelt es sich beim Optimalitätskriterium Φ um ein Maximierungsproblem, dann sind die Φ -optimalen Gewichte diejenigen Gewichte, welche die Kriteriumsfunktion für ein Design mit gegebenen Stützpunkten $\mathbf{x}_1, \dots, \mathbf{x}_l$ maximieren. Ein D -optimales Design

muss mindestens p Stützpunkte besitzen. Bei D_A - und c -Optimalität können auch optimale Designs mit weniger als p Stützpunkten auftreten. Der folgende Satz gibt für die verschiedenen Optimalitätskriterien die optimalen Gewichte unter der Bedingung unabhängiger Vektoren der Regressionsfunktionen $\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_l)$ an (vgl. Pukelsheim und Torsney, 1991).

Satz 3.10

a) Für die D -optimalen Gewichte für ein Design ξ mit p Stützpunkten gilt:

$$w_j^* = \frac{1}{p} \quad \text{für } j = 1, \dots, p. \quad (3.18)$$

b) Es sei \mathbf{A} eine $p \times s$ -Matrix. Für die D_A -optimalen Gewichte für ein Design ξ mit $l \leq p$ Stützpunkten gilt

$$w_j^* = \frac{\sqrt{b_j}}{\sum_{i=1}^l \sqrt{b_i}} \quad \text{für } j = 1, \dots, l, \quad (3.19)$$

wobei b_1, \dots, b_l die nicht negativen Hauptdiagonalelemente der positiv semidefiniten Matrix $\mathbf{B} = \mathbf{V}\mathbf{C}\mathbf{V}^T$ mit $\mathbf{V} = \mathbf{\Lambda}^{-\frac{1}{2}}(\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\mathbf{A}$ und $\mathbf{C} = (\mathbf{A}^T\mathbf{M}(\xi; \boldsymbol{\beta}) - \mathbf{A})^{-1}$ sind.

c) Es sei $\mathbf{c} \in \mathbb{R}^p$. Für die c -optimalen Gewichte für ein Design ξ mit $l \leq p$ Stützpunkten gilt

$$w_j^* = \frac{|v_j|}{\sum_{i=1}^l |v_i|} \quad \text{für } j = 1, \dots, l, \quad (3.20)$$

wobei v_1, \dots, v_l die Komponenten des Vektors $\mathbf{v} = (\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{X}\mathbf{X}^T\mathbf{\Lambda}^{\frac{1}{2}})^{-1}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{X}\mathbf{c}$ sind.

Satz 3.10 für die optimalen Gewichte gilt auch, wenn es sich bei $\mathbf{x}_1, \dots, \mathbf{x}_l$ nicht um die optimalen Stützpunkte handelt.

3.4 Äquivalenzsätze

Der Beweis der Optimalität eines Designs erfolgt häufig unter Verwendung eines Äquivalenzsatzes. Der Begriff geht zurück auf Kiefer und Wolfowitz (1960), welche einen Äquivalenzsatz für D -Optimalität für lineare Modelle hergeleitet haben.

Es wird der Fall eines Optimalitätskriteriums mit zu maximierender Kriteriumsfunktion betrachtet. Dabei sei das Optimalitätskriterium $\Phi(\xi) = \Phi_{\boldsymbol{\beta}}(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$ konkav auf der Menge der Designs Ξ , also gilt

$$\Phi(\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2) \geq \alpha \cdot \Phi(\xi_1) + (1 - \alpha) \cdot \Phi(\xi_2) \quad (3.21)$$

für alle $\alpha \in [0, 1]$ und $\xi_1, \xi_2 \in \Xi$. Die betrachteten Optimalitätskriterien können alle in ein Maximierungsproblem mit konkaver Funktion Φ transformiert werden (vgl. Pukelsheim (1993, S. 151), Pronzato und Pázman (2013, S. 116–117)).

Zur Formulierung des Äquivalenzsatzes für Φ -Optimalität wird die Fréchet-Ableitung $F_{\Phi_{\beta}}(\xi_1, \xi_2) = F_{\Phi}(\xi_1, \xi_2)$ von $\Phi(\cdot)$ an der Stelle ξ_1 in Richtung ξ_2 eingeführt (vgl. Pronzato und Pázman, 2013, S. 126):

$$F_{\Phi}(\xi_1, \xi_2) = \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \cdot [\Phi((1 - \epsilon) \cdot \xi_1 + \epsilon \cdot \xi_2) - \Phi(\xi_1)]. \quad (3.22)$$

Es sei $\delta_{\mathbf{x}}$ das Dirac-Maß in $\mathbf{x} \in \mathcal{X}$. Also weist das Design $\delta_{\mathbf{x}}$ der Versuchseinstellung \mathbf{x} das Gewicht 1 zu. Dann gilt der folgende Äquivalenzsatz für Φ -Optimalität (vgl. Pronzato und Pázman, 2013, S. 131–133).

Satz 3.11 (Äquivalenzsatz)

Es sei $\Phi(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$ konkav auf Ξ und ϕ sei an der Stelle $\mathbf{M}(\xi^*; \boldsymbol{\beta})$ differenzierbar. Ein Design ξ^* ist genau dann Φ -optimal, wenn $F_{\Phi}(\xi^*, \delta_{\mathbf{x}}) \leq 0$ für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

Die Funktion $d(\xi, \mathbf{x}) = F_{\Phi}(\xi, \delta_{\mathbf{x}})$ wird Sensitivitätsfunktion genannt.

Ist $\Phi_{\beta}(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$ konkav auf Ξ , dann ist auch $\Phi_{\text{eff}, \boldsymbol{\beta}}(\xi) = \text{eff}_{\phi}(\xi; \boldsymbol{\beta})$ konkav auf Ξ . Daher ist $\Phi_{\text{min eff}}(\xi) = \min_{\boldsymbol{\beta} \in B} \text{eff}_{\phi}(\xi; \boldsymbol{\beta})$ als das Minimum von auf Ξ konkaven Funktionen ebenfalls konkav. Somit kann auch für standardisiert Maximin Φ -Optimalität ein Äquivalenzsatz angegeben werden (vgl. Pronzato und Pázman, 2013, S. 136, 245).

Satz 3.12 (Äquivalenzsatz für standardisierte Maximin Optimalität)

Es sei $\Phi_{\beta}(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$ konkav auf Ξ für alle $\boldsymbol{\beta} \in B$ und ϕ sei an der Stelle $\mathbf{M}(\xi^*; \boldsymbol{\beta})$ für alle $\boldsymbol{\beta} \in B$ differenzierbar. Es sei ξ^* ein Design und

$$\mathcal{N}(\xi^*) = \left\{ \boldsymbol{\beta} \in B : \boldsymbol{\beta} = \arg \min_{\boldsymbol{\beta} \in B} \text{eff}_{\phi}(\xi^*; \boldsymbol{\beta}) \right\} \quad (3.23)$$

die Menge der Parametervektoren $\boldsymbol{\beta}$, bei denen die ϕ -Effizienz von ξ^* minimal wird. Das Design ξ^* ist genau dann standardisiert Maximin Φ -optimal, wenn eine Wahrscheinlichkeitsverteilung π auf $\mathcal{N}(\xi^*)$ existiert, so dass

$$\int_{\mathcal{N}(\xi^*)} F_{\Phi_{\beta}}(\xi^*, \delta_{\mathbf{x}}) \pi(d\boldsymbol{\beta}) \leq 0 \quad (3.24)$$

für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

4 Information und Optimalität für das Poisson-Gamma-Modell

Es wird das Poisson-Gamma-Modell mit bekannten Parametern a und b der Gamma-Verteilung betrachtet. Die in Gleichung (2.12) gegebene Fisher-Informationsmatrix wird zur Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta}) = m^{-1} \mathbf{I}(\boldsymbol{\beta})$ eines Designs ξ verallgemeinert:

$$\mathbf{M}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \left(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) - \frac{\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} + \frac{b}{m}} \right). \quad (4.1)$$

Dabei ist

$$\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \sum_{j=1}^l w_j \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T \quad (4.2)$$

die Informationsmatrix für das Poisson-Modell. Somit lässt sich die Informationsmatrix für das Poisson-Gamma-Modell unter Verwendung der Informationsmatrix für das Poisson-Modell darstellen. Mit der Designmatrix $\mathbf{X} = (\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_l))^T$ und den Diagonalmatrizen $\mathbf{W} = \text{Diag}(w_1, \dots, w_l)$ und $\mathbf{\Lambda} = \text{Diag}(\exp(\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta}), \dots, \exp(\mathbf{f}(\mathbf{x}_l)^T \boldsymbol{\beta}))$ folgt die Matrix-Darstellung der Informationsmatrix $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{W} \mathbf{\Lambda} \mathbf{X}$ für das Poisson-Modell.

Da der Parameter $a > 0$ der Gamma-Verteilung ein multiplikativer Faktor in der Informationsmatrix für das Poisson-Gamma-Modell ist, kann die Optimierung der Informationsmatrix unabhängig von ihm erfolgen. Somit hängt ein optimales Design nicht von a ab. Der Parameter $b > 0$ der Gamma-Verteilung als auch die Anzahl m der Beobachtungen je statistischer Einheit können hingegen einen Einfluss auf das optimale Design haben, aber b und m beeinflussen das optimale Design nur gemeinsam über den Quotienten b/m .

Lemma 4.1

Die Matrix $\mathbf{L}(\xi; \boldsymbol{\beta})$ sei gegeben durch:

$$\mathbf{L}(\xi; \boldsymbol{\beta}) = \mathbf{I}_p - \frac{\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} + \frac{b}{m}}. \quad (4.3)$$

Dann gelten die folgenden Aussagen:

- (i) $\mathbf{M}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{L}(\xi; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{L}(\xi; \boldsymbol{\beta})^T$.
- (ii) Die Matrix $\mathbf{L}(\xi; \boldsymbol{\beta})$ ist regulär mit $\det(\mathbf{L}(\xi; \boldsymbol{\beta})) = \left(\frac{m}{b} \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + 1\right)^{-1}$.
- (iii) $\mathbf{L}(\xi; \boldsymbol{\beta}) = \mathbf{I}_p - \frac{m}{a} \cdot \mathbf{M}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T$.

Unter Verwendung von Lemma 4.1, dessen Beweis sich im Anhang befindet, lassen sich weitere Zusammenhänge zwischen den Informationsmatrizen für das Poisson-Modell und das Poisson-Gamma-Modell herleiten (vgl. Schmidt und Schwabe, 2020).

Lemma 4.2

Für ein Design ξ besitzen die Informationsmatrizen $\mathbf{M}(\xi; \boldsymbol{\beta})$ für das Poisson-Gamma-Modell und $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ für das Poisson-Modell den gleichen Rang.

Beweis:

Nach Lemma 4.1 kann die Informationsmatrix als $\mathbf{M}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{L}(\xi; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ dargestellt werden, wobei $\mathbf{L}(\xi; \boldsymbol{\beta})$ eine reguläre Matrix ist und somit vollen Rang besitzt. Daher folgt $\text{Rang}(\mathbf{M}(\xi; \boldsymbol{\beta})) = \text{Rang}(\mathbf{L}(\xi; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})) = \text{Rang}(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))$. \square

Satz 4.3

Es sei \mathbf{A} eine $p \times s$ -Matrix mit $\text{Rang}(\mathbf{A}) = s \leq p$. Für ein Design ξ ist $\mathbf{A}^T \boldsymbol{\beta}$ genau dann im Poisson-Gamma-Modell identifizierbar, wenn $\mathbf{A}^T \boldsymbol{\beta}$ für das Design ξ im Poisson-Modell identifizierbar ist.

Beweis:

Es sei \mathbf{H} eine $p \times s$ -Matrix und $\tilde{\mathbf{H}} = \frac{b}{a} \cdot (\mathbf{L}(\xi; \boldsymbol{\beta})^T)^{-1} \mathbf{H}$ mit der regulären Matrix $\mathbf{L}(\xi; \boldsymbol{\beta})$ aus Lemma 4.1. Mit der Darstellung der Informationsmatrix für das Poisson-Gamma-Modell in Lemma 4.1 (i) gilt:

$$\mathbf{A} - \mathbf{M}(\xi; \boldsymbol{\beta}) \tilde{\mathbf{H}} = \mathbf{A} - \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{L}(\xi; \boldsymbol{\beta})^T (\mathbf{L}(\xi; \boldsymbol{\beta})^T)^{-1} \mathbf{H} = \mathbf{A} - \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{H}.$$

Somit ist $\mathbf{A} = \mathbf{M}(\xi; \boldsymbol{\beta}) \tilde{\mathbf{H}}$ äquivalent zu $\mathbf{A} = \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{H}$. Die Identifizierbarkeit von $\mathbf{A}^T \boldsymbol{\beta}$ ist daher in beiden Modellen äquivalent. \square

Auch für die verallgemeinerten Inversen der Informationsmatrizen für das Poisson- und das Poisson-Gamma-Modell lässt sich ein Zusammenhang herstellen, wie im folgenden Satz gezeigt wird (vgl. Schmidt und Schwabe, 2020).

Satz 4.4

Die Matrix

$$\mathbf{M}(\xi; \boldsymbol{\beta})^- = \frac{b}{a} \cdot \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^- + \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \quad (4.4)$$

ist genau dann eine verallgemeinerte Inverse von $\mathbf{M}(\xi; \boldsymbol{\beta})$, wenn $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^-$ eine verallgemeinerte Inverse von $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ ist.

Beweis:

Zur Vereinfachung der Notation seien $\mathbf{M} = \mathbf{M}(\xi; \boldsymbol{\beta})$ und $\mathbf{M}_{\text{Po}} = \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$. Es sei \mathbf{M}_{Po}^- eine $p \times p$ -Matrix und $\mathbf{M}^- = \frac{b}{a} \cdot \mathbf{M}_{\text{Po}}^- + \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T$. Die Matrix \mathbf{M}^- ist eine verallgemeinerte Inverse von \mathbf{M} , wenn sie die Bedingung $\mathbf{M}\mathbf{M}^-\mathbf{M} = \mathbf{M}$ erfüllt. Es gilt:

$$\mathbf{M}\mathbf{M}^-\mathbf{M} = \mathbf{M} \cdot \left(\frac{b}{a} \cdot \mathbf{M}_{\text{Po}}^- + \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \right) \cdot \mathbf{M} = \frac{b}{a} \cdot \mathbf{M}\mathbf{M}_{\text{Po}}^-\mathbf{M} + \frac{m}{a} \cdot \mathbf{M}\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{M}.$$

Mit der Matrix $\mathbf{L} = \mathbf{L}(\xi; \boldsymbol{\beta})$ aus Lemma 4.1 folgt:

$$\frac{b}{a} \cdot \mathbf{M}\mathbf{M}_{\text{Po}}^-\mathbf{M} = \frac{a}{b} \cdot \mathbf{L}\mathbf{M}_{\text{Po}}\mathbf{M}_{\text{Po}}^-\mathbf{M}_{\text{Po}}\mathbf{L}^T.$$

Außerdem gilt mit Lemma 4.1 (i) und (iii):

$$\frac{m}{a} \cdot \mathbf{M}\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{M} - \mathbf{M} = \left(\frac{m}{a} \cdot \mathbf{M}\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T - \mathbf{I}_p \right) \cdot \mathbf{M} = -\mathbf{L}\mathbf{M} = -\frac{a}{b} \cdot \mathbf{L}\mathbf{M}_{\text{Po}}\mathbf{L}^T.$$

Somit gilt:

$$\mathbf{M}\mathbf{M}^-\mathbf{M} - \mathbf{M} = \frac{a}{b} \cdot \mathbf{L} \cdot (\mathbf{M}_{\text{Po}}\mathbf{M}_{\text{Po}}^-\mathbf{M}_{\text{Po}} - \mathbf{M}_{\text{Po}}) \cdot \mathbf{L}^T.$$

Die Matrix \mathbf{L} ist nach Lemma 4.1 (ii) regulär. Daher ist $\mathbf{M}\mathbf{M}^-\mathbf{M} = \mathbf{M}$ äquivalent zu $\mathbf{M}_{\text{Po}}\mathbf{M}_{\text{Po}}^-\mathbf{M}_{\text{Po}} = \mathbf{M}_{\text{Po}}$. \square

Bemerkung 4.5

Nach Lemma 4.2 ist $\mathbf{M}(\xi; \boldsymbol{\beta})$ genau dann regulär, wenn $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ regulär ist. In diesem Fall können die verallgemeinerten Inversen in Satz 4.4 durch die gewöhnlichen Inversen ersetzt werden:

$$\mathbf{M}(\xi; \boldsymbol{\beta})^{-1} = \frac{b}{a} \cdot \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-1} + \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T. \quad (4.5)$$

Für Verallgemeinerte Lineare Modelle wie dem Poisson-Modell ist die Informationsmatrix einer Konvexkombination der Designs gleich der Konvexkombination der Informationsmatrizen dieser Designs (vgl. Fedorov, 1972, S. 66–67). Aufgrund des zufälligen Effekts gilt dies jedoch nicht für das Poisson-Gamma-Modell. Da die Informationsmatrix eine Darstellung der Form $\mathbf{M}(\xi; \boldsymbol{\beta}) = (\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})^{-1} + \mathbf{D})^{-1}$ mit $\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ und einer nicht vom Design ξ abhängigen Matrix $\mathbf{D} = \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T$ besitzt, kann das folgende Resultat gezeigt werden (vgl. Schmelter, 2007).

Satz 4.6

Es seien ξ_1 und ξ_2 zwei Designs. Dann gilt für alle $\alpha \in [0, 1]$ die folgende Ungleichung bezüglich der Loewner-Ordnung:

$$\mathbf{M}(\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2; \boldsymbol{\beta}) \geq \alpha \cdot \mathbf{M}(\xi_1; \boldsymbol{\beta}) + (1 - \alpha) \cdot \mathbf{M}(\xi_2; \boldsymbol{\beta}). \quad (4.6)$$

Ein zu dem Beweis von Schmelter (2007) alternativer Beweis befindet sich im Anhang. Eine Kriteriumsfunktion ϕ ist isoton, falls $\phi(\mathbf{M}_1) \geq \phi(\mathbf{M}_2)$ für alle positiv semidefiniten Matrizen $\mathbf{M}_1 \geq \mathbf{M}_2$ gilt. Alle betrachteten Optimalitätskriterien können in ein Maximierungsproblem mit isotoner Kriteriumsfunktion transformiert werden (vgl. Pronzato und Pázman, 2013, S. 114, 118).

Folgerung 4.7

Es sei $\zeta = \left\{ \begin{smallmatrix} \xi_1 & \dots & \xi_r \\ q_1 & \dots & q_r \end{smallmatrix} \right\}$ ein beliebiges Populationsdesign. Dann gilt $\mathbf{M}(\tilde{\zeta}; \boldsymbol{\beta}) \geq \mathbf{M}(\zeta; \boldsymbol{\beta})$ für das Populationsdesign $\tilde{\zeta} = \left\{ \begin{smallmatrix} \xi \\ 1 \end{smallmatrix} \right\}$, welches dem individuellen Design $\xi = \sum_{i=1}^r q_i \xi_i$ das Gewicht 1 zuweist.

Für ein isotones Optimalitätskriterium gilt somit $\phi(\mathbf{M}(\tilde{\zeta}; \boldsymbol{\beta})) \geq \phi(\mathbf{M}(\zeta; \boldsymbol{\beta}))$.

Da nach Folgerung 4.7 ein optimales individuelles Design ξ^* ein optimales Populationsdesign $\tilde{\zeta}^* = \left\{ \begin{smallmatrix} \xi^* \\ 1 \end{smallmatrix} \right\}$ liefert, welches für alle statistischen Einheiten das Design ξ^* verwendet (vgl. Schmelter, 2007), wird sich im Folgenden auf die Bestimmung von optimalen individuellen Designs beschränkt.

Eine notwendige Bedingung für die Herleitung von Äquivalenzsätzen für das Poisson-Gamma-Modell ist die Konkavität von $\Phi(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$ auf Ξ , welche im nächsten Satz nachgewiesen wird.

Satz 4.8

Es sei $\Phi(\xi) = \phi(\mathbf{M}(\xi; \boldsymbol{\beta}))$ ein Optimalitätskriterium mit isotoner und konkaver Kriteriumsfunktion ϕ . Dann ist Φ konkav auf der Menge der Designs Ξ .

Beweis:

Es seien $\xi_1, \xi_2 \in \Xi$ und $\alpha \in [0, 1]$. Dann gilt:

$$\begin{aligned}
\Phi(\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2) &= \phi(\mathbf{M}(\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2; \boldsymbol{\beta})) \\
&\geq \phi(\alpha \cdot \mathbf{M}(\xi_1; \boldsymbol{\beta}) + (1 - \alpha) \cdot \mathbf{M}(\xi_2; \boldsymbol{\beta})) \\
&\geq \alpha \cdot \phi(\mathbf{M}(\xi_1; \boldsymbol{\beta})) + (1 - \alpha) \cdot \phi(\mathbf{M}(\xi_2; \boldsymbol{\beta})) \\
&= \alpha \cdot \Phi(\xi_1) + (1 - \alpha) \cdot \Phi(\xi_2).
\end{aligned}$$

Dabei wurde bei der ersten Ungleichung Satz 4.6 und die Isotonie von ϕ und bei der zweiten Ungleichung die Konkavität von ϕ verwendet. \square

Die Kriteriumsfunktion $\phi(\cdot) = \ln(\det(\cdot))$ für D -Optimalität ist isoton und konkav. Somit ist $\Phi(\xi) = \ln(\det(\mathbf{M}(\xi; \boldsymbol{\beta}))$ nach Satz 4.8 konkav auf Ξ . Da die Kriteriumsfunktion ϕ außerdem für reguläre Informationsmatrizen differenzierbar ist, kann nach Satz 3.11 ein Äquivalenzsatz für D -Optimalität hergeleitet werden. Es kann auf ein Resultat zum Äquivalenzsatz für Informationsmatrizen von der Form $\mathbf{M}(\xi; \boldsymbol{\beta}) = (\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})^{-1} + \mathbf{D})^{-1}$ von Fedorov und Hackl (1997, S. 78) zurückgegriffen werden. Dabei ist $\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})$ die Informationsmatrix im zugehörigen Modell ohne zufälligen Effekt. Zunächst wird im folgenden Lemma die Fréchet-Ableitung $F_{\Phi_{\boldsymbol{\beta}}}(\xi, \delta_{\mathbf{x}})$ angegeben.

Lemma 4.9

Die Informationsmatrix besitze die Darstellung $\mathbf{M}(\xi; \boldsymbol{\beta}) = (\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})^{-1} + \mathbf{D})^{-1}$, wobei $\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}) = \sum_{j=1}^l w_j \cdot \lambda(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T$ ist. Dann ist die Fréchet-Ableitung $F_{\Phi_{\boldsymbol{\beta}}}(\xi, \delta_{\mathbf{x}})$ gegeben durch:

$$\begin{aligned}
F_{\Phi_{\boldsymbol{\beta}}}(\xi, \delta_{\mathbf{x}}) &= \lambda(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi; \boldsymbol{\beta}) \tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \\
&\quad - \text{Spur}(\mathbf{M}(\xi; \boldsymbol{\beta}) \tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})^{-1})
\end{aligned} \tag{4.7}$$

Mit Satz 3.11 und Lemma 4.9 folgt der Äquivalenzsatz für D -Optimalität für das Poisson-Gamma-Modell (vgl. Fedorov und Hackl, 1997, S. 78).

Satz 4.10 (Äquivalenzsatz)

Ein Design ξ^* ist genau dann D -optimal für das Poisson-Gamma-Modell, wenn

$$\begin{aligned}
&\exp(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) \\
&\leq \text{Spur}(\mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1})
\end{aligned} \tag{4.8}$$

für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

Die standardisierte Maximin D -Optimalität eines Designs kann mit dem folgenden Äquivalenzsatz nachgewiesen werden, welcher aus Satz 3.12 und Lemma 4.9 folgt.

Satz 4.11 (Äquivalenzsatz für standardisierte Maximin Optimalität)

Es sei ξ^* ein Design und

$$\mathcal{N}(\xi^*) = \left\{ \boldsymbol{\beta} \in B : \boldsymbol{\beta} = \arg \min_{\boldsymbol{\beta} \in B} \text{eff}_D(\xi^*; \boldsymbol{\beta}) \right\} \quad (4.9)$$

die Menge der Parametervektoren $\boldsymbol{\beta}$, bei denen die D -Effizienz von ξ^* minimal wird. Das Design ξ^* ist genau dann standardisiert Maximin D -optimal für das Poisson-Gamma-Modell, wenn eine Wahrscheinlichkeitsverteilung π auf $\mathcal{N}(\xi^*)$ existiert, so dass

$$\int_{\mathcal{N}(\xi^*)} \left(\exp(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x})^T \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{f}(\mathbf{x}) - \text{Spur}(\mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1}) \right) \pi(d\boldsymbol{\beta}) \leq 0 \quad (4.10)$$

für alle $\mathbf{x} \in \mathcal{X}$ gilt. An den Stützpunkten von ξ^* gilt Gleichheit.

5 Lokal optimale Designs

In diesem Kapitel werden lokal optimale Designs für verschiedene Optimalitätskriterien für das Poisson-Gamma-Modell mit bekannten Parametern a und b der Gamma-Verteilung hergeleitet. Da das Poisson-Gamma-Modell ein nichtlineares Modell ist, hängen die optimalen Designs vom Parametervektor β ab. Dennoch sind sie auch von praktischem Interesse, beispielsweise wenn eine Vermutung über den wahren Wert des Parametervektors vorliegt oder zur Berechnung von sequentiellen oder standardisiert Maximin optimalen Designs (vgl. Ford et al., 1992).

Nach Folgerung 4.7 kann für alle betrachteten Optimalitätskriterien unter Verwendung eines optimalen individuellen Designs ξ^* ein optimales Populationsdesign $\zeta^* = \{\xi^*\}_1$ erhalten werden. Daher wird sich im Folgenden auf individuelle Designs beschränkt. Zunächst werden D - und D_s -Optimalität und anschließend werden L - und c -Optimalität, welche lineare Optimalitätskriterien sind, betrachtet. Dabei werden nicht nur optimale Designs hergeleitet, sondern auch Zusammenhänge zu Optimalitätskriterien für das Poisson-Modell hergestellt (vgl. Schmidt und Schwabe, 2020).

5.1 D -optimale Designs

Beim D -Optimalitätskriterium wird die Determinante der Informationsmatrix bezüglich des Designs maximiert. Aufbauend auf der Berechnung der Kriteriumsfunktion für D -Optimalität beim Poisson-Gamma-Modell werden Zusammenhänge zu D - und D_s -Optimalität beim Poisson-Modell für ein allgemeines Regressionsmodell hergestellt. Weiterhin erfolgt die Herleitung D -optimaler Gewichte für Designs mit minimalem Träger und die Berechnung D -optimaler Designs für das Regressionsmodell mit einer Kovariablen und für das multiple Regressionsmodell. Es wird gezeigt, dass die D_s -optimalen Designs für das Poisson-Gamma-Modell mit denen für das Poisson-Modell übereinstimmen. Außerdem wird in Beispielen die Effizienz der optimalen Designs für das Poisson- und Poisson-Gamma-Modell untereinander verglichen.

5.1.1 Allgemeines Regressionsmodell

Es wird das Regressionsmodell mit $\mathbf{f} = (1, f_1, \dots, f_{p-1})^T$ betrachtet, wobei für die Regressionsfunktionen f_i nur die Stetigkeit vorausgesetzt wird. Sie können beispielsweise als $f_i(x) = x^i$ für polynomiale Regression mit eindimensionaler Kovariable $x \in \mathbb{R}$ oder $f_i(\mathbf{x}) = x_i$ für multiple Regression mit $\mathbf{x} \in \mathbb{R}^{p-1}$ gewählt werden. Für $p = 2$ und $f_1(x) = x$ ergibt sich als Spezialfall die Regression mit einer Kovariablen $x \in \mathbb{R}$. Mit Lemma 4.1 lässt sich die Determinante der Informationsmatrix berechnen, die im folgenden Satz angegeben wird.

Satz 5.1

Die Kriteriumsfunktion für D -Optimalität für das Poisson-Gamma-Modell ist gegeben durch:

$$\det(\mathbf{M}(\xi; \boldsymbol{\beta})) = \left(\frac{a}{b}\right)^p \cdot \frac{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}{1 + \frac{m}{b} \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p}}. \quad (5.1)$$

Beweis:

Falls $\mathbf{M}(\xi; \boldsymbol{\beta})$ singular ist, dann ist $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ ebenfalls singular und somit gilt Gleichung (5.1).

Es sei $\mathbf{M}(\xi; \boldsymbol{\beta})$ regulär. Dann ist $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ nach Lemma 4.2 ebenfalls regulär. Mit der regulären Matrix $\mathbf{L}(\xi; \boldsymbol{\beta})$ aus Lemma 4.1 gilt $\mathbf{M}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{L}(\xi; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$. Mit der Determinante von $\mathbf{L}(\xi; \boldsymbol{\beta})$ in Satz 4.1 (ii) folgt

$$\begin{aligned} \det(\mathbf{M}(\xi; \boldsymbol{\beta})) &= \left(\frac{a}{b}\right)^p \cdot \det(\mathbf{L}(\xi; \boldsymbol{\beta})) \cdot \det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})) \\ &= \left(\frac{a}{b}\right)^p \cdot \frac{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}{1 + \frac{m}{b} \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p}} \end{aligned} \quad (5.2)$$

und somit die Darstellung der Kriteriumsfunktion. □

Unter Verwendung der Darstellungen der Kriteriumsfunktionen in Satz 5.1 für D -Optimalität und in Satz 3.5 für D_s -Optimalität lässt sich ein Zusammenhang zwischen D -Optimalität für das Poisson-Gamma-Modell und D - und D_s -Optimalität für das Poisson-Modell herstellen. Die Maximierung der Determinante der Informationsmatrix für das Poisson-Gamma-Modell ist äquivalent zur Minimierung der inversen Determinante, welche durch

$$\det(\mathbf{M}(\xi; \boldsymbol{\beta}))^{-1} = \left(\frac{b}{a}\right)^p \cdot \frac{1 + \frac{m}{b} \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p}}{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}$$

$$= \left(\frac{b}{a}\right)^p \cdot \left(\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))^{-1} + \frac{m}{b} \cdot \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p}}{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))} \right) \quad (5.3)$$

gegeben ist.

Wird beim Poisson-Modell D_s -Optimalität für die Parameter $\beta_1, \dots, \beta_{p-1}$ betrachtet, also $\mathbf{A}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1})$, dann gilt mit der entsprechenden Darstellung der Informationsmatrix $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ als 2×2 -Blockmatrix wie in Gleichung (3.7) für die obere linke Teilmatrix

$$M_{11} = \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} = \sum_{j=1}^l w_j \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}) > 0 \quad (5.4)$$

für alle $\xi \in \Xi$. Die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ ist somit nach Satz 3.4 für ein D_s -optimales Design regulär und mit Satz 3.5 folgt für die Kriteriumsfunktion:

$$\det(\mathbf{A}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}) = \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p}}{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}. \quad (5.5)$$

Daher gilt mit $\mathbf{A}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1})$ und Gleichung (5.5) der folgende Zusammenhang:

$$\arg \max_{\xi \in \Xi} \{ \det(\mathbf{M}(\xi; \boldsymbol{\beta})) \} = \arg \min_{\xi \in \Xi} \left\{ \det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))^{-1} + \frac{m}{b} \cdot \det(\mathbf{A}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}) \right\}.$$

Somit ist das D -Optimalitätskriterium für das Poisson-Gamma-Modell äquivalent zu einem kombinierten gewichteten Optimalitätskriterium aus D -Optimalität und D_s -Optimalität für $\beta_1, \dots, \beta_{p-1}$ für das Poisson-Modell. Dabei hängt die Gewichtung der beiden Optimalitätskriterien für das Poisson-Modell vom Parameter b der Gamma-Verteilung ab. Je größer b ist, desto stärker liegt das Gewicht auf D -Optimalität und desto kleiner wird der Einfluss von D_s -Optimalität.

Die D -optimalen Gewichte sind diejenigen Gewichte, welche für ein Design mit gegebenen Stützpunkten die Kriteriumsfunktion maximieren. Bei Modellen mit unabhängigen Beobachtungen, wie etwa beim Poisson-Modell, gilt nach Satz 3.10 a), dass ein D -optimales Design mit minimalem Träger, also genau p Stützpunkten, jedem Stützpunkt das gleiche Gewicht $1/p$ zuordnet. Der folgende Satz gibt die D -optimalen Gewichte für das Poisson-Gamma-Modell an, indem die D -optimalen Gewichte in Verbindung zu den D_s -optimalen Gewichten in Satz 3.10 b) für ein Modell mit unabhängigen Beobachtungen gebracht werden.

Satz 5.2

Es seien $\mathbf{x}_1, \dots, \mathbf{x}_p$ Stützpunkte eines p -Punkt-Designs ξ , wobei $\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_p)$ linear unabhängig seien. Die Gewichte w_1^*, \dots, w_p^* sind genau dann D -optimal für $\mathbf{x}_1, \dots, \mathbf{x}_p$ für das Poisson-Gamma-Modell, wenn sie D_s -optimal (für $\beta_1, \dots, \beta_{p-1}$) für $\mathbf{x}_1, \dots, \mathbf{x}_p$ für ein Modell mit Informationsmatrix $\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}) = \sum_{j=1}^p w_j \cdot \lambda(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}) \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T$ mit $\lambda(z) = 1 + \frac{m}{b} \cdot e^z$ sind.

Beweis:

Für das Design ξ mit Stützpunkten $\mathbf{x}_1, \dots, \mathbf{x}_p$ berechnet sich die Determinante der Informationsmatrix für das Poisson-Gamma-Modell nach Satz 5.1 zu:

$$\det(\mathbf{M}(\xi; \boldsymbol{\beta})) = \left(\frac{a}{b}\right)^p \cdot \frac{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}{1 + \frac{m}{b} \cdot \mathbf{e}_1^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_1}.$$

Dabei ist die Informationsmatrix für das Poisson-Modell durch $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \mathbf{X}$ mit $\mathbf{X} = (\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_p))^T$ und den Diagonalmatrizen $\mathbf{W} = \text{Diag}(w_1, \dots, w_p)$ und $\boldsymbol{\Lambda} = \text{Diag}(\exp(\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta}), \dots, \exp(\mathbf{f}(\mathbf{x}_p)^T \boldsymbol{\beta}))$ gegeben.

Mit $\tilde{\boldsymbol{\Lambda}} = \text{Diag}(1 + \frac{m}{b} \cdot \exp(\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta}), \dots, 1 + \frac{m}{b} \cdot \exp(\mathbf{f}(\mathbf{x}_p)^T \boldsymbol{\beta}))$ folgt:

$$\begin{aligned} \det(\mathbf{M}(\xi; \boldsymbol{\beta})) &= \left(\frac{a}{b}\right)^p \cdot \frac{\det(\mathbf{W}) \cdot \det(\boldsymbol{\Lambda}) \cdot \det(\mathbf{X})^2}{1 + \frac{m}{b} \cdot \sum_{j=1}^p w_j \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \\ &= \left(\frac{a}{b}\right)^p \cdot \frac{\det(\boldsymbol{\Lambda})}{\det(\tilde{\boldsymbol{\Lambda}})} \cdot \frac{\det(\mathbf{W}) \cdot \det(\tilde{\boldsymbol{\Lambda}}) \cdot \det(\mathbf{X})^2}{\sum_{j=1}^p w_j \cdot (1 + \frac{m}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})} \\ &= \left(\frac{a}{b}\right)^p \cdot \frac{\det(\boldsymbol{\Lambda})}{\det(\tilde{\boldsymbol{\Lambda}})} \cdot \frac{\det(\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}))}{\mathbf{e}_1^T \tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}) \mathbf{e}_1}. \end{aligned}$$

Da $\det(\boldsymbol{\Lambda})/\det(\tilde{\boldsymbol{\Lambda}})$ nicht von den Gewichten abhängt und $\det(\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})) / (\mathbf{e}_1^T \tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}) \mathbf{e}_1)$ gleich der Kriteriumsfunktion für D_s -Optimalität für $\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta})$ ist, folgt die Gleichheit der optimalen Gewichte. \square

Da im Allgemeinen die D_s -optimalen Gewichte auch bei einem Design mit minimalem Träger nicht alle gleich sind, folgt aus Satz 5.2, dass beim Poisson-Gamma-Modell die D -optimalen Gewichte für ein Design mit minimalem Träger ebenfalls nicht alle gleich sind.

5.1.2 Regressionsmodell mit einer Kovariablen

Es wird in diesem Abschnitt der Fall des Regressionsmodells mit einer Kovariablen, also $p = 2$, betrachtet. Die Regressionsfunktion sei $\mathbf{f}(x) = (1, x)^T$ und der Parametervektor sei $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$. Der lineare Prädiktor ist also durch $\mathbf{f}(x)^T \boldsymbol{\beta} = \beta_0 + \beta_1 x$ gegeben. Für den Fall einer Kovariablen stimmen c -Optimalität mit $\mathbf{c} = (0, 1)^T$ und D_s -Optimalität für die Steigung β_1 überein. Somit steht D -Optimalität für das Poisson-Gamma-Modell im Zusammenhang zu D -Optimalität und c -Optimalität mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell.

Die D -optimalen Gewichte für das Poisson-Gamma-Modell aus Satz 5.2 lassen sich für den Fall einer Kovariablen vereinfachen. Dies führt zu folgendem Satz, dessen Beweis im Anhang angegeben ist.

Satz 5.3

Es seien x_1 und x_2 zwei verschiedene Stützpunkte eines 2-Punkt-Designs ξ . Dann sind die D -optimalen Gewichte für x_1 und x_2 gegeben durch:

$$w_1^* = \frac{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 x_2)}}{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 x_1)} + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 x_2)}}, \quad (5.6)$$

$$w_2^* = \frac{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 x_1)}}{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 x_1)} + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 x_2)}}. \quad (5.7)$$

Mit den D -optimalen Gewichten in Satz 5.3 ergibt sich direkt das D -optimale Design für einen binären Designraum $\mathcal{X} = \{0, 1\}$.

Satz 5.4

Es sei $\mathcal{X} = \{0, 1\}$ der Designraum. Das D -optimale Design ξ^* ist gegeben durch:

$$\xi^* = \left\{ \begin{array}{cc} 0 & 1 \\ \frac{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1)}}{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0)} + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1)}} & \frac{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0)}}{\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0)} + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1)}} \end{array} \right\}. \quad (5.8)$$

Nun wird ein stetiger Designraum $\mathcal{X} = [u, v]$ betrachtet. Auch bei einem stetigen Designraum ist immer ein 2-Punkt-Design D -optimal, wie das folgende Lemma zeigt.

Lemma 5.5

Es sei $\mathcal{X} = [u, v]$ mit $u < v$ der Designraum. Ein D -optimales Design ξ^* besitzt immer genau zwei verschiedene Stützpunkte. Für $\beta_1 > 0$ ist ein Stützpunkt bei v und für $\beta_1 < 0$ ist ein Stützpunkt bei u . Für $\beta_1 = 0$ sind u und v die Stützpunkte.

Beweis:

Mit dem Äquivalenzsatz 4.10 folgt, dass ein Design ξ^* genau dann D -optimal ist, wenn

$$\begin{aligned} & \exp(\beta_0 + \beta_1 x) \cdot \mathbf{f}(x)^T \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{f}(x) \\ & \leq \text{Spur}(\mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1}) \end{aligned}$$

für alle $x \in [u, v]$ gilt. Dabei gilt Gleichheit in der Ungleichung an den Stützpunkten des optimalen Designs ξ^* . Es seien $g(x) = \mathbf{f}(x)^T \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{f}(x)$ und $h(x) = \exp(\beta_0 + \beta_1 x)$. Dann ist die Bedingung des Äquivalenzsatzes äquivalent zu $k(x) = h(x) \cdot g(x) \leq \text{Spur}(\mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1})$ für alle $x \in [u, v]$. Es gilt:

$$\begin{aligned} k'(x) &= h'(x) \cdot g(x) + h(x) \cdot g'(x) = \beta_1 \cdot h(x) \cdot g(x) + g'(x) \cdot h(x) \\ &= h(x) \cdot (\beta_1 \cdot g(x) + g'(x)). \end{aligned}$$

Die Funktion g ist ein quadratisches Polynom und $g'(x)$ ist somit eine Gerade. Es gilt $h(x) > 0$ für alle $x \in \mathbb{R}$ und $\beta_1 \cdot g(x) = -g'(x)$ kann nur für maximal zwei $x \in \mathbb{R}$ gelten. Also besitzt $k'(x)$ maximal zwei Nullstellen und $k(x)$ hat somit maximal zwei Extrema für $x \in \mathbb{R}$. Dann kann $k(x)$ höchstens zwei Extrema im Inneren des Intervalls $[u, v]$ besitzen. Im Fall zweier Extrema muss ein Extremum ein Maximum und das andere Extremum ein Minimum sein. Somit besitzt das D -optimale Design maximal zwei Stützpunkte. Da zur Schätzbarkeit des Parametervektors $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ mindestens zwei Stützpunkte benötigt werden, besteht das D -optimale Design aus genau zwei Stützpunkten.

Da $\mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1}$ positiv definit ist, ist $g(x)$ konvex und es gilt $\lim_{x \rightarrow \infty} g(x) = \infty$. Für $\beta_1 \geq 0$ gilt $\lim_{x \rightarrow \infty} k(x) = \infty$, somit muss v ein Stützpunkt sein. Analog folgt für $\beta_1 \leq 0$, dass u ein Stützpunkt sein muss. Insbesondere folgt für $\beta_1 = 0$, dass ξ^* die Stützpunkte u und v besitzt. \square

Unter Verwendung von Lemma 5.5 für die D -optimalen Stützpunkte und Satz 5.3 für die D -optimalen Gewichte wird im folgenden Satz das D -optimale Design für einen stetigen Designraum hergeleitet.

Satz 5.6

Es sei $\mathcal{X} = [u, v]$ der Designraum.

a) Es sei $\beta_1 \neq 0$. Für $\beta_1 > 0$ sei $d = v$ und für $\beta_1 < 0$ sei $d = u$. Die Gleichung

$$0 = z - 2 \cdot \left(1 + \frac{\frac{m}{b} \cdot \exp(\beta_0 + \beta_1 d - z)}{1 + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 d - z)} \cdot \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 d)}} \right) \quad (5.9)$$

besitzt eine eindeutige Lösung z^* . Für $\beta_1 > 0$ sei $x^* = \max\{u, v - z^*/\beta_1\}$ und für $\beta_1 < 0$ sei $x^* = \min\{u - z^*/\beta_1, v\}$. Dann ist das D -optimale Design ξ^* gegeben durch:

$$\xi^* = \left\{ \begin{array}{c} d \\ \frac{\sqrt{1+\frac{m}{b} \cdot \exp(\beta_0+\beta_1 x^*)}}{\sqrt{1+\frac{m}{b} \cdot \exp(\beta_0+\beta_1 d)} + \sqrt{1+\frac{m}{b} \cdot \exp(\beta_0+\beta_1 x^*)}} \end{array} \quad \begin{array}{c} x^* \\ \frac{\sqrt{1+\frac{m}{b} \cdot \exp(\beta_0+\beta_1 d)}}{\sqrt{1+\frac{m}{b} \cdot \exp(\beta_0+\beta_1 d)} + \sqrt{1+\frac{m}{b} \cdot \exp(\beta_0+\beta_1 x^*)}} \end{array} \right\}. \quad (5.10)$$

b) Für $\beta_1 = 0$ ist das D -optimale Design ξ^* gegeben durch:

$$\xi^* = \left\{ \begin{array}{cc} u & v \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}. \quad (5.11)$$

Beweis:

Da nach Lemma 5.5 das D -optimale Design ξ^* zwei Stützpunkte besitzt, wird ein Design mit Stützpunkten x_1 und x_2 , wobei $x_1 < x_2$ gelte, sowie zugehörigen Gewichten w_1 und w_2 betrachtet. Es sei $\lambda_i = \exp(\beta_0 + \beta_1 x_i)$ und $\delta_i = 1 + \frac{m}{b} \cdot \lambda_i$ für $i = 1, 2$. Mit der Matrixzerlegung der Informationsmatrix $\mathbf{M}_{\text{Po}}(\xi; \beta) = \mathbf{X}^T \mathbf{W} \mathbf{\Lambda} \mathbf{X}$ ist die Determinante der Informationsmatrix nach Satz 5.1 durch

$$\begin{aligned} \det(\mathbf{M}(\xi; \beta)) &= \left(\frac{a}{b}\right)^2 \cdot \frac{\det(\mathbf{M}_{\text{Po}}(\xi; \beta))}{1 + \frac{m}{b} \cdot \mathbf{e}_1^T \mathbf{M}_{\text{Po}}(\xi; \beta) \mathbf{e}_1} = \left(\frac{a}{b}\right)^2 \cdot \frac{\det(\mathbf{W}) \cdot \det(\mathbf{X})^2 \cdot \det(\mathbf{\Lambda})}{1 + \frac{m}{b} \cdot \mathbf{e}_1^T \mathbf{M}_{\text{Po}}(\xi; \beta) \mathbf{e}_1} \\ &= \left(\frac{a}{b}\right)^2 \cdot \frac{w_1 \cdot w_2 \cdot \lambda_1 \cdot \lambda_2 \cdot (x_2 - x_1)^2}{1 + \frac{m}{b} \cdot (w_1 \cdot \lambda_1 + w_2 \cdot \lambda_2)} \\ &= \left(\frac{a}{b}\right)^2 \cdot \frac{w_1 \cdot w_2 \cdot \lambda_1 \cdot \lambda_2 \cdot (x_2 - x_1)^2}{w_1 \cdot \delta_1 + w_2 \cdot \delta_2} \end{aligned} \quad (5.12)$$

gegeben. Einsetzen der D -optimalen Gewichte aus Satz 5.3 ergibt:

$$\begin{aligned} \det(\mathbf{M}(\xi; \beta)) &= \left(\frac{a}{b}\right)^2 \cdot \frac{\frac{\sqrt{\delta_2}}{\sqrt{\delta_1+\sqrt{\delta_2}}} \cdot \frac{\sqrt{\delta_1}}{\sqrt{\delta_1+\sqrt{\delta_2}}} \cdot \lambda_1 \cdot \lambda_2 \cdot (x_2 - x_1)^2}{\frac{\sqrt{\delta_2}}{\sqrt{\delta_1+\sqrt{\delta_2}}} \cdot \delta_1 + \frac{\sqrt{\delta_1}}{\sqrt{\delta_1+\sqrt{\delta_2}}} \cdot \delta_2} \\ &= \left(\frac{a}{b}\right)^2 \cdot \frac{\lambda_1 \cdot \lambda_2 \cdot (x_2 - x_1)^2}{(\sqrt{\delta_1} + \sqrt{\delta_2})^2}. \end{aligned} \quad (5.13)$$

Es wird der Fall $\beta_1 > 0$ betrachtet. Nach Lemma 5.5 ist ein Stützpunkt bei $x_2 = v$. Durch Weglassen positiver Konstanten bezüglich x_1 ergibt sich als vereinfachte Kriteriumsfunktion:

$$k(x_1) = \frac{\lambda_1 \cdot (v - x_1)^2}{(\sqrt{\delta_1} + \sqrt{\delta_2})^2}.$$

Es gilt $d\lambda_1/dx_1 = \beta_1 \cdot \lambda_1$, $d\delta_1/dx_1 = \frac{m}{b} \cdot \beta_1 \cdot \lambda_1$ und $d\delta_2/dx_1 = 0$. Die Ableitung von k nach x_1 berechnet sich zu:

$$\frac{dk(x_1)}{dx_1} = \frac{(v - x_1) \cdot \lambda_1 \cdot \left[(\beta_1 \cdot (v - x_1) - 2) \cdot (\sqrt{\delta_1} + \sqrt{\delta_2}) - \frac{(v - x_1) \cdot \beta_1 \cdot \frac{m}{b} \cdot \lambda_1}{\sqrt{\delta_1}} \right]}{(\sqrt{\delta_1} + \sqrt{\delta_2})^3}.$$

Um die Kriteriumsfunktion zu maximieren, wird die Ableitung gleich Null gesetzt. Sie ist für $x_1 < v$ genau dann gleich Null, wenn der Term in eckigen Klammern gleich Null ist. Ausklammern von $\beta_1 \cdot (v - x_1)$ und Multiplikation mit $\sqrt{\delta_1}$ ergibt:

$$\begin{aligned} 0 &= \beta_1 \cdot (v - x_1) \cdot \left(\delta_1 + \sqrt{\delta_1} \cdot \sqrt{\delta_2} - \frac{m}{b} \cdot \lambda_1 \right) - 2 \cdot \left(\delta_1 + \sqrt{\delta_1} \cdot \sqrt{\delta_2} \right) \\ &= \beta_1 \cdot (v - x_1) \cdot \left(1 + \sqrt{\delta_1} \cdot \sqrt{\delta_2} \right) - 2 \cdot \left(\frac{m}{b} \cdot \lambda_1 + 1 + \sqrt{\delta_1} \cdot \sqrt{\delta_2} \right). \end{aligned}$$

Äquivalent gilt:

$$\begin{aligned} 0 &= \beta_1 \cdot (v - x_1) - 2 \cdot \left(1 + \frac{\frac{m}{b} \cdot \lambda_1}{1 + \sqrt{\delta_1} \cdot \sqrt{\delta_2}} \right) \\ &= \beta_1 \cdot (v - x_1) - 2 \cdot \left(1 + \frac{\frac{m}{b}}{e^{-\beta_0 - \beta_1 x_1} + \sqrt{e^{-\beta_0 - \beta_1 x_1} + \frac{m}{b}} \cdot \sqrt{e^{-\beta_0 - \beta_1 x_1} + \frac{m}{b}} \cdot e^{\beta_1 \cdot (v - x_1)}}} \right). \end{aligned}$$

Mit der Transformation $x_1 \rightarrow z = \beta_1 \cdot (v - x_1)$ ergibt sich die transformierte Gleichung:

$$0 = z - 2 \cdot \left(1 + \frac{\frac{m}{b} \cdot e^{\beta_0 + \beta_1 v - z}}{1 + \sqrt{1 + \frac{m}{b}} \cdot e^{\beta_0 + \beta_1 v - z}} \cdot \sqrt{1 + \frac{m}{b}} \cdot e^{\beta_0 + \beta_1 v}} \right)$$

Die rechte Seite der Gleichung ist streng monoton wachsend in z und hat somit höchstens eine Lösung z^* . Für $z \rightarrow \infty$ konvergiert die rechte Seite der Gleichung gegen unendlich und für $z = 0$ ist die rechte Seite kleiner als null. Somit existiert nach dem Zwischenwertsatz eine Lösung z^* im Intervall $(0, \infty)$. Folglich ist $x^* = v - z^*/\beta_1$ die eindeutige Nullstelle von $dk(x_1)/dx_1$. Für $x_1 \rightarrow v$ und $x_1 \rightarrow -\infty$ konvergiert die Kriteriumsfunktion $\det(\mathbf{M}(\xi; \boldsymbol{\beta}))$ gegen Null. Da zudem $\det(\mathbf{M}(\xi; \boldsymbol{\beta})) > 0$ gilt, handelt es sich bei x^* um ein Maximum. Ist $x^* \geq u$, so ist $x_1^* = x^*$. Falls $x^* < u$ gilt, dann ist die Kriteriumsfunktion auf dem Intervall $[u, v]$ streng monoton fallend und somit für $x_1^* = u$ maximal. Für $\beta_1 < 0$ folgt das D -optimale Design analog. Für $\beta_1 = 0$ ergibt sich das D -optimale Design direkt aus Lemma 5.5 und Satz 5.3. \square

Das D -optimale Design ist immer ein 2-Punkt-Design, wobei ein Stützpunkt der Randpunkt des Designraums ist, an dem $\exp(\beta_0 + \beta_1 x)$ maximal wird. Zur Bestimmung des

zweiten Stützpunkts muss Gleichung (5.9) numerisch gelöst werden. Der Stützpunkt kann sowohl innerhalb als auch auf dem anderen Randpunkt des Designraums liegen.

Als Spezialfälle des D -optimalen Designs für das Poisson-Gamma-Modell ergeben sich für $b \rightarrow \infty$ das D -optimale Design und für $b \rightarrow 0$ das c -optimale Design mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell. Dabei vereinfacht sich Gleichung (5.9) so, dass sie analytisch lösbar ist. Diese optimalen Designs für das Poisson-Modell sind bereits bekannt (vgl. Ford et al. (1992), Rodríguez-Torreblanca und Rodríguez-Díaz (2007)). Für den Designraum $\mathcal{X} = [u, v]$ werden sie im Folgenden angegeben.

Für $\beta_1 < 0$ sei $x^* = \min \{u - 2/\beta_1, v\}$. Dann ist das D -optimale Design für das Poisson-Modell durch

$$\xi^* = \left\{ \begin{array}{cc} u & x^* \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\} \quad (5.14)$$

gegeben. Für $\beta_1 > 0$ besitzt das D -optimale Design für das Poisson-Modell die Stützpunkte v und $x^* = \max \{u, v - 2/\beta_1\}$ mit gleichen Gewichten $1/2$.

Für $\beta_1 < 0$ sei $x^* = \min \{u - 2 \cdot (1 + W(e^{-1}))/\beta_1, v\}$. Dabei ist W der Hauptzweig der Lambertschen W -Funktion, welcher die Umkehrfunktion von $w \mapsto w \cdot e^w$ für $w \geq -1$ ist (vgl. Corless et al., 1996). Weiterhin gilt $2 \cdot (1 + W(e^{-1})) \approx 2.557$. Dann ist das c -optimale Design mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell durch

$$\xi^* = \left\{ \begin{array}{cc} \frac{u}{\frac{\sqrt{\exp(\beta_1 x^*)}}{\sqrt{\exp(\beta_1 u) + \sqrt{\exp(\beta_1 x^*)}}}} & \frac{x^*}{\frac{\sqrt{\exp(\beta_1 u)}}{\sqrt{\exp(\beta_1 u) + \sqrt{\exp(\beta_1 x^*)}}}} \end{array} \right\} \quad (5.15)$$

gegeben. Für $\beta_1 > 0$ besitzt das c -optimale Design mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell die Stützpunkte v und $x^* = \max \{u, v - 2 \cdot (1 + W(e^{-1}))/\beta_1\}$.

Für $\beta_1 = 0$ besitzen das D -optimale Design und das c -optimale Design mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell die Stützpunkte u und v mit gleichen Gewichten $1/2$.

Im Gegensatz zum Poisson-Gamma-Modell hängen die optimalen Designs beim Poisson-Modell nicht vom Parameter β_0 ab, da er als Faktor $\exp(\beta_0)$ aus der Informationsmatrix herausgezogen werden kann.

Im folgenden Beispiel wird das D -optimale Design im Poisson-Gamma-Modell berechnet. Außerdem werden das D -optimale Design und das c -optimale Design für β_1 im Poisson-Modell betrachtet.

Beispiel 5.7

Der Designraum sei $\mathcal{X} = [0, 3]$. Die Regressionsfunktion sei $\mathbf{f}(x) = (1, x)^T$ und der Parametervektor sei durch $\boldsymbol{\beta} = (0, -1)^T$ gegeben. Weiterhin seien $m = 10$ und $b = 1$.

Modell	Kriterium	Optimales Design	Effizienzen		
			Po D	P-G D	Po c
Poisson	D	$\begin{Bmatrix} 0 & 2 \\ 1/2 & 1/2 \end{Bmatrix}$	1	0.925	0.769
Poisson-Gamma	D	$\begin{Bmatrix} 0 & 2.341 \\ 0.297 & 0.703 \end{Bmatrix}$	0.902	1	0.974
Poisson	c	$\begin{Bmatrix} 0 & 2.557 \\ 0.218 & 0.782 \end{Bmatrix}$	0.799	0.981	1

Tabelle 5.1: Optimale Designs in Beispiel 5.7 und Vergleich der Effizienzen für das Poisson-Gamma-Modell (P-G) und das Poisson-Modell (Po) im Fall einer Kovariablen

Das D -optimale Design für das Poisson-Gamma-Modell lässt sich mit Satz 5.6 berechnen. Die Lösung von Gleichung (5.9) berechnet sich zu $z^* = 2.341$. In Tabelle 5.1 ist das D -optimale Design für das Poisson-Gamma-Modell angegeben. Zudem wurden für das Poisson-Modell das D -optimale Design aus Gleichung (5.14) und das c -optimale Design für β_1 aus Gleichung (5.15) gefolgert. Weiterhin werden die optimalen Designs bezüglich ihrer Effizienzen in der Tabelle verglichen.

Die drei optimalen Designs sind von der gleichen Struktur. Sie besitzen zwei Stützpunkte, wobei ein Stützpunkt bei 0 ist. Dies ist der Punkt des Designraums, an dem $\exp(\beta_0 + \beta_1 x)$ maximal wird. Der andere Stützpunkt des D -optimalen Designs für das Poisson-Gamma-Modell liegt zwischen den Stützpunkten der D - und c -optimalen Designs für das Poisson-Modell. Beim c -optimalen Design im Poisson-Modell ist der Abstand des Stützpunkts zum Stützpunkt 0 größer als beim D -optimalen Design im Poisson-Modell.

Im Vergleich zum D -optimalen Design im Poisson-Modell besitzt das c -optimale Design in diesem Beispiel eine höhere Effizienz bezüglich des D -optimalen Designs im Poisson-Gamma-Modell. Die Effizienzen sind relativ hoch, es tritt jedoch ein Effizienzverlust auf. Werden für ein Design für das Poisson-Gamma-Modell die D -optimalen Stützpunkte 0 und 2 für das Poisson-Modell verwendet und nur die Gewichte optimiert, so ergeben sich die optimalen Gewichte nach Satz 5.3 zu 0.316 und 0.684. Die D -Effizienz des Designs im Vergleich zum D -optimalen Design im Poisson-Gamma-Modell beträgt 0.985. Es lässt sich somit ein deutlicher Effizienzgewinn durch Optimierung der Gewichte erzielen.

In den folgenden Abschnitten werden für verschiedene Optimalitätskriterien weitere Resultate zu den Beziehungen zwischen den optimalen Designs für das Poisson-Gamma- und Poisson-Modell hergeleitet, welche auch auf das Regressionsmodell mit einer Kovariablen anwendbar sind.

5.1.3 Multiples Regressionsmodell

Im Folgenden erfolgt die Betrachtung des multiplen Regressionsmodells mit der Regressionsfunktion $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$, wobei $\mathbf{x} = (x_1, \dots, x_{p-1})^T \in \mathbb{R}^{p-1}$ ist, und dem Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^T$.

Um auch für das multiple Regressionsmodell D -optimale Designs zu bestimmen, werden zunächst die Gewichte betrachtet. Für ein Design mit minimalem Träger liefert der nächste Satz die D -optimalen Gewichte, wobei der Beweis sich im Anhang befindet (vgl. Schmidt und Schwabe, 2020).

Satz 5.8

Es seien $\mathbf{x}_1, \dots, \mathbf{x}_p$ affin unabhängige Stützpunkte eines Designs ξ , wobei $\mathbf{x}_1, \dots, \mathbf{x}_{p-1}$ auf einer Hyperebene $\{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = c\}$ liegen. Für \mathbf{x}_p gelte die Ungleichung $\mathbf{f}(\mathbf{x}_p)^T \boldsymbol{\beta} > c$. Dann ist das D -optimale Gewicht für \mathbf{x}_p gegeben durch:

$$w_p^* = \frac{2}{p + \sqrt{(p-2)^2 + 4 \cdot (p-1) \cdot \frac{1 + \frac{m}{b} \cdot \exp(\mathbf{f}(\mathbf{x}_p)^T \boldsymbol{\beta})}{1 + \frac{m}{b} \cdot \exp(c)}}}. \quad (5.16)$$

Die D -optimalen Gewichte für $\mathbf{x}_1, \dots, \mathbf{x}_{p-1}$ sind durch $w_1^* = \dots = w_{p-1}^* = (1 - w_p^*) / (p-1)$ gegeben.

Die D -optimalen Gewichte erfüllen die Ungleichung:

$$0 < w_p^* < \frac{1}{p} < w_1^* = \dots = w_{p-1}^* < \frac{1}{p-1}. \quad (5.17)$$

Um das D -optimale Design zu charakterisieren, ist die folgende Bemerkung hilfreich.

Bemerkung 5.9

Es wird der rechteckige Designraum $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und $\beta_i \neq 0$ für $i = 1, \dots, p-1$ betrachtet. Es sei $\mathbf{d} = (d_1, \dots, d_{p-1})$ mit $d_i = v_i$ für $\beta_i > 0$ und $d_i = u_i$ für $\beta_i < 0$. Beispielsweise ist $\mathbf{d} = (v_1, \dots, v_{p-1})$, falls $\beta_i > 0$ für $i = 1, \dots, p-1$ gilt. Weiterhin wird mit $\mathbf{e}_{i,p-1} \in \mathbb{R}^{p-1}$ der i -te Standard-Einheitsvektor bezeichnet und es sei $z > 0$. Ein Design mit einem Stützpunkt auf der Ecke \mathbf{d} von \mathcal{X} und weiteren Stützpunkten $\mathbf{d} - (z/\beta_1) \cdot \mathbf{e}_{1,p-1}, \dots, \mathbf{d} - (z/\beta_{p-1}) \cdot \mathbf{e}_{p-1,p-1}$ auf den Kanten erfüllt die Bedingungen von Satz 5.8 mit $c = \mathbf{f}(\mathbf{d})^T \boldsymbol{\beta} - z$. Somit hängen die D -optimalen Gewichte $w_1^*(z), \dots, w_p^*(z)$ aus Satz 5.8 für dieses Design von dem Distanzfaktor z zur Ecke \mathbf{d} ab. Dabei muss $z \leq \min_{i=1, \dots, p-1} (|\beta_i| \cdot (v_i - u_i))$ gelten, damit alle Stützpunkte im Designraum \mathcal{X} liegen.

Mit den D -optimalen Gewichten in Satz 5.8 und Bemerkung 5.9 lässt sich das D -optimale Design im nächsten Satz angeben (vgl. Schmidt und Schwabe, 2020).

Satz 5.10

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ der Designraum und $\beta_i \neq 0$ für $i = 1, \dots, p-1$. Sei $\mathbf{d} = (d_1, \dots, d_{p-1})$ mit $d_i = v_i$ für $\beta_i > 0$ und $d_i = u_i$ für $\beta_i < 0$. Für beliebiges $z > 0$ seien $w_1^*(z) = \dots = w_{p-1}^*(z)$ und $w_p^*(z)$ die D -optimalen Gewichte aus Bemerkung 5.9. Die Gleichung

$$0 = m \cdot \left((p-1) \cdot w_1^*(z) \cdot e^{\mathbf{f}(\mathbf{d})^T \boldsymbol{\beta} - z} + w_p^*(z) \cdot e^{\mathbf{f}(\mathbf{d})^T \boldsymbol{\beta}} \right) \cdot (z \cdot (p-1) \cdot w_1^*(z) - 2) + b \cdot (z \cdot p \cdot w_1^*(z) - 2) \quad (5.18)$$

besitzt eine eindeutige Lösung z^* in $(0, \infty)$. Gilt $z^* \leq \min_{i=1, \dots, p-1} (|\beta_i| \cdot (v_i - u_i))$, dann ist das D -optimale Design ξ^* gegeben durch:

$$\xi^* = \left\{ \begin{array}{ccc} \mathbf{d} - (z^*/\beta_1) \cdot \mathbf{e}_{1,p-1} & \dots & \mathbf{d} - (z^*/\beta_{p-1}) \cdot \mathbf{e}_{p-1,p-1} & \mathbf{d} \\ w_1^*(z^*) & \dots & w_{p-1}^*(z^*) & w_p^*(z^*) \end{array} \right\}. \quad (5.19)$$

Beweis:

Zur Vereinfachung der Notation seien $\mathbf{M} = \mathbf{M}(\xi^*, \boldsymbol{\beta})$ und $\mathbf{M}_{\text{Po}} = \mathbf{M}_{\text{Po}}(\xi^*, \boldsymbol{\beta})$. Da das Optimierungsproblem nicht von der positiven Konstante a abhängt, wird außerdem $a = b$ gewählt.

Es wird der erweiterte Designraum $\mathcal{X}_{\text{erw}} = \mathcal{X}_{1,\text{erw}} \times \dots \times \mathcal{X}_{p-1,\text{erw}}$ mit $\mathcal{X}_{i,\text{erw}} = (-\infty, v_i]$ für $\beta_i > 0$ und $\mathcal{X}_{i,\text{erw}} = [u_i, \infty)$ für $\beta_i < 0$ für $i = 1, \dots, p-1$ betrachtet. Da die Matrix $\mathbf{M}_{\text{Po}}^{-1} \mathbf{M} \mathbf{M}_{\text{Po}}^{-1}$ positiv definit ist und die Mengen $\{\mathbf{x} \in \mathbb{R}^{p-1} : \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = c\} \cap \mathcal{X}_{\text{erw}}$ für alle $c \in \mathbb{R}$ beschränkt sind, wird die linke Seite der Bedingung im Äquivalenzsatz 4.10 auf den Kanten des erweiterten Designraums \mathcal{X}_{erw} maximal. Somit genügt es zum Nachweis der D -Optimalität des Designs ξ^* zu zeigen, dass die Bedingung des Äquivalenzsatzes auf den Kanten von \mathcal{X}_{erw} erfüllt ist (vgl. Schmidt und Schwabe (2017), Schmidt (2019)). Daher ist das Design ξ^* genau dann D -optimal, falls

$$e^{\mathbf{f}(\mathbf{d} + (x_i - d_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{d} + (x_i - d_i) \cdot \mathbf{e}_{i,p-1})^T \mathbf{M}_{\text{Po}}^{-1} \mathbf{M} \mathbf{M}_{\text{Po}}^{-1} \mathbf{f}(\mathbf{d} + (x_i - d_i) \cdot \mathbf{e}_{i,p-1}) \leq \text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1})$$

für alle $x_i \in \mathcal{X}_{i,\text{erw}}$ gilt, $i = 1, \dots, p-1$. Es werden die folgenden zwei Funktionen

eingeführt:

$$g_i(x) = \mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1})^T \mathbf{M}_{\text{Po}}^{-1} \mathbf{M} \mathbf{M}_{\text{Po}}^{-1} \mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1}),$$

$$h_i(x) = g_i(x) - \exp(-\mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}) \cdot \text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1}).$$

Die Bedingung im Äquivalenzsatz ist äquivalent zu

$$h_i(x_i) = g_i(x_i) - \exp(-\mathbf{f}(\mathbf{d} + (x_i - d_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}) \cdot \text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1}) \leq 0 \quad (5.20)$$

für alle $x_i \in \mathcal{X}_{i,\text{erw}}$, $i = 1, \dots, p-1$.

Mit der unteren Dreiecksmatrix $\mathbf{L} = \mathbf{I}_p - (\mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T) / (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m})$ aus Lemma 4.1 gilt $\mathbf{M} \mathbf{M}_{\text{Po}}^{-1} = \mathbf{L} \mathbf{M}_{\text{Po}} \mathbf{M}_{\text{Po}}^{-1} = \mathbf{L}$. Somit folgt:

$$\mathbf{M}_{\text{Po}}^{-1} \mathbf{M} \mathbf{M}_{\text{Po}}^{-1} = \mathbf{M}_{\text{Po}}^{-1} \mathbf{L} = \mathbf{M}_{\text{Po}}^{-1} - \frac{\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}},$$

$$\text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1}) = \text{Spur}(\mathbf{L}) = p - \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} = \frac{(p-1) \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + p \cdot \frac{b}{m}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}}.$$

Es sei $\lambda_1 = \exp(\mathbf{f}(\mathbf{d})^T \boldsymbol{\beta} - z^*)$ und $\lambda_p = \exp(\mathbf{f}(\mathbf{d})^T \boldsymbol{\beta})$. Für das Design ξ^* besitzt die Informationsmatrix für das Poisson-Modell die Matrix-Darstellung $\mathbf{M}_{\text{Po}} = \mathbf{X}^T \boldsymbol{\Lambda} \mathbf{W} \mathbf{X}$ mit $\boldsymbol{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_1, \lambda_p)$, $\mathbf{W} = \text{Diag}(w_1^*, \dots, w_1^*, w_p^*)$ und

$$\mathbf{X} = \begin{pmatrix} \mathbf{1}_{p-1} & \mathbf{1}_{p-1} \cdot \mathbf{d}^T - z^* \cdot \text{Diag}(\tilde{\boldsymbol{\beta}}_{\text{rez}}) \\ 1 & \mathbf{d}^T \end{pmatrix}$$

mit $\mathbf{1}_{p-1} = (1, \dots, 1)^T \in \mathbb{R}^{p-1}$ und $\tilde{\boldsymbol{\beta}}_{\text{rez}} = (1/\beta_1, \dots, 1/\beta_{p-1})^T$, welcher aus den Kehrwerten der Komponenten von $\tilde{\boldsymbol{\beta}} = (\beta_1, \dots, \beta_{p-1})^T$ besteht. Die Inverse von \mathbf{X} ist durch

$$\mathbf{X}^{-1} = \frac{1}{z^*} \cdot \begin{pmatrix} (\text{Diag}(\tilde{\boldsymbol{\beta}}) \mathbf{d})^T & z^* - \mathbf{d}^T \tilde{\boldsymbol{\beta}} \\ -\text{Diag}(\tilde{\boldsymbol{\beta}}) & \tilde{\boldsymbol{\beta}} \end{pmatrix}$$

gegeben.

Es wird nun gezeigt, dass an den Stützpunkten des Designs ξ^* Gleichheit in Bedingung (5.20) gilt. Zunächst wird der Stützpunkt \mathbf{d} betrachtet. Für die Inverse der Informationsmatrix für das Poisson-Modell gilt $\mathbf{M}_{\text{Po}}^{-1} = \mathbf{X}^{-1} \mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} (\mathbf{X}^T)^{-1}$. Mit dem Zusammenhang $(\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{d}) = (\mathbf{X}^T)^{-1} \mathbf{X}^T \mathbf{e}_{p,p} = \mathbf{e}_{p,p}$ folgt:

$$g_i(d_i) = \mathbf{f}(\mathbf{d} + (d_i - d_i) \cdot \mathbf{e}_{i,p-1})^T \mathbf{M}_{\text{Po}}^{-1} \mathbf{M} \mathbf{M}_{\text{Po}}^{-1} \mathbf{f}(\mathbf{d} + (d_i - d_i) \cdot \mathbf{e}_{i,p-1})$$

$$\begin{aligned}
&= \mathbf{f}(\mathbf{d})^T \cdot \left(\mathbf{X}^{-1} \mathbf{W}^{-1} \mathbf{\Lambda}^{-1} (\mathbf{X}^T)^{-1} - \frac{\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right) \cdot \mathbf{f}(\mathbf{d}) \\
&= \mathbf{e}_{p,p}^T \mathbf{W}^{-1} \mathbf{\Lambda}^{-1} \mathbf{e}_{p,p} - \frac{1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} = \frac{1}{w_p^* \cdot \lambda_p} - \frac{1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \\
&= \frac{\frac{1}{w_p^* \cdot \lambda_p} \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}) - 1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}}.
\end{aligned}$$

Somit gilt:

$$\begin{aligned}
h_i(d_i) &= g_i(d_i) - \exp(-\mathbf{f}(\mathbf{d} + (d_i - d_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}) \cdot \text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1}) \\
&= \frac{\frac{1}{w_p^* \cdot \lambda_p} \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}) - 1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} - \frac{1}{\lambda_p} \cdot \frac{(p-1) \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + p \cdot \frac{b}{m}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \\
&= \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} - w_p^* \cdot \lambda_p - (p-1) \cdot w_p^* \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m} \cdot (1 - p \cdot w_p^*)}{w_p^* \cdot \lambda_p \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m})}.
\end{aligned}$$

Mit $\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} = (p-1) \cdot w_1^* \cdot \lambda_1 + w_p^* \cdot \lambda_p$ und $(1 - p \cdot w_p^*) / (p-1) = -1 + p \cdot w_1^*$ folgt:

$$\begin{aligned}
h_i(d_i) &= \frac{(p-1) \cdot w_1^* \cdot \lambda_1 - (p-1) \cdot w_p^* \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m} \cdot (1 - p \cdot w_p^*)}{w_p^* \cdot \lambda_p \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m})} \\
&= (p-1) \cdot \frac{w_1^* \cdot \lambda_1 - w_p^* \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m} \cdot (-1 + p \cdot w_1^*)}{w_p^* \cdot \lambda_p \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m})}.
\end{aligned}$$

Da nach Lemma A.5 der Zähler gleich Null ist, gilt $h_i(d_i) = 0$.

Nun wird gezeigt, dass an den Stützpunkten $\mathbf{d} - (z^*/\beta_i) \cdot \mathbf{e}_{i,p-1}$ für $i = 1, \dots, p-1$ ebenfalls Gleichheit in Bedingung (5.20) gilt. Unter Verwendung des Zusammenhangs $(\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{d} - (z^*/\beta_i) \cdot \mathbf{e}_{i,p-1}) = (\mathbf{X}^T)^{-1} \mathbf{X}^T \mathbf{e}_{i,p} = \mathbf{e}_{i,p}$ folgt:

$$\begin{aligned}
&g_i\left(d_i - \frac{z^*}{\beta_i}\right) \\
&= \mathbf{f}\left(\mathbf{d} - \frac{z^*}{\beta_i} \cdot \mathbf{e}_{i,p-1}\right)^T \cdot \left(\mathbf{X}^{-1} \mathbf{W}^{-1} \mathbf{\Lambda}^{-1} (\mathbf{X}^T)^{-1} - \frac{\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right) \cdot \mathbf{f}\left(\mathbf{d} - \frac{z^*}{\beta_i} \cdot \mathbf{e}_{i,p-1}\right) \\
&= \mathbf{e}_{i,p}^T \mathbf{W}^{-1} \mathbf{\Lambda}^{-1} \mathbf{e}_{i,p} - \frac{1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} = \frac{1}{w_1^* \cdot \lambda_1} - \frac{1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \\
&= \frac{\frac{1}{w_1^* \cdot \lambda_1} \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}) - 1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}}.
\end{aligned}$$

Mit $\exp(\mathbf{f}(\mathbf{d} - (z^*/\beta_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}) = \exp(\mathbf{f}(\mathbf{d})^T \boldsymbol{\beta} - z^*) = \lambda_1$ gilt somit:

$$\begin{aligned} h_i \left(d_i - \frac{z^*}{\beta_i} \right) &= \frac{\frac{1}{w_1^* \lambda_1} \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}) - 1}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} - \frac{1}{\lambda_1} \cdot \frac{(p-1) \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + p \cdot \frac{b}{m}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \\ &= \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} \cdot (1 - (p-1) \cdot w_1^*) - w_1^* \cdot \lambda_1 + \frac{b}{m} \cdot (1 - p \cdot w_1^*)}{w_1^* \cdot \lambda_1 \cdot (\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m})}. \end{aligned}$$

Da $1 - (p-1) \cdot w_1^* = w_p^*$ gilt, folgt $h_i(d_i - z^*/\beta_i) = 0$ für $i = 1, \dots, p-1$ mit Lemma A.5. Die erste und zweite Ableitung von h_i sind durch

$$\begin{aligned} h_i'(x) &= g_i'(x) + \beta_i \cdot \exp(-\mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}) \cdot \text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1}), \\ h_i''(x) &= g_i''(x) - \beta_i^2 \cdot \exp(-\mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1})^T \boldsymbol{\beta}) \cdot \text{Spur}(\mathbf{M} \mathbf{M}_{\text{Po}}^{-1}) \end{aligned}$$

gegeben. Da g_i ein quadratisches Polynom ist, ist g_i'' konstant. Die Exponentialfunktion ist injektiv, somit ist h_i'' ebenfalls injektiv und hat höchstens eine Nullstelle. Nach dem Satz von Rolle kann h_i' maximal zwei Nullstellen besitzen. Daher hat h_i höchstens zwei Extrema. Außerdem gilt für den Grenzwert $\lim_{x \rightarrow \pm\infty} h_i(x) = \pm\infty$ für $\beta_i > 0$ und $\lim_{x \rightarrow \pm\infty} h_i(x) = \mp\infty$ für $\beta_i < 0$. Somit existieren keine Sattelpunkte. Das Design ξ^* ist D -optimal, wenn für $i = 1, \dots, p-1$ bei $d_i - z^*/\beta_i$ ein Maximum von h_i liegt. Wegen $h_i(d_i) = h_i(d_i - z^*/\beta_i) = 0$ befindet sich dann zwischen den beiden Stützpunkten ein Minimum. Bei $d_i - z^*/\beta_i$ liegt ein Maximum vor, wenn h_i' bei $d_i - z^*/\beta_i$ eine Nullstelle hat. Die Ableitung von g_i berechnet sich zu:

$$\begin{aligned} g_i'(x) &= 2 \cdot \mathbf{e}_{i+1,p}^T \mathbf{M}_{\text{Po}}^{-1} \mathbf{M} \mathbf{M}_{\text{Po}}^{-1} \mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1}) \\ &= 2 \cdot \mathbf{e}_{i+1,p}^T \cdot \left(\mathbf{X}^{-1} \mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} (\mathbf{X}^T)^{-1} - \frac{\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right) \cdot \mathbf{f}(\mathbf{d} + (x - d_i) \cdot \mathbf{e}_{i,p-1}). \end{aligned}$$

Mit $\mathbf{e}_{i+1,p}^T \mathbf{X}^{-1} = (\beta_i/z^*) \cdot (-\mathbf{e}_{i,p-1}^T, 1)$ und $(\mathbf{X}^T)^{-1} \mathbf{f}(\mathbf{d} - (z^*/\beta_i) \cdot \mathbf{e}_{i,p-1}) = \mathbf{e}_{i,p}$ folgt:

$$\begin{aligned} g_i' \left(d_i - \frac{z^*}{\beta_i} \right) &= 2 \cdot \mathbf{e}_{i+1,p}^T \cdot \left(\mathbf{X}^{-1} \mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} (\mathbf{X}^T)^{-1} - \frac{\mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right) \cdot \mathbf{f} \left(\mathbf{d} - \frac{z^*}{\beta_i} \cdot \mathbf{e}_{i,p-1} \right) \\ &= 2 \cdot \frac{\beta_i}{z^*} \cdot (-\mathbf{e}_{i,p-1}^T, 1) \mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} \mathbf{e}_{i,p} - 2 \cdot \frac{\mathbf{e}_{i+1,p}^T \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{f}(\mathbf{d} - \frac{z^*}{\beta_i} \cdot \mathbf{e}_{i,p-1})}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}}. \end{aligned}$$

Wegen $\mathbf{e}_{i+1,p}^T \mathbf{e}_{1,p} = 0$ folgt:

$$g'_i \left(d_i - \frac{z^*}{\beta_i} \right) = -2 \cdot \frac{\beta_i}{z^*} \cdot \frac{1}{w_1^* \cdot \lambda_1}.$$

Somit gilt:

$$\begin{aligned} h'_i \left(d_i - \frac{z^*}{\beta_i} \right) &= -2 \cdot \frac{\beta_i}{z^*} \cdot \frac{1}{w_1^* \cdot \lambda_1} + \frac{\beta_i}{\lambda_1} \cdot \frac{(p-1) \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + p \cdot \frac{b}{m}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \\ &= \frac{\beta_i}{\lambda_1} \cdot \frac{-\frac{2}{z^* \cdot w_1^*} \cdot \left(\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m} \right) + (p-1) \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + p \cdot \frac{b}{m}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}}. \end{aligned}$$

Die Gleichung $h'_i(d_i - z^*/\beta_i) = 0$ ist äquivalent zu:

$$\begin{aligned} 0 &= -2 \cdot \left(m \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + b \right) + z^* \cdot w_1^* \cdot \left(m \cdot (p-1) \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + p \cdot b \right) \\ &= m \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} \cdot \left((p-1) \cdot z^* \cdot w_1^* - 2 \right) + b \cdot \left(p \cdot z^* \cdot w_1^* - 2 \right) \\ &= m \cdot \left((p-1) \cdot w_1^* \cdot \lambda_1 + w_p^* \cdot \lambda_p \right) \cdot \left((p-1) \cdot z^* \cdot w_1^* - 2 \right) + b \cdot \left(p \cdot z^* \cdot w_1^* - 2 \right). \end{aligned}$$

Nun wird mit dem Zwischenwertsatz gezeigt, dass diese Gleichung eine Lösung besitzt.

Wegen $\lim_{z^* \rightarrow 0} \lambda_1 = \lambda_p$ gilt:

$$\begin{aligned} &\lim_{z^* \rightarrow 0} m \cdot \left((p-1) \cdot w_1^* \cdot \lambda_1 + w_p^* \cdot \lambda_p \right) \cdot \left((p-1) \cdot z^* \cdot w_1^* - 2 \right) + b \cdot \left(p \cdot z^* \cdot w_1^* - 2 \right) \\ &= -2 \cdot m \cdot \lambda_p - 2 \cdot b < 0. \end{aligned}$$

Nach Satz 5.8 gilt $1/p < w_1^* < 1/(p-1)$. Daher gilt:

$$\lim_{z^* \rightarrow \infty} m \cdot \left((p-1) \cdot w_1^* \cdot \lambda_1 + w_p^* \cdot \lambda_p \right) \cdot \left((p-1) \cdot z^* \cdot w_1^* - 2 \right) + b \cdot \left(p \cdot z^* \cdot w_1^* - 2 \right) = \infty.$$

Somit besitzt die Gleichung eine Lösung im Intervall $(0, \infty)$ und das Design ξ^* ist D -optimal auf \mathcal{X}_{erw} .

Hätte die Gleichung mehr als eine Lösung im Intervall $(0, \infty)$, so würde ein weiteres D -optimales Design existieren. Da die Kriteriumsfunktion isoton und konkav ist, wäre nach Satz 4.8 jede Konvexkombination dieser Designs ebenfalls D -optimal. Das resultierende Design hätte mehr als zwei Stützpunkte auf einer Kante von \mathcal{X}_{erw} , was einen Widerspruch zum Verlauf von h_i darstellt. Somit ist die Lösung der Gleichung im Intervall $(0, \infty)$ eindeutig und das D -optimale Design ist ebenfalls eindeutig.

Falls $z^* \leq \min_{i=1, \dots, p-1} (|\beta_i| \cdot (v_i - u_i))$ gilt, befinden sich alle Stützpunkte von ξ^* innerhalb des Designraums $\mathcal{X} \subseteq \mathcal{X}_{\text{erw}}$. Somit ist das Design ξ^* auch D -optimal auf \mathcal{X} . \square

Die Struktur des D -optimalen Designs für das Poisson-Gamma-Modell ist für den Fall von drei Kovariablen in Abbildung 5.1 dargestellt. Ein Stützpunkt befindet sich auf der Ecke \mathbf{d} des Designraums, wo die Intensitätsfunktion $\exp(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta})$ maximal ist. Da $z^* \leq \min_{i=1, \dots, p-1} (|\beta_i| \cdot (v_i - u_i))$ gilt, befinden sich die anderen Stützpunkte innerhalb des Designraums. Sie liegen auf den Kanten, welche inzident zu \mathbf{d} sind. Für das Poisson-Gamma-Modell ist die Distanz vom Stützpunkt \mathbf{d} zu den anderen Stützpunkten durch $z^*/|\beta_i|$ gegeben. Gleichung (5.18) kann nur erfüllt sein, wenn $z \cdot (p-1) \cdot w_1^*(z) - 2 < 0$ und $z \cdot p \cdot w_1^*(z) - 2 > 0$ gilt. Da $1/p < w_1^*(z) < 1/(p-1)$ nach Satz 5.8 gilt, folgt für die Lösung z^* von Gleichung (5.18):

$$2 \cdot \frac{p-1}{p} < z^* < 2 \cdot \frac{p}{p-1}. \quad (5.21)$$

Die D -optimalen Gewichte für die ersten $p-1$ Stützpunkte sind gleich, aber unterscheiden sich von dem für die Ecke \mathbf{d} . Dabei ist das Gewicht für den Stützpunkt \mathbf{d} geringer als die Gewichte für die anderen Stützpunkte.

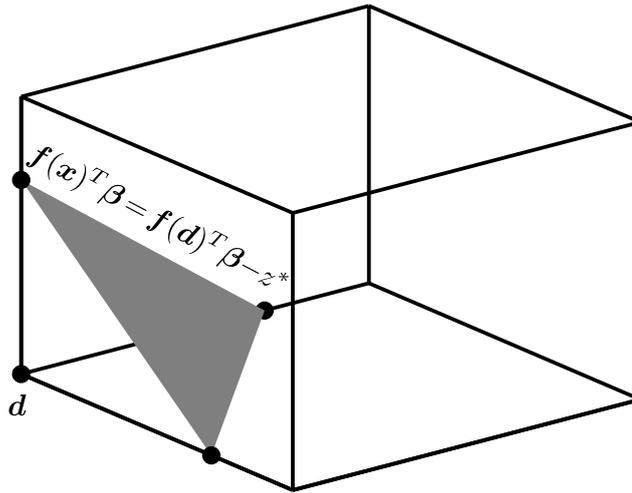


Abbildung 5.1: Grafische Darstellung des D -optimalen Designs für das Poisson-Gamma-Modell für den Fall von drei Kovariablen

Für das Poisson-Modell mit Designraum $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$, $\beta_i \neq 0$, $i = 1, \dots, p-1$, und $2 \leq \min_{i=1, \dots, p-1} (|\beta_i| \cdot (v_i - u_i))$ ist das D -optimale Design durch

$$\xi^* = \left\{ \begin{array}{cccc} \mathbf{d} - (2/\beta_1) \cdot \mathbf{e}_{1,p-1} & \dots & \mathbf{d} - (2/\beta_{p-1}) \cdot \mathbf{e}_{p-1,p-1} & \mathbf{d} \\ 1/p & \dots & 1/p & 1/p \end{array} \right\} \quad (5.22)$$

gegeben (vgl. Russell et al., 2009). Dabei ist \mathbf{d} wie in Satz 5.10 für das Poisson-Gamma-Modell definiert. Die Struktur der D -optimalen Designs für das Poisson- und Poisson-Gamma-Modell ist ähnlich. Die Distanz von der Ecke \mathbf{d} zu den anderen Stützpunkten

auf den Kanten unterscheidet sich aber bei beiden Modellen. Für das Poisson-Modell ist sie durch $2/|\beta_i|$ gegeben, also ist $z^* = 2$. Weiterhin unterscheiden sich die Gewichte der optimalen Designs bei beiden Modellen, da beim D -optimalen Design für das Poisson-Gamma-Modell von der Gleichverteilungsregel der Gewichte für das Poisson-Modell abgewichen wird.

Das D -optimale Design für das Poisson-Gamma-Modell hängt nicht vom Parameter a der Gamma-Verteilung ab. Wird $a = b^{(p-1)/p}$ gewählt, so vereinfacht sich die Kriteriumsfunktion für D -Optimalität für das Poisson-Gamma-Modell in Gleichung (5.1) zu $\det(\mathbf{M}(\xi; \boldsymbol{\beta})) = \det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})) / (b + m \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p})$. Somit lässt sich für $b \rightarrow 0$ das D_s -optimale Design für die Parameter $\beta_1, \dots, \beta_{p-1}$ im Poisson-Modell aus Satz 5.10 folgern (vgl. Schmidt und Schwabe, 2020).

Satz 5.11

Es sei $\mathcal{X} = [u_1, v_1] \times \dots \times [u_{p-1}, v_{p-1}]$ und $\beta_i \neq 0$ für $i = 1, \dots, p-1$. Sei $\mathbf{d} = (d_1, \dots, d_{p-1})$ mit $d_i = v_i$ für $\beta_i > 0$ und $d_i = u_i$ für $\beta_i < 0$. Für beliebiges $z > 0$ sei

$$w_p^*(z) = \frac{2}{p + \sqrt{(p-2)^2 + 4 \cdot (p-1) \cdot \exp(z)}} \quad (5.23)$$

und $w_1^*(z) = \dots = w_{p-1}^*(z) = (1 - w_p^*(z))/(p-1)$. Die Gleichung

$$z \cdot (1 - w_p^*(z)) - 2 = 0 \quad (5.24)$$

besitzt eine eindeutige Lösung z^* in $(0, \infty)$. Gilt $z^* \leq \min_{i=1, \dots, p-1} (|\beta_i| \cdot (v_i - u_i))$, dann ist das D_s -optimale Design ξ^* für $\beta_1, \dots, \beta_{p-1}$ im Poisson-Modell gegeben durch:

$$\xi^* = \left\{ \begin{array}{cccc} \mathbf{d} - (z^*/\beta_1) \cdot \mathbf{e}_{1,p-1} & \dots & \mathbf{d} - (z^*/\beta_{p-1}) \cdot \mathbf{e}_{p-1,p-1} & \mathbf{d} \\ w_1^*(z^*) & \dots & w_{p-1}^*(z^*) & w_p^*(z^*) \end{array} \right\}. \quad (5.25)$$

Gleichung (5.24) für z^* lässt sich als $z = 2/(1 - w_p^*(z))$ schreiben. Da $0 < w_p^*(z) < 1/p$ nach Satz 5.8 gilt, folgt für die Lösung z^* :

$$2 < z^* < 2 \cdot \frac{p}{p-1}. \quad (5.26)$$

Insbesondere ist im Poisson-Modell der Abstand der Stützpunkte auf den Kanten zu dem Stützpunkt auf der Ecke beim D_s -optimalen Design für die Parameter $\beta_1, \dots, \beta_{p-1}$ größer als beim D -optimalen Design. Zudem sind im Gegensatz zu D -Optimalität die D_s -optimalen Gewichte nicht alle gleich.

Satz 5.12

Es seien $\mathbf{f} = (1, f_1, \dots, f_{p-1})^T$ und $\mathbf{A}^T = (\mathbf{0}_s, \tilde{\mathbf{A}})$, wobei $\tilde{\mathbf{A}}$ eine $s \times (p-1)$ -Matrix mit $\text{Rang}(\tilde{\mathbf{A}}) = s$ ist. Dann sind die Kriteriumsfunktionen für D_A -Optimalität für das Poisson- und Poisson-Gamma-Modell bis auf einen positiven konstanten Faktor identisch. Somit ist ein Design ξ^* genau dann D_A -optimal im Poisson-Gamma-Modell, wenn das Design ξ^* im Poisson-Modell D_A -optimal ist.

Beweis:

Mit Satz 4.4 für die verallgemeinerte Inverse der Informationsmatrix folgt die Äquivalenz zwischen den Kriteriumsfunktionen für das Poisson-Gamma-Modell und für das Poisson-Modell:

$$\begin{aligned} \det(\mathbf{A}^T \mathbf{M}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}) &= \det\left(\mathbf{A}^T \cdot \left(\frac{b}{a} \cdot \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-1} + \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T\right) \cdot \mathbf{A}\right) \\ &= \left(\frac{b}{a}\right)^s \cdot \det(\mathbf{A}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-1} \mathbf{A}). \end{aligned}$$

Dabei wurde der Zusammenhang $\mathbf{A}^T \mathbf{e}_{1,p} = \mathbf{0}_s$ verwendet. Da die Identifizierbarkeitsbedingung nach Satz 4.3 im Poisson-Modell und Poisson-Gamma-Modell äquivalent ist, stimmen die D_A -optimalen Designs in beiden Modellen überein. \square

Bemerkung 5.13

Satz 5.12 gilt für beliebige Regressionsfunktionen f_i , $i = 1, \dots, p-1$. Für das multiple Regressionsmodell mit $f_i(\mathbf{x}) = x_i$ ist das D_s -optimale Design für $\beta_1, \dots, \beta_{p-1}$ im Poisson-Gamma-Modell somit durch das Design (5.25) in Satz 5.11 gegeben.

Im folgenden Beispiel werden das D -optimale Design im Poisson-Gamma-Modell und das D - und D_s -optimale Design im Poisson-Modell berechnet.

Beispiel 5.14

Es wird das multiple Regressionsmodell mit zwei Kovariablen betrachtet. Der Designraum sei $\mathcal{X} = [0, 3] \times [0, 3]$, die Regressionsfunktion sei $\mathbf{f}(\mathbf{x}) = (1, x_1, x_2)^T$ und der Parametervektor sei durch $\boldsymbol{\beta} = (0, -1, -1)^T$ gegeben. Es seien $m = 10$ und $b = 1$.

Das D -optimale Design für das Poisson-Gamma-Modell folgt aus Satz 5.10. Um das D -optimale Design zu bestimmen, wird die Lösung von Gleichung (5.18) berechnet, welche sich zu $z^* = 2.240$ ergibt.

Das D_s -optimale Design für die Parameter β_1 und β_2 für das Poisson- und Poisson-Gamma-Modell ergibt sich aus Satz 5.11 und Satz 5.12. Die Lösung von Gleichung (5.24) berechnet sich zu $z^* = 2.385$.

Modell	Kriterium	Optimales Design	Effizienzen		
			Po D	P-G D	Po D_s
Poisson	D	$\left\{ \begin{matrix} (2, 0) & (0, 2) & (0, 0) \\ 1/3 & 1/3 & 1/3 \end{matrix} \right\}$	1	0.956	0.886
Poisson-Gamma	D	$\left\{ \begin{matrix} (2.240, 0) & (0, 2.240) & (0, 0) \\ 0.396 & 0.396 & 0.208 \end{matrix} \right\}$	0.950	1	0.988
Poisson Poisson-Gamma	D_s	$\left\{ \begin{matrix} (2.385, 0) & (0, 2.385) & (0, 0) \\ 0.419 & 0.419 & 0.162 \end{matrix} \right\}$	0.895	0.990	1

Tabelle 5.2: Optimale Designs in Beispiel 5.14 und Vergleich der Effizienzen für das Poisson-Gamma-Modell (P-G) und das Poisson-Modell (Po)

In Tabelle 5.2 ist das D -optimale Design für das Poisson-Gamma-Modell angegeben. Außerdem wurden das D -optimale Design und das D_s -optimale Design für die Parameter β_1 und β_2 für das Poisson-Modell berechnet. Zudem werden die optimalen Designs bezüglich ihrer Effizienzen verglichen.

Alle drei optimalen Designs sind von der gleichen Struktur. Ein Stützpunkt ist bei $(0, 0)$. Dies ist der Punkt des Designraums, an dem $\exp(\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta})$ maximal wird. Die beiden anderen Stützpunkte des D -optimalen Designs für das Poisson-Gamma-Modell liegen auf den Kanten des Designraums zwischen den entsprechenden Stützpunkten der D - und D_s -optimalen Designs für das Poisson-Modell. Beim D_s -optimalen Design im Poisson-Modell ist der Abstand der Stützpunkte auf den Kanten zum Stützpunkt $(0, 0)$ auf der Ecke des Designraums größer als beim D -optimalen Design im Poisson-Modell.

Im Vergleich zum D -optimalen Design im Poisson-Modell besitzt das D_s -optimale Design in diesem Beispiel eine höhere Effizienz bezüglich des D -optimalen Designs im Poisson-Gamma-Modell. Die Effizienzen sind höher als beim Modell mit einer Kovariablen, es tritt jedoch immer noch ein Effizienzverlust auf. Dies lässt sich dadurch erklären, dass die Stützpunkte der optimalen Designs für die verschiedenen Optimalitätskriterien mit steigender Anzahl an Kovariablen näher beieinander liegen.

5.2 Lineare Optimalitätskriterien

Es sei $\mathbf{f} = (1, f_1, \dots, f_{p-1})^T$ mit beliebigen Regressionsfunktionen f_i , etwa $f_i(x) = x^i$ für polynomiale Regression mit eindimensionaler Kovariable $x \in \mathbb{R}$ oder $f_i(\mathbf{x}) = x_i$ für multiple Regression mit $\mathbf{x} \in \mathbb{R}^{p-1}$.

Für spezielle Lineare Gemischte Modelle wie dem Random-Coefficient-Regressionsmodell ist bereits bekannt, dass für lineare Optimalitätskriterien die optimalen Designs in den Modellen mit und ohne zufällige Effekte gleich sind (vgl. Liski et al., 2002, S. 24). Unter

Verwendung von Satz 4.4 wird im nächsten Satz gezeigt, dass die c - und L -optimalen Designs im Poisson-Gamma-Modell mit denen im Poisson-Modell übereinstimmen (vgl. Schmidt und Schwabe, 2020).

Satz 5.15

Ein Design ξ^* ist genau dann im Poisson-Gamma-Modell L -optimal (c -optimal), wenn das Design ξ^* im Poisson-Modell L -optimal (c -optimal) ist.

Beweis:

Es sei $\mathbf{B} = \mathbf{A}\mathbf{A}^T$. Nach Satz 4.3 ist für ein Design ξ die Identifizierbarkeit von $\mathbf{A}^T\boldsymbol{\beta}$ in beiden Modellen äquivalent. Mit dem in Satz 4.4 gezeigten Zusammenhang zwischen den verallgemeinerten Inversen von $\mathbf{M}(\xi; \boldsymbol{\beta})$ und $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ gilt für ein Design ξ :

$$\begin{aligned} \text{Spur}(\mathbf{M}(\xi; \boldsymbol{\beta})^{-} \mathbf{B}) &= \text{Spur} \left(\left(\frac{b}{a} \cdot \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-} + \frac{m}{a} \cdot \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \right) \cdot \mathbf{B} \right) \\ &= \frac{b}{a} \cdot \text{Spur}(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})^{-} \mathbf{B}) + \frac{m}{a} \cdot \mathbf{e}_{1,p}^T \mathbf{B} \mathbf{e}_{1,p}. \end{aligned}$$

Da der zweite Summand nicht vom Design abhängt, folgt die Äquivalenz der Optimalität eines Designs in beiden Modellen. Mit $\mathbf{B} = \mathbf{c}\mathbf{c}^T$ ergibt sich c -Optimalität als Spezialfall von L -Optimalität. Daher folgt die Äquivalenz auch für c -Optimalität. \square

Da das A -Optimalitätskriterium und das $IMSE$ -Optimalitätskriterium, bei dem die Matrizen \mathbf{B} und \mathbf{B}_{Po} bis auf einen multiplikativen Faktor übereinstimmen, Spezialfälle des L -Optimalitätskriteriums sind, gilt Satz 5.15 und somit die Gleichheit der optimalen Designs im Poisson- und Poisson-Gamma-Modell auch für diese Optimalitätskriterien. Satz 5.15 zeigt, dass man sich bei der Bestimmung von c - und L -optimalen Designs auf das Poisson-Modell beschränken kann, was die Berechnung von optimalen Designs vereinfacht. Insbesondere können die für das Poisson-Modell erzielten Resultate zu c -optimalen Designs von Ford et al. (1992) und Rodríguez-Torreblanca und Rodríguez-Díaz (2007) für eine Kovariable und zu c - und L -optimalen Designs von Schmidt (2019) für multiple Regression auf das Poisson-Gamma-Modell übertragen werden.

6 Standardisiert Maximin optimale Designs

Es wird weiterhin das Poisson-Gamma-Modell mit bekannten Parametern a und b der Gamma-Verteilung betrachtet. Beim Poisson-Gamma-Modell handelt es sich um ein nichtlineares Modell. Daher hängen die optimalen Designs vom Parametervektor $\boldsymbol{\beta}$ ab. Bei Misspezifikation von $\boldsymbol{\beta}$ können die für $\boldsymbol{\beta}$ optimalen Designs bezüglich des optimalen Designs für den wahren Wert von $\boldsymbol{\beta}$ eine geringere Effizienz aufweisen.

Dieser Effizienzverlust wird für den Designraum $\mathcal{X} = [0, 3]$, die Regressionsfunktion $\boldsymbol{f}(x) = (1, x)^T$ und den Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ sowie $m = 10$ und $b = 1$ untersucht. Das lokal D -optimale Design $\xi_{\boldsymbol{\beta}}^*$ für den Parametervektor $\boldsymbol{\beta} = (0, -1)^T$ ist in Beispiel 5.7 gegeben. In Abbildung 6.1 wird die D -Effizienz des Designs $\xi_{\boldsymbol{\beta}}^*$ bei Parametermisspezifikation von β_0 und β_1 dargestellt.

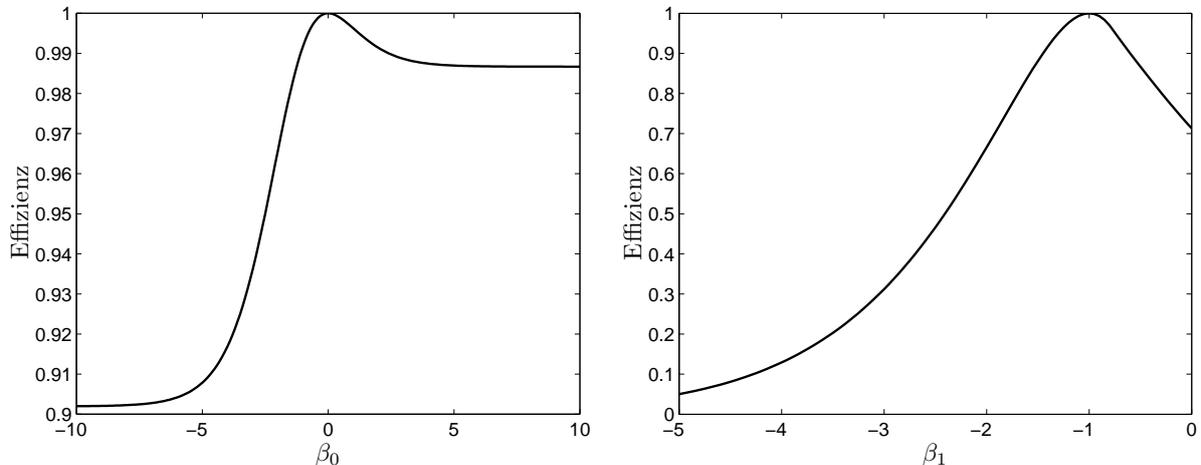


Abbildung 6.1: D -Effizienz des für $\boldsymbol{\beta} = (0, -1)^T$ lokal D -optimalen Designs $\xi_{\boldsymbol{\beta}}^*$ bei Variation von β_0 (links) und β_1 (rechts)

In der linken Abbildung wird der wahre Wert von β_0 von -10 bis 10 variiert, wobei $\beta_1 = -1$ konstant ist. Für $\beta_0 = 0$ ist $\xi_{\boldsymbol{\beta}}^*$ das lokal D -optimale Design und besitzt somit eine D -Effizienz von 1. Die D -Effizienz von $\xi_{\boldsymbol{\beta}}^*$ fällt mit dem Abstand des wahren Werts von β_0 zum Wert 0, der für die Berechnung des lokal D -optimalen Designs $\xi_{\boldsymbol{\beta}}^*$ angenommen wurde. Wie bereits in Abschnitt 2.4 gezeigt, lässt sich die Dichte für das

Poisson-Gamma-Modell so darstellen, dass die Parameter b und β_0 immer als Produkt $(1/b) \cdot e^{\beta_0}$ auftreten. Dies gilt somit auch für die Informationsmatrix. Daher ergibt sich für $\beta_0 \rightarrow -\infty$ das D -optimale Design und für $\beta_0 \rightarrow \infty$ das c -optimale Design mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell. Nach Tabelle 5.1 in Beispiel 5.7 ergibt sich somit die Effizienz für $\beta_0 \rightarrow -\infty$ zu 0.902 und für $\beta_0 \rightarrow \infty$ zu 0.974. Das Design ξ_{β}^* besitzt also immer eine D -Effizienz, welche größer als 0.9 ist.

In der rechten Abbildung wird der wahre Wert von β_1 von -5 bis 0 variiert, wobei $\beta_0 = 0$ konstant ist. Das Design ξ_{β}^* ist für $\beta_1 = -1$ das lokal D -optimale Design und besitzt eine D -Effizienz von 1. Die D -Effizienz von ξ_{β}^* nimmt mit zunehmendem Abstand des wahren Werts von β_1 zum Wert -1 ab. Da der Designraum durch $\mathcal{X} = [0, 3]$ gegeben ist und somit für $\beta_1 \rightarrow 0$ die D -optimalen Stützpunkte bei 0 und 3 liegen, fällt die D -Effizienz für $\beta_1 \rightarrow 0$ nicht so stark wie für $\beta_1 \rightarrow -\infty$ ab. Beispielsweise liegt die D -Effizienz für $\beta_1 = -5$ unter 0.1 .

Das einführende Beispiel zeigt, dass der Effizienzverlust bei Missspezifikation von β_1 deutlich größer als bei Missspezifikation von β_0 ist. Daher wird im Folgenden der Parameter β_0 als beliebig, aber fest angesehen und für die Effektparameter ein Parameterintervall B vorgegeben. Es werden also standardisiert Maximin D - und c -optimale Designs bei Vorgabe eines Parameterbereiches $\tilde{B} = \{\beta_0\} \times B$ berechnet.

6.1 Standardisiert Maximin D -optimale Designs

In diesem Abschnitt wird das standardisierte Maximin-Kriterium für D -Optimalität betrachtet. Für das Regressionsmodell mit einer Kovariablen wird der Fall eines binären und eines stetigen Designraums untersucht. Die Berechnung der optimalen Designs erfolgt sowohl analytisch als auch numerisch. Bei dem multiplen Regressionsmodell mit zwei Kovariablen ist das Optimierungsproblem so komplex, dass die optimalen Designs nur noch numerisch bestimmt werden.

6.1.1 Regressionsmodell mit einer binären Kovariablen

Im Folgenden wird das Regressionsmodell mit Regressionsfunktion $\mathbf{f}(x) = (1, x)^T$ mit $x \in \mathbb{R}$ und Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ betrachtet. Das lokal D -optimale Design bei einem binären Designraum ist in Satz 5.4 gegeben. Es besitzt immer die Stützpunkte 0 und 1 , hängt aber über die D -optimalen Gewichte vom Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ ab. Um diese Abhängigkeit zu verringern, wird im folgenden Satz das standardisiert Maximin D -optimale Design bei Vorgabe eines Parameterbereiches $B = [b_0, b_1]$ für β_1 berechnet, wobei der Parameter β_0 beliebig, aber fest ist.

Satz 6.1

Es sei $\mathcal{X} = \{0, 1\}$. Für festes β_0 und $\beta_1 \in [b_0, b_1]$ mit $b_0 < b_1$ ist das standardisiert Maximin D -optimale Design ξ^* gegeben durch:

$$\xi^* = \frac{1}{2} \cdot (\xi_{\beta_0}^* + \xi_{\beta_1}^*) \quad (6.1)$$

Dabei sind $\xi_{\beta_0}^*$ und $\xi_{\beta_1}^*$ die lokal D -optimalen Designs für die Parameter $\beta_0 = (\beta_0, b_0)^T$ und $\beta_1 = (\beta_0, b_1)^T$.

Beweis:

Es sei $\lambda_j = \lambda_j(\beta_1) = \exp(\beta_0 + \beta_1 x_j)$ und $\delta_j = \delta_j(\beta_1) = 1 + \frac{m}{b} \cdot \lambda_j(\beta_1)$ für $j = 1, 2$, wobei $x_1 = 0$ und $x_2 = 1$ ist. Zur Vereinfachung der Notation wird als Kriteriumsfunktion die quadrierte Effizienz $\text{eff}_D(\xi; \beta)^2 = \det(\mathbf{M}(\xi; \beta)) / \det(\mathbf{M}(\xi_{\beta}^*; \beta))$ betrachtet, wobei ξ_{β}^* das lokal D -optimale Design für β ist. Nach Gleichung (5.12) ist die Determinante der Informationsmatrix für ein Design ξ mit Stützpunkten $x_1 = 0$ und $x_2 = 1$ und zugehörigen Gewichten w_1 und w_2 durch

$$\det(\mathbf{M}(\xi; \beta)) = \left(\frac{a}{b}\right)^2 \cdot \frac{w_1 \cdot w_2 \cdot \lambda_1 \cdot \lambda_2}{w_1 \cdot \delta_1 + w_2 \cdot \delta_2}$$

gegeben. Für das lokal D -optimale Design ξ_{β}^* für den binären Designraum mit D -optimalen Gewichten

$$w_{1,\beta}^* = \frac{\sqrt{\delta_2}}{\sqrt{\delta_1} + \sqrt{\delta_2}} \quad \text{und} \quad w_{2,\beta}^* = \frac{\sqrt{\delta_1}}{\sqrt{\delta_1} + \sqrt{\delta_2}}$$

gilt für die Determinante der Informationsmatrix nach Gleichung (5.13):

$$\det(\mathbf{M}(\xi_{\beta}^*; \beta)) = \left(\frac{a}{b}\right)^2 \cdot \frac{w_{1,\beta}^* \cdot w_{2,\beta}^* \cdot \lambda_1 \cdot \lambda_2}{w_{1,\beta}^* \cdot \delta_1 + w_{2,\beta}^* \cdot \delta_2} = \left(\frac{a}{b}\right)^2 \cdot \frac{\lambda_1 \cdot \lambda_2}{(\sqrt{\delta_1} + \sqrt{\delta_2})^2}.$$

Die Kriteriumsfunktion für standardisierte Maximin D -Optimalität für ein Design ξ mit Stützpunkten $x_1 = 0$ und $x_2 = 1$ und zugehörigen Gewichten w und $1 - w$ ergibt sich zu:

$$\begin{aligned} g(w, \beta_1) &= \frac{\det(\mathbf{M}(\xi; \beta))}{\det(\mathbf{M}(\xi_{\beta}^*; \beta))} = \frac{w \cdot (1 - w) \cdot \lambda_1 \cdot \lambda_2}{w \cdot \delta_1 + (1 - w) \cdot \delta_2} \cdot \left(\frac{\lambda_1 \cdot \lambda_2}{(\sqrt{\delta_1} + \sqrt{\delta_2})^2} \right)^{-1} \\ &= \frac{w \cdot (1 - w) \cdot (\sqrt{\delta_1} + \sqrt{\delta_2})^2}{w \cdot \delta_1 + (1 - w) \cdot \delta_2}. \end{aligned} \quad (6.2)$$

Unter Verwendung der lokal D -optimalen Gewichte $w_{1,\beta}^*$ und $w_{2,\beta}^*$ lässt sich die Funktion

$g(w, \beta_1)$ auch darstellen als:

$$g(w, \beta_1) = \frac{w \cdot (1 - w)}{w \cdot (w_{2,\beta}^*)^2 + (1 - w) \cdot (w_{1,\beta}^*)^2}. \quad (6.3)$$

Es wird nun gezeigt, dass die Kriteriumsfunktion $g(w, \beta_1)$ für beliebiges, aber festes w ihre Minima bezüglich β_1 im Intervall $[b_0, b_1]$ immer in der Menge der Randpunkte $\{b_0, b_1\}$ annimmt. Es gilt $\partial\delta_1/\partial\beta_1 = 0$ und $\partial\delta_2/\partial\beta_1 = (m/b) \cdot \lambda_2$. Ableiten der Kriteriumsfunktion $g(w, \beta_1)$ in Gleichung (6.2) nach β_1 ergibt:

$$\begin{aligned} & \frac{\partial g(w, \beta_1)}{\partial \beta_1} \\ &= w \cdot (1 - w) \cdot \left(\frac{(\sqrt{\delta_1} + \sqrt{\delta_2}) \cdot \frac{m}{b} \cdot \lambda_2}{\sqrt{\delta_2} \cdot (w \cdot \delta_1 + (1 - w) \cdot \delta_2)} - \frac{(\sqrt{\delta_1} + \sqrt{\delta_2})^2 \cdot (1 - w) \cdot \frac{m}{b} \cdot \lambda_2}{(w \cdot \delta_1 + (1 - w) \cdot \delta_2)^2} \right). \end{aligned}$$

Es gilt $\partial g(w, \beta_1)/\partial \beta_1 = 0$, genau dann wenn

$$\begin{aligned} 0 &= (w \cdot \delta_1 + (1 - w) \cdot \delta_2) - (\sqrt{\delta_1} + \sqrt{\delta_2}) \cdot (1 - w) \cdot \sqrt{\delta_2} \\ &= w \cdot \delta_1 - (1 - w) \cdot \sqrt{\delta_1} \cdot \sqrt{\delta_2} \end{aligned}$$

erfüllt ist. Umstellen liefert die Gleichung $\sqrt{\delta_2} = (w/(1 - w)) \cdot \sqrt{\delta_1}$ und mit Quadrieren folgt $\delta_2 = (w/(1 - w))^2 \cdot \delta_1$. Es gilt:

$$\begin{aligned} 1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1) &= \left(\frac{w}{1 - w} \right)^2 \cdot \left(1 + \frac{m}{b} \cdot \exp(\beta_0) \right) \\ \Leftrightarrow \exp(\beta_0 + \beta_1) &= \left(\left(\frac{w}{1 - w} \right)^2 \cdot \left(1 + \frac{m}{b} \cdot \exp(\beta_0) \right) - 1 \right) \cdot \frac{b}{m}. \end{aligned}$$

Für $(w/(1 - w))^2 > (1 + \frac{m}{b} \cdot \exp(\beta_0))^{-1}$ ergibt sich als eindeutige Lösung:

$$\beta_1^* = \ln \left(\left(\left(\frac{w}{1 - w} \right)^2 \cdot \left(1 + \frac{m}{b} \cdot \exp(\beta_0) \right) - 1 \right) \cdot \frac{b}{m} \right) - \beta_0.$$

Für $\beta_1 = \beta_1^*$ gilt der Zusammenhang $\sqrt{\delta_2} = (w/(1 - w)) \cdot \sqrt{\delta_1}$. Aufgelöst nach w ergibt sich $w = \sqrt{\delta_2}/(\sqrt{\delta_1} + \sqrt{\delta_2})$. Somit ist w gleich dem optimalen Gewicht $w_{1,\beta}^*$ des lokal D -optimalen Designs ξ_{β}^* für $\beta = (\beta_0, \beta_1^*)^T$ und es gilt $g(w, \beta_1^*) = 1$. Da $g(w, \beta_1) \leq 1$ gilt, liegt für festes w bei $\beta_1 = \beta_1^*$ ein Maximum vor. Da die Ableitung von $g(w, \beta_1)$ keine weiteren Nullstellen besitzt, existieren keine weiteren Extrema im offenen Intervall (b_0, b_1) . Somit ist $g(w, \beta_1)$ bezüglich β_1 unimodal und wird im Intervall $[b_0, b_1]$ auf dem

Rand minimal.

Für $(w/(1-w))^2 \leq (1 + \frac{m}{b} \cdot \exp(\beta_0))^{-1}$ besitzt die Ableitung der Kriteriumsfunktion nach β_1 keine Nullstelle und ist für festes w für alle $\beta_1 \in \mathbb{R}$ negativ. Somit wird $g(w, \beta_1)$ im Intervall $[b_0, b_1]$ bei b_1 minimal.

Die Kriteriumsfunktion besitzt ihre Minima bezüglich β_1 immer in der Menge $\{b_0, b_1\}$. Um die minimale Effizienz zu maximieren, wird $\min \{g(w, b_0), g(w, b_1)\}$ bezüglich w maximiert. Für die Ableitung von $g(w, b_i)$, $i = 0, 1$, nach w gilt:

$$\frac{\partial g(w, b_i)}{\partial w} = \left(\sqrt{\delta_1} + \sqrt{\delta_2} \right)^2 \cdot \left(\frac{(1 - 2 \cdot w)}{w \cdot \delta_1 + (1 - w) \cdot \delta_2} - \frac{w \cdot (1 - w) \cdot (\delta_1 - \delta_2)}{(w \cdot \delta_1 + (1 - w) \cdot \delta_2)^2} \right).$$

Es gilt $\partial g(w, b_i)/\partial w = 0$, genau dann wenn

$$\begin{aligned} 0 &= (1 - 2 \cdot w) \cdot (w \cdot \delta_1 + (1 - w) \cdot \delta_2) - w \cdot (1 - w) \cdot (\delta_1 - \delta_2) \\ &= -w^2 \cdot (\delta_1 - \delta_2) - 2 \cdot w \cdot \delta_2 + \delta_2 \end{aligned}$$

erfüllt ist. Für $b_i = 0$ gilt $\delta_1 = \delta_2$ und als Lösung der Gleichung ergibt sich $w_0 = 1/2$. Für $b_i \neq 0$ gilt $\delta_1 \neq \delta_2$ und die quadratische Gleichung besitzt die Lösungen:

$$\begin{aligned} w_{b_i;+,-} &= \frac{-\delta_2 \pm \sqrt{\delta_2^2 + (\delta_1 - \delta_2) \cdot \delta_2}}{\delta_1 - \delta_2} = \frac{-\delta_2 \pm \sqrt{\delta_1 \cdot \delta_2}}{\delta_1 - \delta_2} = \frac{\sqrt{\delta_2} \cdot (-\sqrt{\delta_2} \pm \sqrt{\delta_1})}{(\sqrt{\delta_1} + \sqrt{\delta_2}) \cdot (\sqrt{\delta_1} - \sqrt{\delta_2})} \\ &= \frac{\pm \sqrt{\delta_2}}{\sqrt{\delta_1} \pm \sqrt{\delta_2}}. \end{aligned}$$

Für $b_i < 0$ gilt $\delta_1 > \delta_2$ und für $b_i > 0$ gilt $\delta_1 < \delta_2$. Die Lösung $w_{b_i;-}$ ist für $b_i < 0$ negativ und für $b_i > 0$ größer als 1. Somit gilt $w_{b_i;-} \notin [0, 1]$. Die Lösung $w_{b_i;+}$ ist gleich dem optimalen Gewicht w_{1,β_i}^* des lokal D -optimalen Designs $\xi_{b_i}^* = \xi_{\beta_i}^*$ für $\beta_i = (\beta_0, b_i)^T$. Für $b_i = 0$ gilt $w_{b_i;+} = 1/2 = w_0$ und ebenfalls $w_{b_i;+} = w_{1,\beta_i}^*$ mit $\beta_i = (\beta_0, 0)^T$. Somit ist für $w = w_{b_i;+}$ die Effizienz bei $\beta_i = (\beta_0, b_i)^T$ gleich 1 und es handelt sich um ein Maximum von $g(w, b_i)$. Daher ist $g(w, b_i)$ bezüglich w unimodal im Intervall $(0, 1)$.

Die Funktionen $k(\beta_1) = \sqrt{\delta_2(\beta_1)}$ und $h(z) = z / (\sqrt{\delta_1} + z)$ für $z \geq 0$ sind streng monoton wachsend. Somit ist auch die Komposition

$$h(k(\beta_1)) = \frac{\sqrt{\delta_2(\beta_1)}}{\sqrt{\delta_1} + \sqrt{\delta_2(\beta_1)}}$$

streng monoton wachsend und für $b_0 < b_1$ gilt $w_{b_0;+} < w_{b_1;+}$. Für $w \leq w_{b_0;+}$ sind $g(w, b_0)$ und $g(w, b_1)$ streng monoton wachsend in w und für $w \geq w_{b_1;+}$ sind $g(w, b_0)$ und

$g(w, b_1)$ streng monoton fallend in w . Somit wird $\min \{g(w, b_0), g(w, b_1)\}$ bezüglich w im Intervall $[w_{b_0;+}, w_{b_1;+}]$ maximal. In diesem Intervall ist $g(w, b_0)$ streng monoton fallend und $g(w, b_1)$ streng monoton wachsend. Da außerdem $g(w_{b_0;+}, b_0) = g(w_{b_1;+}, b_1) = 1$ gilt, existiert genau ein $w^* \in [w_{b_0;+}, w_{b_1;+}]$ mit $g(w^*, b_0) = g(w^*, b_1)$. Für dieses w^* wird das Minimum von $g(w, b_0)$ und $g(w, b_1)$ maximal. Dieser Sachverhalt wird in Abbildung 6.2 grafisch veranschaulicht.

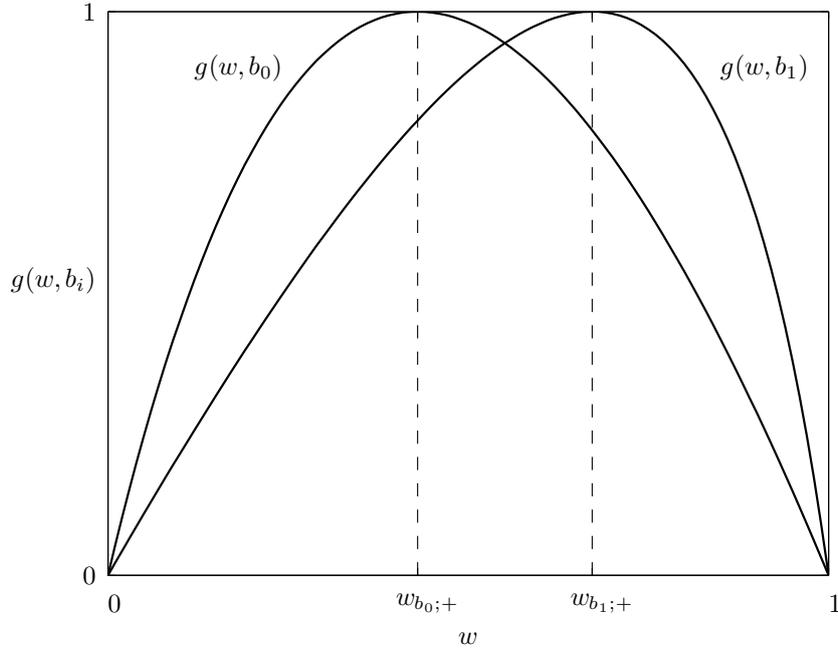


Abbildung 6.2: Kriteriumsfunktion bei $\beta_1 = b_0$ und $\beta_1 = b_1$ in Abhängigkeit vom Gewicht w für den Stützpunkt $x_1 = 0$

Mit der Darstellung von $g(w, b_i)$ in Gleichung (6.3) berechnet sich das optimale Gewicht w^* somit als Lösung der folgenden Gleichung:

$$\frac{w^* \cdot (1 - w^*)}{w^* \cdot (1 - w_{1,\beta_0}^*)^2 + (1 - w^*) \cdot (w_{1,\beta_0}^*)^2} = \frac{w^* \cdot (1 - w^*)}{w^* \cdot (1 - w_{1,\beta_1}^*)^2 + (1 - w^*) \cdot (w_{1,\beta_1}^*)^2}.$$

Dabei ist w_{1,β_i}^* das Gewicht des lokal D -optimalen Designs $\xi_{\beta_i}^*$ für $\beta_i = (\beta_0, b_i)$, $i = 1, 2$. Äquivalent gilt:

$$\begin{aligned} w^* - 2 \cdot w^* \cdot w_{1,\beta_0}^* + (w_{1,\beta_0}^*)^2 &= w^* - 2 \cdot w^* \cdot w_{1,\beta_1}^* + (w_{1,\beta_1}^*)^2 \\ \Leftrightarrow w^* &= \frac{(w_{1,\beta_1}^*)^2 - (w_{1,\beta_0}^*)^2}{2 \cdot (w_{1,\beta_1}^* - w_{1,\beta_0}^*)} = \frac{w_{1,\beta_1}^* + w_{1,\beta_0}^*}{2}. \end{aligned}$$

Somit ist das Design $\xi^* = \frac{1}{2} \cdot (\xi_{\beta_0}^* + \xi_{\beta_1}^*)$ standardisiert Maximin D -optimal. \square

Als Grenzfall des lokal D -optimalen Designs für das Poisson-Gamma-Modell folgt für $b \rightarrow \infty$ das D -optimale Design für das Poisson-Modell. Ebenso ergibt sich beim binären Designraum $\mathcal{X} = \{0, 1\}$ aus dem standardisiert Maximin D -optimalen Design für das Poisson-Gamma-Modell für $b \rightarrow \infty$ das standardisiert Maximin D -optimale Design für das Poisson-Modell. Es weist den Stützpunkten 0 und 1 gleiche Gewichte $1/2$ zu und ist somit identisch zum lokal D -optimalen Design, da dieses nicht vom Parametervektor β abhängt. Das standardisiert Maximin D -optimale Design für das Poisson-Modell kann daher ebenfalls als Mittelwert der für $\beta_0 = (\beta_0, b_0)^T$ und $\beta_1 = (\beta_1, b_1)^T$ lokal D -optimalen Designs dargestellt werden.

6.1.2 Regressionsmodell mit einer stetigen Kovariablen

Es wird nun ein stetiger Designraum $\mathcal{X} = [u, v]$ mit $u < v$ betrachtet. Die Regressionsfunktion ist weiterhin durch $f(x) = (1, x)^T$ und der Parametervektor durch $\beta = (\beta_0, \beta_1)^T$ gegeben. Der Parameter β_0 wird als beliebig, aber fest vorgegeben angesehen. Der Parameterbereich $B = [b_0, b_1]$ für β_1 wird so gewählt, dass alle $\beta_1 \in B$ das gleiche Vorzeichen besitzen. Weiterhin gelte für den Designraum, dass $v = 0$ für positive β_1 und $u = 0$ für negative β_1 ist. Dann ist nach Satz 5.6 ein Stützpunkt des lokal D -optimalen Designs bei 0. Ist \mathcal{X} groß genug, ergibt sich der zweite Stützpunkt als $-z^*/\beta_1$, wobei z^* die Lösung der Gleichung

$$0 = z - 2 \cdot \left(1 + \frac{\frac{m}{b} \cdot \exp(\beta_0 - z)}{1 + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 - z)} \cdot \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0)}} \right) \quad (6.4)$$

ist. Diese Lösung z^* hängt nicht von β_1 ab, der Stützpunkt $-z^*/\beta_1$ des lokal D -optimalen Designs hingegen wird von β_1 beeinflusst.

Zunächst wird untersucht, wie sich das standardisiert Maximin D -optimale Design bei Transformation des Intervalls $[b_0, b_1]$ der möglichen Werte für β_1 verhält. Dabei wird für eine Konstante $c \in \mathbb{R}$ und eine Menge M die Notation $c \cdot M = \{c \cdot m : m \in M\}$ verwendet.

Satz 6.2

Es seien $\mathcal{X} = [u, v]$, β_0 fest und $\beta_1 \in [b_0, b_1]$ mit $b_0 < b_1$, wobei $b_0 > 0$ und $v = 0$ oder $b_1 < 0$ und $u = 0$ gelte. Außerdem sei die Ungleichung $z^* \leq \min\{|b_0|, |b_1|\} \cdot (v - u)$ erfüllt, wobei $z^* > 0$ die Lösung der Gleichung (6.4) ist. Das Design

$$\xi^* = \left\{ \begin{array}{ccc} x_1^* & \dots & x_l^* \\ w_1^* & \dots & w_l^* \end{array} \right\} \quad (6.5)$$

sei standardisiert Maximin D -optimal auf $\mathcal{X} = [u, v]$.

Es sei $c \neq 0$. Dann ist für $\beta_1 \in c \cdot [b_0, b_1]$ und Designraum $\mathcal{X}_c = \frac{1}{c} \cdot [u, v]$ das Design

$$\xi_c^* = \left\{ \begin{array}{ccc} \frac{x_1^*}{c} & \cdots & \frac{x_l^*}{c} \\ w_1^* & \cdots & w_l^* \end{array} \right\} \quad (6.6)$$

standardisiert Maximin D -optimal. Außerdem besitzen die Designs ξ^* und ξ_c^* die gleiche minimale Effizienz.

Der Beweis zu Satz 6.2 befindet sich im Anhang.

Für das Poisson-Modell wurde unter gewissen Voraussetzungen von Konstantinou et al. (2014) das standardisiert Maximin D -optimale Design für einen stetigen Designraum in der Klasse der 2-Punkt-Designs hergeleitet. Dabei war durch die Beschränkung auf 2-Punkt-Designs bereits bekannt, dass die Gewichte des standardisiert Maximin D -optimalen Designs gleich und somit durch $1/2$ gegeben sind. Daher waren nur noch die Stützpunkte zu optimieren.

Bei Betrachtung der Poisson-Gamma-Modells mit einem binären Designraum sind die Stützpunkte des standardisiert Maximin D -optimalen Designs bereits bekannt und nur die Gewichte müssen optimiert werden.

Hingegen müssen für die Berechnung des standardisiert Maximin D -optimalen Designs für das Poisson-Gamma-Modell mit einem stetigen Designraum sowohl die optimalen Stützpunkte als auch die optimalen Gewichte bestimmt werden, was die Komplexität des Optimierungsproblems erhöht. Im folgenden Satz wird das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs für einen stetigen Designraum angegeben, wobei der Beweis sich im Anhang befindet.

Satz 6.3

Es sei $\mathcal{X} = [u, v]$, β_0 fest und $\beta_1 \in [b_0, b_1]$ mit $b_0 < b_1$, wobei $b_0 > 0$ und $v = 0$ oder $b_1 < 0$ und $u = 0$ gelte. Außerdem sei die Ungleichung $z^* \leq \min\{|b_0|, |b_1|\} \cdot (v - u)$ erfüllt, wobei $z^* > 0$ die Lösung der Gleichung 6.4 ist. Es seien

$$w(x) = \frac{1}{1 - \frac{1 + \frac{m}{b} \cdot \exp(\beta_0)}{1 + \frac{m}{b} \cdot \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\exp(\beta_0 + b_1 x)} - \frac{b_1^2}{\exp(\beta_0 + b_0 x)}}}} \quad (6.7)$$

und

$$h(x) = \frac{m}{b} \cdot \exp(\beta_0 + (b_0 + b_1)x) \cdot (b_1^2 - b_0^2) + \exp(b_1x) \cdot b_1^2 - \exp(b_0x) \cdot b_0^2. \quad (6.8)$$

Zudem sei $\tilde{x} = 2 \cdot \ln(b_1/b_0)/(b_0 - b_1)$. Für $b_1 < 0$ sei $x_0 > \tilde{x}$ die eindeutige Nullstelle von $h(x)$ im Intervall $(0, \infty)$ und es sei $I = (\tilde{x}, x_0)$. Für $b_0 > 0$ sei $x_0 < \tilde{x}$ die eindeutige Nullstelle von $h(x)$ im Intervall $(-\infty, 0)$ und es sei $I = (x_0, \tilde{x})$. Außerdem sei

$$x^* = \arg \max_{x \in I} \frac{w(x) \cdot (1 - w(x)) \cdot x^2}{\frac{m}{b} \cdot w(x) \cdot (\exp(-b_0x) - 1) + \exp(-\beta_0 - b_0x) + \frac{m}{b}}. \quad (6.9)$$

Falls $x^* \in \mathcal{X}$ gilt, dann ist das standardisiert Maximin D -optimale Design ξ^* in der Klasse der 2-Punkt-Designs gegeben durch:

$$\xi^* = \left\{ \begin{array}{cc} 0 & x^* \\ w(x^*) & 1 - w(x^*) \end{array} \right\}. \quad (6.10)$$

Bei allen berechneten Beispielen war die Bedingung $x^* \in \mathcal{X}$ in Satz 6.3 stets erfüllt. Daher scheint diese Bedingung im Allgemeinen keine Einschränkung darzustellen. Die Bedingung $z^* \leq \min\{|b_0|, |b_1|\} \cdot (v - u)$ sichert, dass die lokal D -optimalen Designs immer die Stützpunkte 0 und $-z^*/\beta_1$ besitzen. Unter dieser Voraussetzung ist die Effizienz bezüglich β_1 unimodal und wird auf dem Rand von $[b_0, b_1]$ minimal.

In der folgenden Bemerkung wird der Fall betrachtet, dass der Parameterbereich B kein Intervall ist.

Bemerkung 6.4

Es sei $B \subseteq [b_0, b_1]$ mit $b_0, b_1 \in B$. In den Beweisen von Satz 6.1 für das standardisiert Maximin D -optimale Design bei einem binären Designraum und Satz 6.3 für das standardisiert Maximin D -optimale Design bei einem stetigen Designraum wurde gezeigt, dass die Kriteriumsfunktion für ein 2-Punkt-Design bezüglich β_1 unimodal auf $[b_0, b_1]$ ist und dass die Effizienz des optimalen Designs bei b_0 und b_1 minimal wird. Somit sind die standardisiert Maximin D -optimalen Designs auf $B \subseteq [b_0, b_1]$ mit $b_0, b_1 \in B$ und $[b_0, b_1]$ identisch und durch Satz 6.1 und Satz 6.3 gegeben. Insbesondere besitzt das standardisiert Maximin D -optimale Design für den Parameterbereich $B = \{b_0, b_1\}$ die gleiche Effizienz an den beiden möglichen Parameterwerten b_0 und b_1 .

Um Satz 6.3 zu illustrieren, wird im folgenden Beispiel für einen exemplarischen Parameterbereich das standardisiert Maximin D -optimale Design im Poisson-Gamma-Modell berechnet.

Beispiel 6.5

Es sei $\mathcal{X} = [0, 3]$ der Designraum und $\mathbf{f}(x) = (1, x)^T$ die Regressionsfunktion. Für die Komponenten des Parametervektors $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ gelte $\beta_0 = 0$ und $\beta_1 \in [b_0, b_1]$. Weiterhin seien $m = 10$ und $b = 1$.

Der Parameterbereich sei $[b_0, b_1] = [-3, -1]$. Das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs folgt aus Satz 6.3. Die Lösung von Gleichung (6.4) ergibt sich zu $z^* = 2.341$. Außerdem ist $\tilde{x} = \ln(3) = 1.099$ und die Nullstelle von $h(x)$ in Gleichung (6.8) ist durch $x_0 = \ln(5) = 1.609$ gegeben. Somit ist die in Abbildung 6.3 dargestellte minimale Effizienz für ein Design mit Stützpunkten 0 und x sowie Gewichten $w(x)$ und $1 - w(x)$ auf dem Intervall $I = (\tilde{x}, x_0) = (\ln(3), \ln(5)) = (1.099, 1.609)$ bezüglich x zu maximieren. Die Lösung ist durch Gleichung (6.9) gegeben, bei der bezüglich x konstante Faktoren weggelassen wurden, und ergibt sich zu $x^* = 1.266$. Die optimalen Gewichte zu den Stützpunkten 0 und 1.266 berechnen sich zu $w(x^*) = w(1.266) = 0.325$ und $1 - w(x^*) = 1 - w(1.266) = 0.675$. Somit ist das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs durch

$$\xi^* = \begin{Bmatrix} 0 & 1.266 \\ 0.325 & 0.675 \end{Bmatrix}$$

gegeben. In Abbildung 6.4 wird die Effizienz von ξ^* für $\beta_1 \in [-3, -1]$ grafisch dargestellt. Das Design ξ^* besitzt eine minimale Effizienz von 0.825 und eine maximale Effizienz von 0.998. Somit ist ξ^* für kein $\beta_1 \in [-3, -1]$ lokal D -optimal.

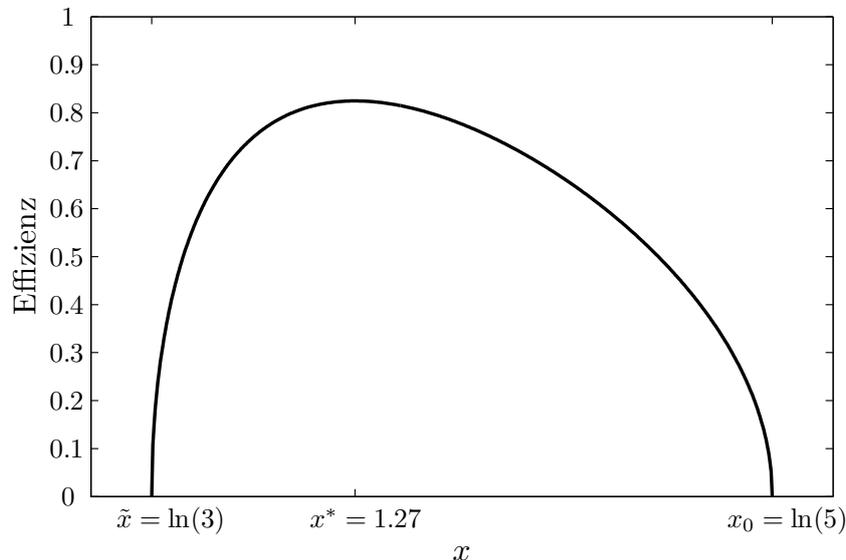


Abbildung 6.3: Minimale Effizienz für ein Design mit Stützpunkten 0 und x sowie Gewichten $w(x)$ und $1 - w(x)$, welche bei $b_0 = -3$ und $b_1 = -1$ angenommen wird

Die Menge der Parameter β_1 , bei denen die Effizienz von ξ^* minimal wird, ist durch $\mathcal{N}(\xi^*) = \{-3, -1\}$ gegeben. Mit der Wahrscheinlichkeitsverteilung π auf $\mathcal{N}(\xi^*)$, welche als $\pi(-3) = 0.403$ und $\pi(-1) = 0.597$ gewählt wird, gilt

$$\begin{aligned}
 d(\xi^*, x) &= \int_{\mathcal{N}(\xi^*)} \left(\exp(\beta_1 x) \cdot \mathbf{f}(x)^T \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1} \mathbf{f}(x) \right. \\
 &\quad \left. - \text{Spur}(\mathbf{M}(\xi^*; \boldsymbol{\beta}) \mathbf{M}_{\text{Po}}(\xi^*; \boldsymbol{\beta})^{-1}) \right) \pi(d\beta_1) \\
 &\leq 0
 \end{aligned}$$

für alle $x \in \mathcal{X}$, was numerisch gezeigt werden kann und in Abbildung 6.4 grafisch veranschaulicht wird. Somit ist nach dem Äquivalenzsatz 4.11 das Design ξ^* standardisiert Maximin D -optimal in der Klasse aller Designs.

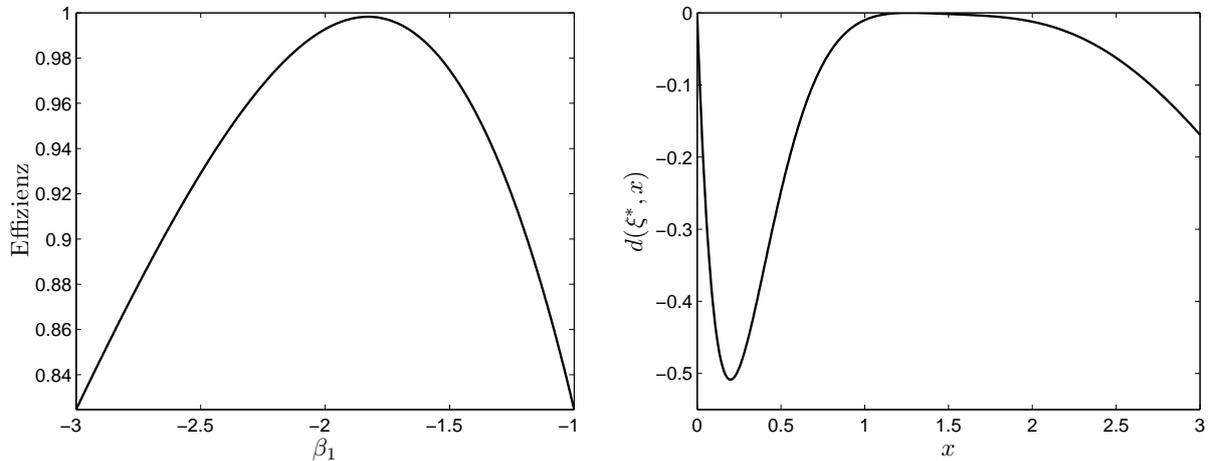


Abbildung 6.4: Effizienz (links) und Sensitivitätsfunktion (rechts) für das Design ξ^*

Es wird untersucht, wie sich die standardisiert Maximin D -optimalen Designs bei Variation des Parameterbereichs $B = [b_0, b_1]$ für β_1 verändern. Dazu wird die Situation aus Beispiel 6.5 betrachtet. Für die Komponenten des Parametervektors $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ gelte $\beta_0 = 0$ und $\beta_1 \in [b_0, b_1]$. Während die obere Intervallgrenze weiterhin bei $b_1 = -1$ ist, wird die untere Intervallgrenze b_0 von -1.5 in 0.5 -Schritten bis -41 verändert.

Die standardisiert Maximin D -optimalen 2-Punkt-Designs wurden mit Satz 6.3 berechnet. Ab einer gewissen Größe des Parameterbereichs sind die 2-Punkt-Designs nicht mehr standardisiert Maximin D -optimal in der Klasse aller Designs. Dann sind Designs mit mehr als zwei Stützpunkten optimal, welche numerisch bestimmt wurden. Dazu wird der folgende Algorithmus verwendet.

Algorithmus 6.6 (Berechnung standardisiert Maximin D -optimaler Designs)**Eingabe:**Fehlerschranke ε , Konstante $\lambda \in (0, \frac{1}{2}]$ Designraum \mathcal{X} , Parameterbereich B Intervalle $I_1 = [c_1, d_1], \dots, I_n = [c_n, d_n]$ mit $I_i \cap I_j = \emptyset, i, j \in \{1, \dots, n\}, i \neq j$ **Ablauf:**

Führe wiederholt aus:

$$\theta_j = \frac{c_j + d_j}{2}, j = 1, \dots, n$$

$$\boldsymbol{\theta}_j = (\beta_0, \theta_j)^T, j = 1, \dots, n$$

$$\xi^* = \arg \max_{\xi} \text{eff}_D(\mathbf{M}(\xi; \boldsymbol{\theta}_1))$$

$$\text{u.d.N. } \text{eff}_D(\mathbf{M}(\xi; \boldsymbol{\theta}_1)) = \dots = \text{eff}_D(\mathbf{M}(\xi; \boldsymbol{\theta}_n))$$

$$(\beta_0, \theta_{j,\min})^T = \arg \min_{\boldsymbol{\beta}=(\beta_0, \beta_1)^T, \beta_1 \in I_j} \text{eff}_D(\xi^*; \boldsymbol{\beta}), j = 1, \dots, n$$

Falls $\max_{j=1, \dots, n} |\theta_{j,\min} - \theta_j| < \varepsilon$ gilt, beende den Algorithmus.Für $j = 1, \dots, n$:

$$\text{Falls } \theta_{j,\min} < \theta_j: d_j = d_j - \lambda \cdot (d_j - c_j)$$

$$\text{Falls } \theta_{j,\min} > \theta_j: c_j = c_j + \lambda \cdot (d_j - c_j)$$

Ausgabe:Standardisiert Maximin D -optimales Design ξ^*

Bei Algorithmus 6.6 werden disjunkte Intervalle $I_j \subseteq B, j = 1, \dots, n$, vorgegeben, in denen jeweils genau ein Parameter β_1 liegt, für den das standardisiert Maximin D -optimale Design ξ^* seine minimale Effizienz annimmt. Insbesondere wird somit $|\mathcal{N}(\xi^*)|$ vorgegeben. In jedem Iterationsschritt des Algorithmus wird die D -Effizienz unter der Nebenbedingung, dass Gleichheit der D -Effizienzen an den Mittelpunkten der Intervalle I_j gilt, maximiert. Für das berechnete Design wird in jedem Intervall I_j das β_1 bestimmt, für welches die D -Effizienz minimal ist. Liegt das Minimum links von der Intervallmitte, wird die rechte Intervallgrenze verkleinert. Liegt das Minimum rechts von der Intervallmitte, wird die linke Intervallgrenze vergrößert. Somit wird in jedem Iterationsschritt die Länge der Intervalle I_j um den Faktor λ verkleinert. Beispielsweise führt der Fall $\lambda = 1/2$ zu einer Intervallhalbierung. Je größer $|\mathcal{N}(\xi^*)|$ ist, desto kleiner sollte λ gewählt werden, da die Minimalstellen der D -Effizienz der berechneten Designs stärker variieren können. Es zeigt sich, dass der Mittelpunkt der Intervalle gegen die minimalen β_1 des standardisiert Maximin D -optimalen Designs ξ^* konvergiert. Der Algorithmus wird beendet, wenn der Abstand der Mittelpunkte der Intervalle zu den minimalen β_1 kleiner als eine vorgegebene Schranke ε ist. Das berechnete optimale Design ξ^* ist dann das

standardisiert Maximin D -optimale Design.

In Abbildung 6.5 sind die Stützpunkte der berechneten standardisiert Maximin D -optimalen Designs in Abhängigkeit von der Intervalllänge $b_1 - b_0$ des Parameterbereichs $[b_0, b_1]$ dargestellt, wobei die Farbe der Punkte ihr jeweiliges Gewicht angibt. Weiterhin wird die Anzahl $|\mathcal{N}| = |\mathcal{N}(\xi^*)|$ der Parameter, bei denen die minimale Effizienz des jeweiligen optimalen Designs ξ^* angenommen wird, angegeben.

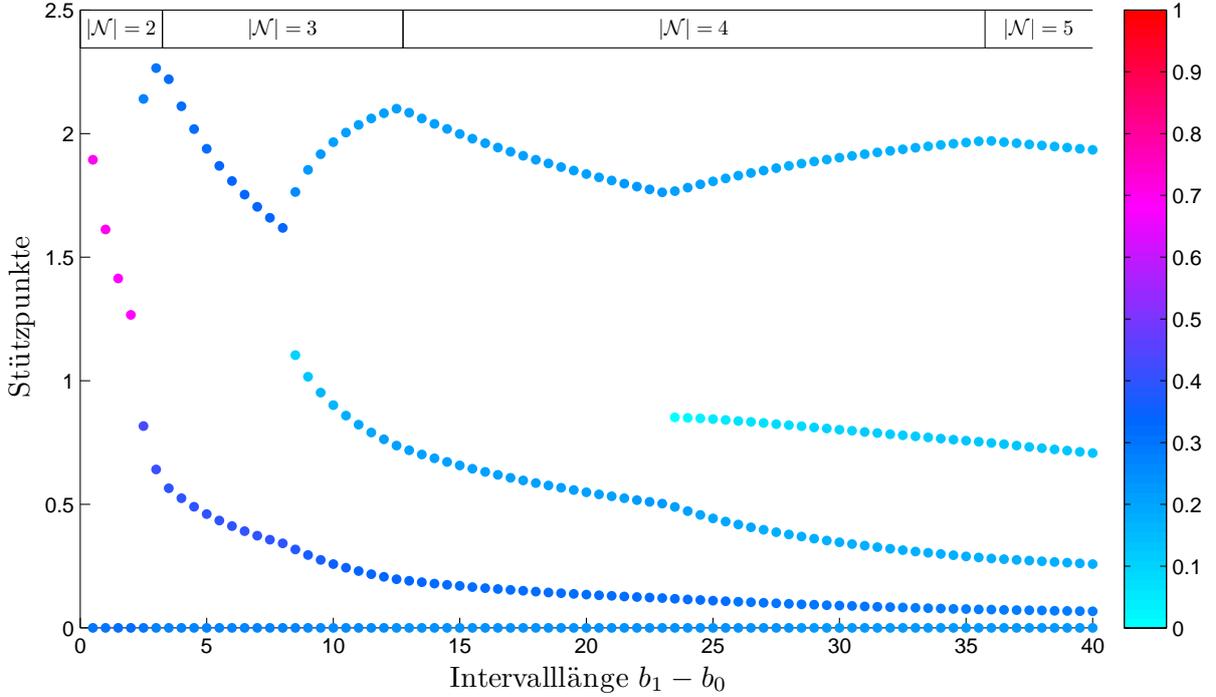


Abbildung 6.5: Standardisiert Maximin D -optimale Designs in Abhängigkeit von der Intervalllänge des Parameterbereichs $[b_0, b_1]$ mit $b_1 = -1$

Bei Intervalllängen $b_1 - b_0 \leq 2$ des Parameterbereichs sind die 2-Punkt-Designs aus Satz 6.3 standardisiert Maximin D -optimal. Für $2.5 \leq b_1 - b_0 \leq 8$ ist ein 3-Punkt-Design optimal, für $8.5 \leq b_1 - b_0 \leq 23$ ist ein 4-Punkt-Design optimal und für $23.5 \leq b_1 - b_0 \leq 40$ ist ein 5-Punkt-Design optimal. Die Effizienz wird bei den optimalen 2-Punkt-Designs immer an beiden Randpunkten b_0 und b_1 des Parameterbereichs minimal und es gilt $|\mathcal{N}(\xi^*)| = 2$. Bei optimalen Designs mit $l \geq 3$ Stützpunkten ist ebenfalls $b_0, b_1 \in \mathcal{N}(\xi^*)$, die Effizienz kann aber auch bei weiteren Parameterwerten β_1 minimal sein. Tritt bei Erhöhung der Intervalllänge das erste Mal ein optimales Design mit $l \geq 3$ Stützpunkten auf, so gilt $|\mathcal{N}(\xi^*)| = l - 1$. Wird die Intervalllänge weiter vergrößert, erhöht sich die Anzahl der Parameterwerte, an denen die Effizienz minimal ist, zu $|\mathcal{N}(\xi^*)| = l$.

In Tabelle 6.1 sind für eine Auswahl an Parameterbereichen die standardisiert Maximin D -optimalen Designs und ihre minimale Effizienz angegeben. Weiterhin ist die Verteilung π auf $\mathcal{N}(\xi^*)$ angegeben, mit der die Optimalität des jeweiligen Designs unter Verwen-

$[b_0, b_1]$	Optimales Design	Min. Eff.	π
$[-4, -1]$	$\begin{Bmatrix} 0 & 0.64 & 2.26 \\ 0.28 & 0.41 & 0.31 \end{Bmatrix}$	0.776	$\begin{Bmatrix} -4 & -1 \\ 0.45 & 0.55 \end{Bmatrix}$
$[-6, -1]$	$\begin{Bmatrix} 0 & 0.46 & 1.94 \\ 0.27 & 0.40 & 0.33 \end{Bmatrix}$	0.747	$\begin{Bmatrix} -6 & -2.69 & -1 \\ 0.35 & 0.20 & 0.44 \end{Bmatrix}$
$[-12, -1]$	$\begin{Bmatrix} 0 & 0.23 & 0.82 & 2.04 \\ 0.25 & 0.34 & 0.19 & 0.21 \end{Bmatrix}$	0.689	$\begin{Bmatrix} -12 & -3.80 & -1 \\ 0.30 & 0.33 & 0.37 \end{Bmatrix}$
$[-18, -1]$	$\begin{Bmatrix} 0 & 0.15 & 0.61 & 1.93 \\ 0.24 & 0.32 & 0.22 & 0.22 \end{Bmatrix}$	0.666	$\begin{Bmatrix} -18 & -6.80 & -3.45 & -1 \\ 0.26 & 0.19 & 0.22 & 0.33 \end{Bmatrix}$
$[-35, -1]$	$\begin{Bmatrix} 0 & 0.08 & 0.30 & 0.77 & 1.95 \\ 0.23 & 0.29 & 0.18 & 0.13 & 0.17 \end{Bmatrix}$	0.631	$\begin{Bmatrix} -35 & -11.51 & -3.87 & -1 \\ 0.23 & 0.25 & 0.24 & 0.28 \end{Bmatrix}$
$[-41, -1]$	$\begin{Bmatrix} 0 & 0.07 & 0.26 & 0.71 & 1.93 \\ 0.23 & 0.29 & 0.18 & 0.13 & 0.17 \end{Bmatrix}$	0.624	$\begin{Bmatrix} -41 & -14.00 & -7.94 & -3.81 & -1 \\ 0.23 & 0.21 & 0.08 & 0.21 & 0.27 \end{Bmatrix}$

Tabelle 6.1: Standardisiert Maximin D -optimale Designs für verschiedene Parameterbereiche $B = [b_0, b_1]$

derung des Äquivalenzsatzes 4.11 nachgewiesen werden kann. Dies geschieht grafisch in den Abbildungen 6.6 und 6.7, in denen auch der Effizienzverlauf für die verschiedenen Designs dargestellt wird.

In Abbildung 6.8 wird die minimale Effizienz der standardisiert Maximin D -optimalen Designs für die verschiedenen Parameterbereiche $[b_0, b_1]$ dargestellt. Die minimale Effizienz sinkt mit größer werdendem Parameterbereich, wobei die Abnahme der minimalen Effizienz immer geringer wird.

Beispiel 6.7

Die 2-Punkt-Designs aus Satz 6.3 sind ab einer gewissen Größe des Parameterbereichs nicht mehr standardisiert Maximin D -optimal. Daher wird in diesem Beispiel die Effizienz dieser 2-Punkt-Designs mit den optimalen Designs verglichen. Dazu wird erneut die Situation aus Beispiel 6.5 betrachtet.

a) Der Parameterbereich wird als $[b_0, b_1] = [-4, -1]$ gewählt. Das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs ist nach Satz 6.3 durch

$$\xi^* = \begin{Bmatrix} 0 & 1.057 \\ 0.339 & 0.661 \end{Bmatrix}$$

gegeben. Die minimale Effizienz von ξ^* beträgt 0.741 und wird bei $\beta_1 = b_0 = -4$ und $\beta_1 = b_1 = -1$ angenommen. Das Design ξ^* ist jedoch nicht standardisiert Maximin D -optimal in der Klasse aller Designs. Das global optimale Design ist das in Tabelle 6.1 angegebene 3-Punkt-Design mit einer minimalen Effizienz von 0.776. Somit besitzt das

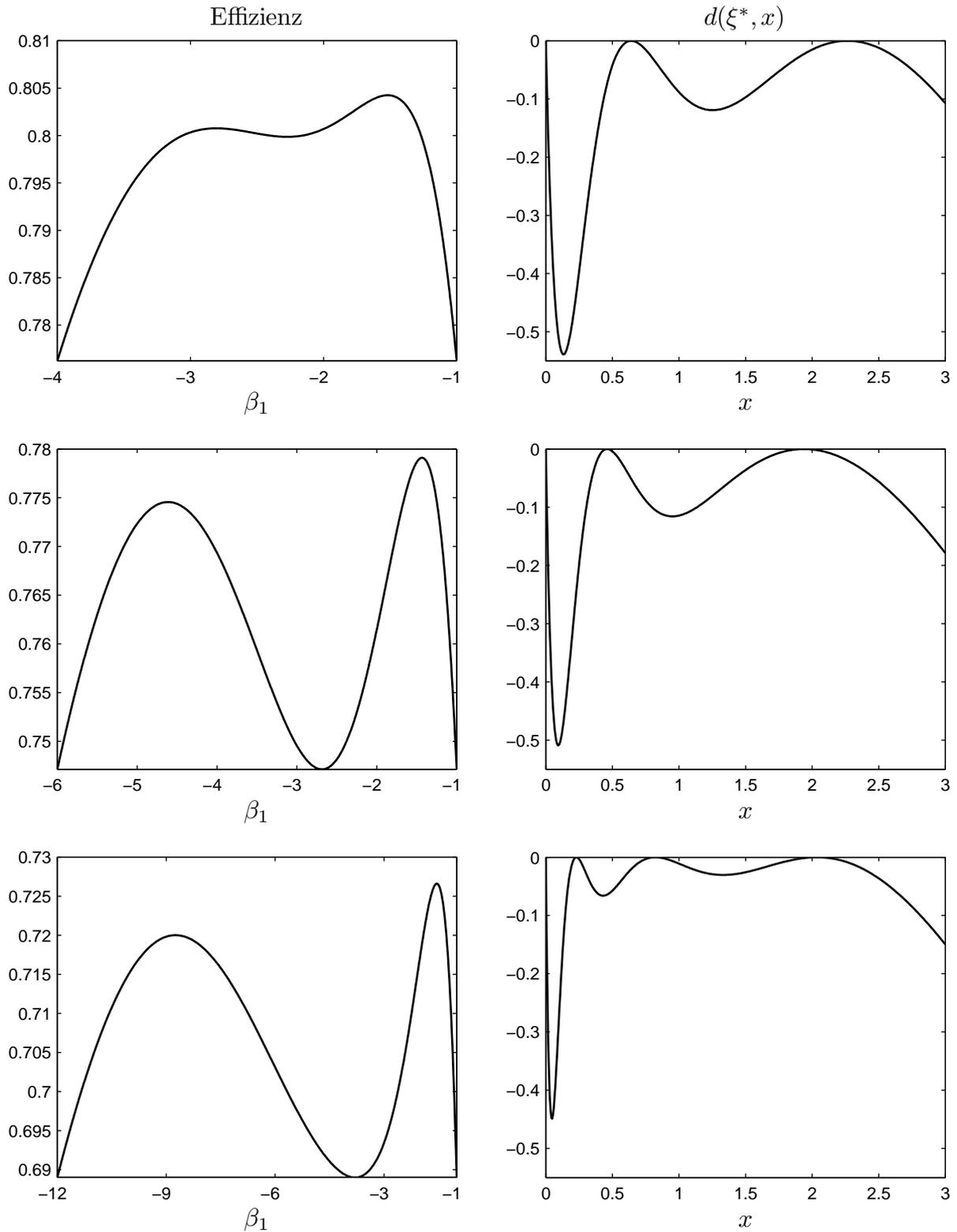


Abbildung 6.6: Effizienz (links) und Sensitivitätsfunktion $d(\xi^*, x)$ (rechts) für die standardisiert Maximin D -optimalen Designs in Tabelle 6.1 für die Parameterbereiche $[-4, -1]$, $[-6, -1]$ und $[-12, -1]$ (von oben nach unten)

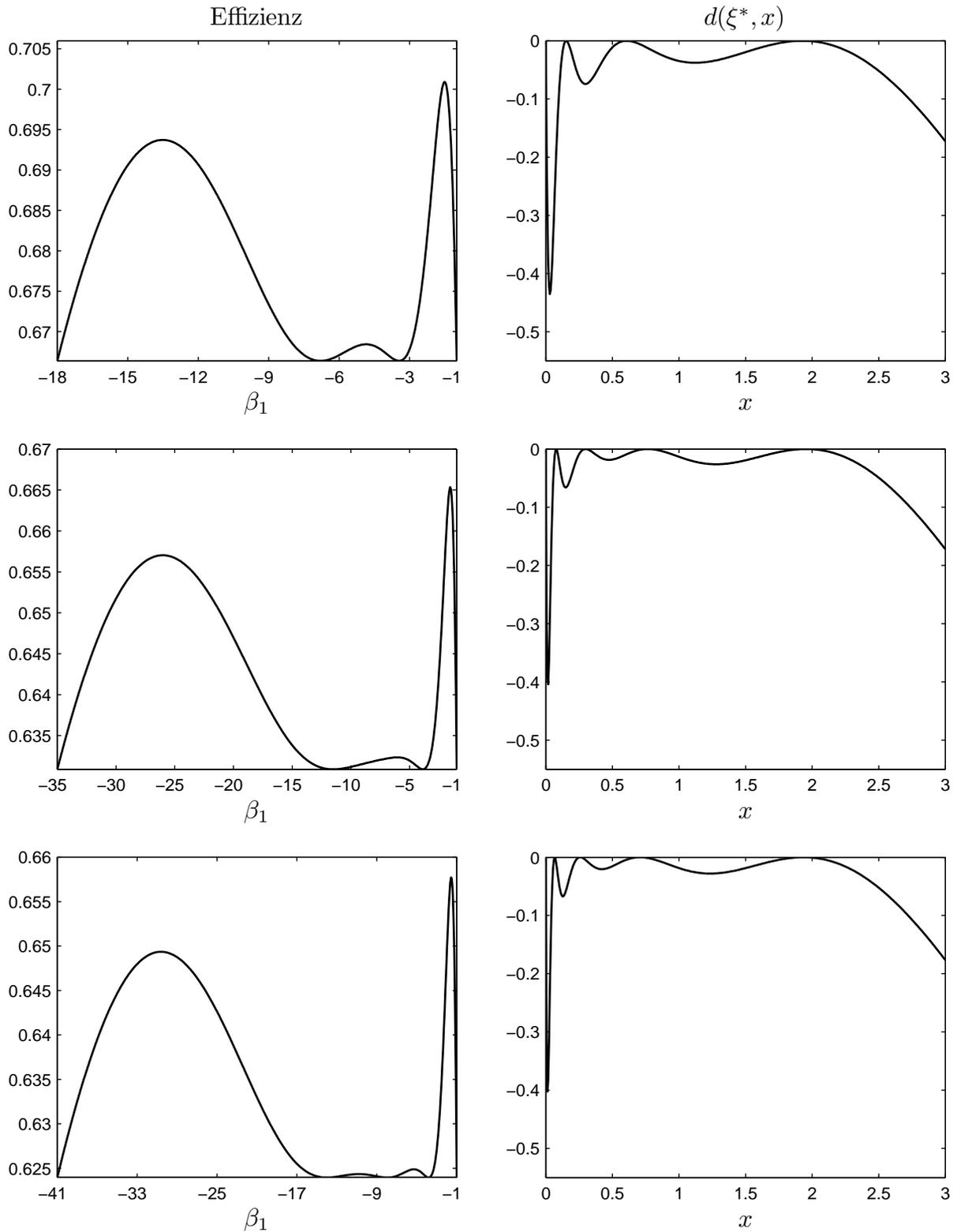


Abbildung 6.7: Effizienz (links) und Sensitivitätsfunktion $d(\xi^*, x)$ (rechts) für die standardisiert Maximin D -optimalen Designs in Tabelle 6.1 für die Parameterbereiche $[-18, -1]$, $[-35, -1]$ und $[-41, -1]$ (von oben nach unten)

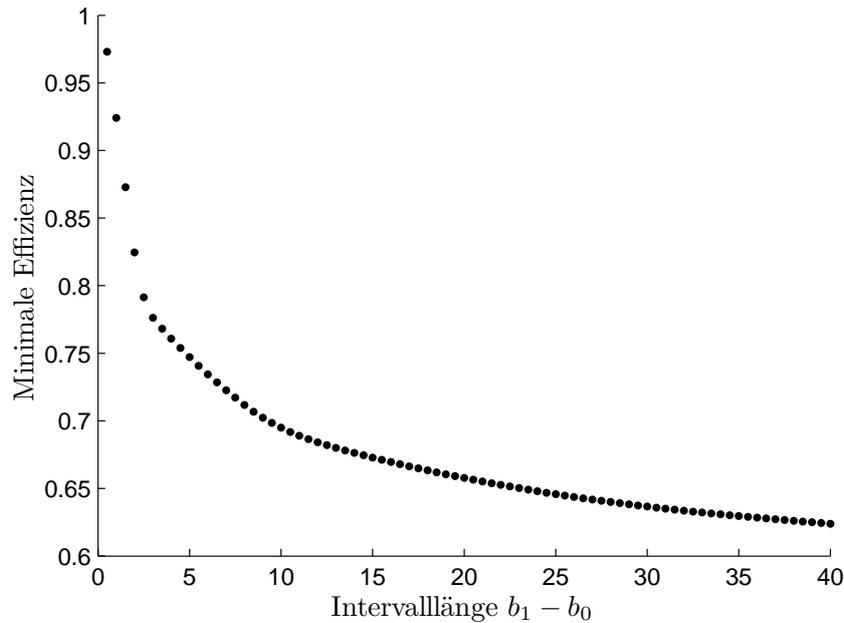


Abbildung 6.8: Minimale Effizienz der standardisiert Maximin D -optimalen Designs in Abhängigkeit von der Intervalllänge des Parameterbereichs $[b_0, b_1]$ mit $b_1 = -1$

2-Punkt-Design ξ^* eine nur geringfügig kleinere minimale Effizienz. In Abbildung 6.9 wird die Effizienz des Designs ξ^* und des global optimalen Designs für $\beta_1 \in [-4, -1]$ grafisch dargestellt. Es zeigt sich, dass das 2-Punkt-Design ξ^* für die meisten Werte von β_1 eine deutlich höhere Effizienz als das global optimale Design besitzt.

b) Der Parameterbereich wird nun als $[-12, -1]$ gewählt. In diesem Fall ist das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs nach Satz 6.3 durch

$$\xi^* = \begin{Bmatrix} 0 & 0.503 \\ 0.389 & 0.611 \end{Bmatrix}$$

gegeben. Es besitzt eine minimale Effizienz von 0.423, welche bei $\beta_1 = b_0 = -12$ und $\beta_1 = b_1 = -1$ angenommen wird. Auch in diesem Fall ist ξ^* nicht standardisiert Maximin D -optimal in der Klasse aller Designs. Das global optimale Design ist in Tabelle 6.1 angegeben. Es handelt sich um ein 4-Punkt-Design mit einer minimalen Effizienz von 0.689. Mit größer werdendem Parameterbereich wird der Unterschied zwischen den minimalen Effizienzen beim 2-Punkt-Design und beim global optimalen Design größer. Für $\beta_1 \in [-12, -1]$ wird die Effizienz des Designs ξ^* und des global optimalen Designs in Abbildung 6.9 dargestellt. Der Anteil der Werte von β_1 , bei denen die Effizienz des 2-Punkt-Designs ξ^* größer als beim global optimalen Design ist, nimmt mit größer werdendem Parameterbereich ab.

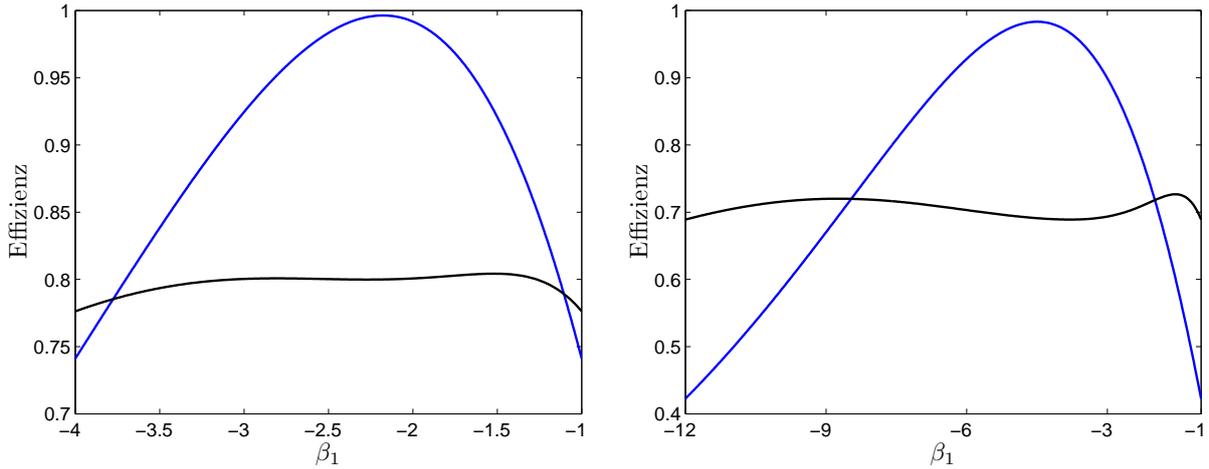


Abbildung 6.9: Effizienzen der standardisiert Maximin D -optimalen Designs in der Klasse der 2-Punkt-Designs (blau) und in der Klasse aller Designs (schwarz) für die Parameterbereiche $[-4, -1]$ (links) und $[-12, -1]$ (rechts)

Im folgenden Beispiel wird Satz 6.2 angewendet, um das standardisiert Maximin D -optimale Design bei Transformation des Parameterbereichs zu erhalten.

Beispiel 6.8

Es sei $\mathcal{X} = [0, 6]$, $\mathbf{f}(x) = (1, x)^T$ und $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ mit $\beta_0 = 0$ und $\beta_1 \in [-2, -\frac{1}{2}]$. Außerdem seien $m = 10$ und $b = 1$. Mit $c = 1/2$ gilt $\mathcal{X} = [0, 6] = \frac{1}{c} \cdot [0, 3]$ und $\beta_1 \in [-2, -\frac{1}{2}] = c \cdot [-4, -1]$. Somit lässt sich das standardisiert Maximin D -optimale Design ξ_c^* nach Satz 6.2 aus dem standardisiert Maximin D -optimalen Design ξ^* für den Designraum $[0, 3]$ und Parameterbereich $[-4, 1]$ für β_1 herleiten, welches in Tabelle 6.1 gegeben ist. Es folgt

$$\xi_c^* = \begin{Bmatrix} 0 & 1.28 & 4.53 \\ 0.28 & 0.41 & 0.31 \end{Bmatrix}$$

als das standardisiert Maximin D -optimale Design. Die minimale Effizienz von ξ_c^* stimmt mit der von ξ^* überein und beträgt 0.776.

6.1.3 Multiples Regressionsmodell

In diesem Abschnitt wird das multiple Regressionmodell mit der Regressionsfunktion $\mathbf{f}(\mathbf{x}) = (1, \mathbf{x}^T)^T$, wobei $\mathbf{x} \in \mathbb{R}^2$ ist, und dem Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ betrachtet. Der Designraum sei $\mathcal{X} = [0, 3] \times [0, 3]$ und für die Komponenten des Parametervektors $\boldsymbol{\beta}$ gelte $\beta_0 = 0$ und $(\beta_1, \beta_2)^T \in B = [\mathbf{b}_0, \mathbf{b}_1] = [b_{01}, b_{11}] \times [b_{02}, b_{12}]$. Der Parameterbereich ist somit $\tilde{B} = \{\beta_0\} \times B$. Weiterhin seien $m = 10$ und $b = 1$.

$[\mathbf{b}_0, \mathbf{b}_1]$	Optimales Design	Min. Eff.
$[-1.1, -1]^2$	$\left\{ \begin{array}{ccc} (2.13, 0) & (0, 2.13) & (0, 0) \\ 0.40 & 0.40 & 0.21 \end{array} \right\}$	0.997
$[-2, -1]^2$	$\left\{ \begin{array}{ccc} (1.53, 0) & (0, 1.53) & (0, 0) \\ 0.39 & 0.39 & 0.22 \end{array} \right\}$	0.862
$[-3.4, -1]^2$	$\left\{ \begin{array}{ccccc} (0.89, 0) & (1.98, 0) & (0, 0.89) & (0, 1.98) & (0, 0) \\ 0.28 & 0.12 & 0.28 & 0.12 & 0.21 \end{array} \right\}$	0.645
$[-4, -1]^2$	$\left\{ \begin{array}{cccccc} (0.63, 0) & (2.19, 0) & (0, 0.63) & (0, 2.19) & (0.45, 0.45) & (0, 0) \\ 0.24 & 0.16 & 0.24 & 0.16 & 0.01 & 0.18 \end{array} \right\}$	0.611
$[-3, -1] \times [-2, -1]$	$\left\{ \begin{array}{ccc} (1.21, 0) & (0, 1.52) & (0, 0) \\ 0.38 & 0.39 & 0.22 \end{array} \right\}$	0.774
$[-4, -1] \times [-2, -1]$	$\left\{ \begin{array}{cccc} (0.64, 0) & (2.19, 0) & (0, 1.52) & (0, 0) \\ 0.25 & 0.16 & 0.39 & 0.20 \end{array} \right\}$	0.725

Tabelle 6.2: Standardisiert Maximin D -optimale Designs für verschiedene Parameterbereiche $B = [\mathbf{b}_0, \mathbf{b}_1]$ für multiples Regressionsmodell in Abschnitt 6.1.3

In Tabelle 6.2 sind für einige ausgewählte Parameterbereiche $[\mathbf{b}_0, \mathbf{b}_1]$ für $(\beta_1, \beta_2)^T$ die standardisiert Maximin D -optimalen Designs und ihre zugehörigen minimalen Effizienzen angegeben. Die Optimalität der Designs kann mit dem Äquivalenzsatz 4.11 nachgewiesen werden, wobei die dafür jeweils notwendige Verteilung π auf $\mathcal{N}(\xi^*)$ in Tabelle 6.3 angegeben ist. In Abbildung 6.10 wird für einige der Designs sowohl die D -Effizienz als auch der Äquivalenzsatz grafisch dargestellt.

$[\mathbf{b}_0, \mathbf{b}_1]$	π
$[-1.1, -1]^2$	$\left\{ \begin{array}{cc} (-1.1, -1) & (-1, -1.1) \\ 0.50 & 0.50 \end{array} \right\}$
$[-2, -1]^2$	$\left\{ \begin{array}{ccc} (-2, -1) & (-1, -2) & (-1, -1) \\ 0.45 & 0.45 & 0.10 \end{array} \right\}$
$[-3.4, -1]^2$	$\left\{ \begin{array}{ccc} (-3.4, -1) & (-1, -3.4) & (-1, -1) \\ 0.43 & 0.43 & 0.14 \end{array} \right\}$
$[-4, -1]^2$	$\left\{ \begin{array}{cccc} (-4, -4) & (-4, -1) & (-1, -4) & (-1, -1) \\ 0.06 & 0.40 & 0.40 & 0.14 \end{array} \right\}$
$[-3, -1] \times [-2, -1]$	$\left\{ \begin{array}{ccc} (-3, -1) & (-1, -2) & (-1, -1) \\ 0.41 & 0.46 & 0.13 \end{array} \right\}$
$[-4, -1] \times [-2, -1]$	$\left\{ \begin{array}{ccc} (-4, -1) & (-1, -2) & (-1, -1) \\ 0.46 & 0.46 & 0.08 \end{array} \right\}$

Tabelle 6.3: Verteilung π zum Nachweis der Maximin D -Optimalität für die Designs in Tabelle 6.2 mit dem Äquivalenzsatz

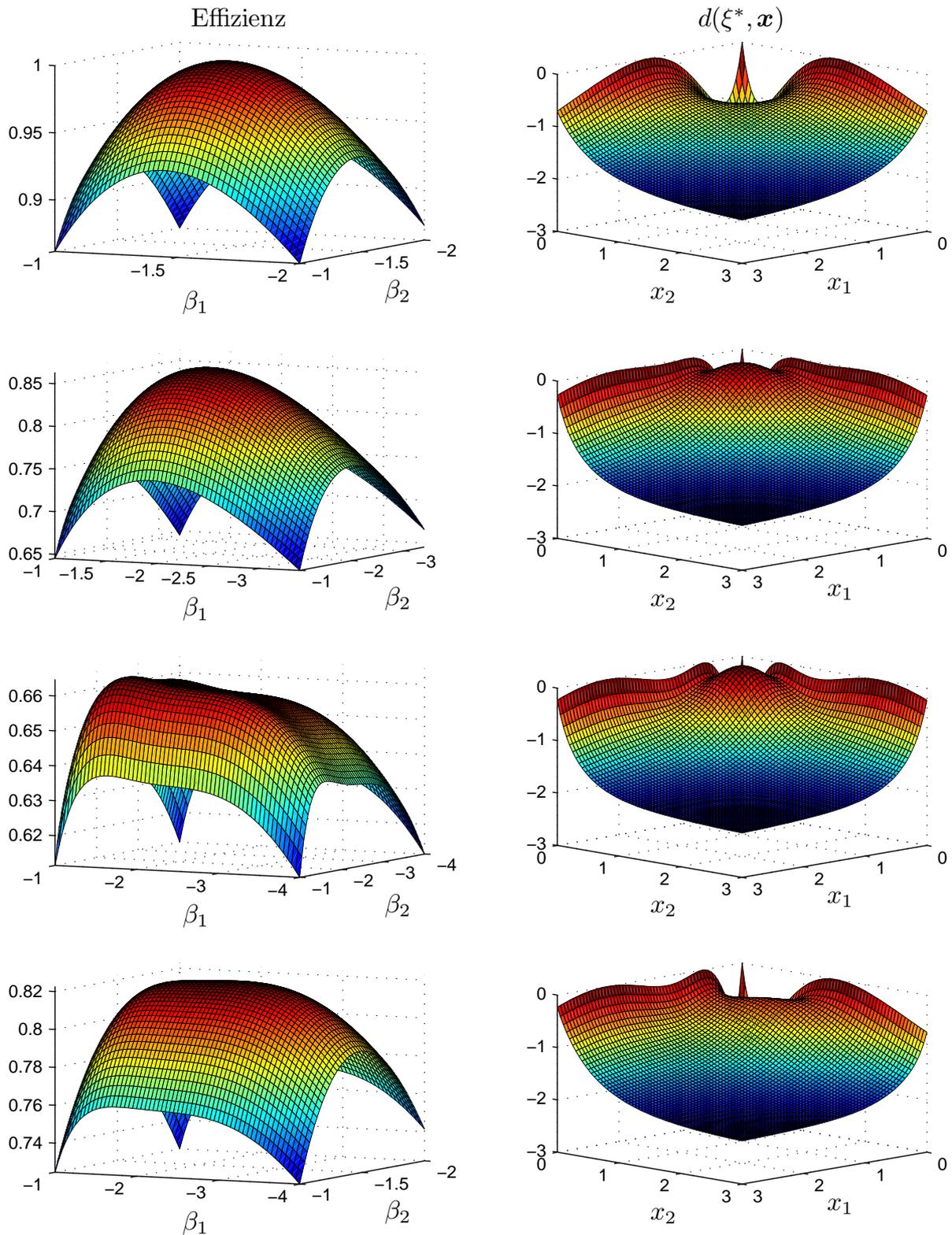


Abbildung 6.10: D -Effizienz (links) und Sensitivitätsfunktion (rechts) für die standardisiert Maximin D -optimalen Designs in Tabelle 6.2 für die Parameterbereiche $[-2, -1]^2$, $[-3.4, -1]^2$, $[-4, -1]^2$ und $[-4, -1] \times [-2, -1]$ (von oben nach unten)

Ein Stützpunkt der standardisiert Maximin D -optimalen Designs ist immer der Punkt $(0, 0)$. Sowohl die Anzahl der Stützpunkte als auch die Anzahl der Minimalstellen der Effizienz $|\mathcal{N}(\xi^*)|$ steigt mit der Größe des Parameterbereiches B . Dabei besitzen die optimalen Designs zunächst nur Stützpunkte bei der Ecke $(0, 0)$ und den zu $(0, 0)$ inzidenten Kanten. Mit größer werdendem Parameterbereich für mindestens einen der Parameter β_1 und β_2 steigt die Anzahl der Stützpunkte auf der jeweiligen Kante. Im Gegensatz zu den lokal D -optimalen Designs kann ab einer gewissen Größe des Parameterbereiches B für $(\beta_1, \beta_2)^T$ auch ein Stützpunkt im Inneren des Designraums auftreten.

6.2 Standardisiert Maximin D_A -optimale Designs

Es wird das Regressionsmodell mit $\mathbf{f} = (1, f_1, \dots, f_{p-1})^T$ betrachtet, wobei die Regressionsfunktionen f_i beliebige Funktionen seien.

In Satz 5.12 wurde gezeigt, dass unter gewissen Voraussetzungen an die Matrix \mathbf{A} die lokal D_A -optimalen Designs im Poisson-Gamma-Modell und im Poisson-Modell übereinstimmen. Diese Äquivalenz der optimalen Designs in beiden Modellen lässt sich auch auf standardisierte Maximin D_A -Optimalität übertragen.

Satz 6.9

Es sei $\mathbf{A}^T = (\mathbf{0}_s, \tilde{\mathbf{A}})$, wobei $\tilde{\mathbf{A}}$ eine $s \times (p-1)$ -Matrix mit $\text{Rang}(\tilde{\mathbf{A}}) = s$ ist. Ein Design ξ^* ist genau dann standardisiert Maximin D_A -optimal im Poisson-Gamma-Modell, wenn das Design ξ^* im Poisson-Modell standardisiert Maximin D_A -optimal ist.

Beweis:

Nach Satz 5.12 sind die Kriteriumsfunktionen für D_A -Optimalität für das Poisson- und Poisson-Gamma-Modell bis auf einen positiven konstanten Faktor gleich und die D_A -optimalen Designs stimmen in beiden Modellen überein. Somit sind die Kriteriumsfunktionen für standardisierte Maximin D_A -Optimalität für das Poisson- und Poisson-Gamma-Modell identisch. Zudem ist die Identifizierbarkeitsbedingung nach Satz 4.3 in beiden Modellen äquivalent. Daher stimmen die standardisiert Maximin D_A -optimalen Designs im Poisson- und Poisson-Gamma-Modell überein. \square

Das standardisierte Maximin D_s -Optimalitätskriterium für $\beta_1, \dots, \beta_{p-1}$ ist ein Spezialfall des standardisierten Maximin D_A -Optimalitätskriteriums, wenn $\mathbf{A}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1})$ gewählt wird. Somit sind nach Satz 6.9 auch die standardisiert Maximin D_s -optimalen Designs für $\beta_1, \dots, \beta_{p-1}$ für das Poisson- und Poisson-Gamma-Modell gleich.

6.3 Standardisiert Maximin c -optimale Designs

Die Regressionsfunktion sei $\mathbf{f}(x) = (1, x)^T$ und der Parametervektor sei $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$. Bei einem binären Designraum $\mathcal{X} = \{0, 1\}$ ergibt sich für festes β_0 und $\beta_1 \in [b_0, b_1]$ aus dem standardisiert Maximin D -optimalen Design für das Poisson-Gamma-Modell für $b \rightarrow 0$ das standardisiert Maximin c -optimale Design mit $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell, welches bereits bekannt ist (vgl. Konstantinou et al., 2014). Es lässt sich ebenfalls als Mittelwert $\xi^* = \frac{1}{2} \cdot (\xi_{\beta_0}^* + \xi_{\beta_1}^*)$ der für $\boldsymbol{\beta}_0 = (\beta_0, b_0)^T$ und $\boldsymbol{\beta}_1 = (\beta_0, b_1)^T$ lokal c -optimalen Designs $\xi_{\beta_0}^*$ und $\xi_{\beta_1}^*$ konstruieren. Da die lokal c -optimalen Designs für das Poisson-Modell nicht vom Parameter β_0 abhängen, hängt auch das standardisiert Maximin c -optimale Design für das Poisson-Modell nicht von β_0 ab.

Bei einem stetigen Designraum $\mathcal{X} = [u, v]$ lässt sich das standardisiert Maximin c -optimale Design für $\mathbf{c} = (0, 1)^T$ für das Poisson-Modell analog zum standardisiert Maximin D -optimalen Design für das Poisson-Gamma-Modell für $b \rightarrow 0$ folgern. Dazu muss der Beweis von Satz 6.3 an einigen Stellen angepasst werden. Der Beweis wird im Anhang angegeben.

Satz 6.10

Es sei $\mathcal{X} = [u, v]$ und $\beta_1 \in [b_0, b_1]$ mit $b_0 < b_1$, wobei $b_0 > 0$ und $v = 0$ oder $b_1 < 0$ und $u = 0$ gelte. Außerdem gelte $2 \cdot (1 + W(e^{-1})) \leq \min\{|b_0|, |b_1|\} \cdot (v - u)$. Es sei

$$w(x) = \frac{1}{1 - \frac{1}{b_0^2 - b_1^2} \cdot \left(\frac{b_0^2}{\exp(b_1 x)} - \frac{b_1^2}{\exp(b_0 x)} \right)} \quad (6.11)$$

und $\tilde{x} = 2 \cdot \ln(b_1/b_0)/(b_0 - b_1)$. Für $b_1 < 0$ sei $I = (\tilde{x}, \infty)$ und für $b_0 > 0$ sei $I = (-\infty, \tilde{x})$. Außerdem sei

$$x^* = \arg \max_{x \in I} \frac{w(x) \cdot (1 - w(x)) \cdot x^2}{w(x) \cdot (\exp(-b_0 x) - 1) + 1}. \quad (6.12)$$

Falls $x^* \in \mathcal{X}$ gilt, dann ist im Poisson-Modell das standardisiert Maximin c -optimale Design ξ^* für $\mathbf{c} = (0, 1)^T$ in der Klasse der 2-Punkt-Designs gegeben durch:

$$\xi^* = \left\{ \begin{array}{cc} 0 & x^* \\ w(x^*) & 1 - w(x^*) \end{array} \right\}. \quad (6.13)$$

Für die Regressionsfunktion $\mathbf{f}(x) = (1, x)^T$ und Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ ist das c -Optimalitätskriterium für $\mathbf{c} = (0, 1)^T$ gleich dem D_s -Optimalitätskriterium für β_1 . Somit folgt als Spezialfall von Satz 6.9 die folgende Aussage.

Satz 6.11

Es sei $\mathbf{c} = (0, 1)^T$. Ein Design ξ^* ist genau dann standardisiert Maximin c -optimal im Poisson-Gamma-Modell, wenn das Design ξ^* im Poisson-Modell standardisiert Maximin c -optimal ist.

Insbesondere ist das in Satz 6.10 gegebene Design auch standardisiert Maximin c -optimal für $\beta_1 \in [b_0, b_1]$ in der Klasse der 2-Punkt-Designs für das Poisson-Gamma-Modell.

Beim Poisson-Modell kann der Parameter β_0 als Faktor $\exp(\beta_0)$ aus der Informationsmatrix herausgezogen werden. Daher hängen die optimalen Designs beim Poisson-Modell nicht von β_0 ab. Da die standardisierten Maximin c -optimalen Designs beim Poisson-Gamma- und Poisson-Modell nach Satz 6.11 übereinstimmen, ist das standardisiert Maximin c -optimale Design für das Poisson-Gamma-Modell ebenfalls unabhängig von β_0 . Somit behalten die in diesem Kapitel erzielten Resultate zu den standardisiert Maximin c -optimalen Designs auch Gültigkeit, wenn für den Parameter β_0 ebenfalls ein beliebiger Parameterbereich vorgegeben wird.

7 Optimale Designs bei unbekanntem Parametern der Gamma-Verteilung

In diesem Kapitel wird der Fall unbekannter Parameter a und b der Gamma-Verteilung betrachtet. Das Verhältnis der Parameter sei $a/b = c$ mit bekanntem $c \in \mathbb{R}$. Die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta}, a) = m^{-1} \mathbf{I}(\boldsymbol{\beta}, a)$ für ein Design ξ , wobei $\mathbf{I}(\boldsymbol{\beta}, a)$ die Informationsmatrix in Satz 2.6 ist, besitzt die Form:

$$\mathbf{M}(\xi; \boldsymbol{\beta}, a) = \begin{pmatrix} \mathbf{M}(\xi; \boldsymbol{\beta}) & \mathbf{0}_p \\ \mathbf{0}_p^T & M(\xi; a) \end{pmatrix}. \quad (7.1)$$

Die Matrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ ist die Informationsmatrix für den Parametervektor $\boldsymbol{\beta}$, welche in Gleichung (4.1) gegeben ist. Die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta}, a)$ ist wie die Fisher-Informationsmatrix $\mathbf{I}(\boldsymbol{\beta}, a)$ eine Blockdiagonalmatrix.

Die Beobachtungen zwischen den einzelnen statistischen Einheiten sind unabhängig. Daher ergibt sich die Informationsmatrix für ein Populationsdesign $\zeta = \{\xi_1 \dots \xi_r\}$ zu $\mathbf{M}(\zeta; \boldsymbol{\beta}) = \sum_{i=1}^r q_i \mathbf{M}(\xi_i; \boldsymbol{\beta})$. Somit ist die Informationsmatrix für das Populationsdesign ζ von der Form

$$\mathbf{M}(\zeta; \boldsymbol{\beta}, a) = \begin{pmatrix} \mathbf{M}(\zeta; \boldsymbol{\beta}) & \mathbf{0}_p \\ \mathbf{0}_p^T & M(\zeta; a) \end{pmatrix}, \quad (7.2)$$

wobei $\mathbf{M}(\zeta; \boldsymbol{\beta})$ die Informationsmatrix für den Parametervektor $\boldsymbol{\beta}$ bei Verwendung des Populationsdesigns ζ ist. Bei $\mathbf{M}(\zeta; \boldsymbol{\beta}, a)$ handelt es sich ebenfalls um eine Blockdiagonalmatrix.

Es soll der Parametervektor $\boldsymbol{\beta}$ oder eine Linearkombinationen $\tilde{\mathbf{A}}^T \boldsymbol{\beta}$ geschätzt werden. Dabei ist $\tilde{\mathbf{A}}$ eine $p \times s$ -Matrix mit $\text{Rang}(\tilde{\mathbf{A}}) = s$. Der gesamte Parametervektor ist durch $(\boldsymbol{\beta}^T, a)^T$ gegeben. Mit $\mathbf{A}^T = (\mathbf{I}_p, \mathbf{0}_p)$ ergibt sich $\mathbf{A}^T (\boldsymbol{\beta}^T, a)^T = \boldsymbol{\beta}$. Soll nur $\beta_1, \dots, \beta_{p-1}$ geschätzt werden, wird $\mathbf{A}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1}, \mathbf{0}_{p-1})$ gewählt. Allgemein wird $\mathbf{A}^T = (\tilde{\mathbf{A}}^T, \mathbf{0}_s)$ betrachtet, um die Linearkombinationen $\tilde{\mathbf{A}}^T \boldsymbol{\beta}$ bestmöglich zu schätzen.

Die Informationsmatrix $\mathbf{M}(\zeta; \boldsymbol{\beta}, a)$ ist eine Blockdiagonalmatrix mit den Blockmatrizen $\mathbf{M}(\zeta; \boldsymbol{\beta})$ und $M(\zeta; a)$. Daher ist die Identifizierbarkeit von $\tilde{\mathbf{A}}^T \boldsymbol{\beta}$ im Poisson-Gamma-Modell mit bekannten und unbekanntem Parametern der Gamma-Verteilung äquivalent.

Satz 7.1

- a) Ein Populationsdesign ζ^* ist genau dann (standardisiert Maximin) D_s -optimal für $\boldsymbol{\beta}$ im Poisson-Gamma-Modell mit unbekanntem Parametern $a/b = c$, wenn das Design ζ^* (standardisiert Maximin) D -optimal im Poisson-Gamma-Modell mit bekannten Parametern a und b ist.
- b) Ein Populationsdesign ζ^* ist genau dann (standardisiert Maximin) D_s -optimal für $\beta_1, \dots, \beta_{p-1}$ im Poisson-Gamma-Modell mit unbekanntem Parametern $a/b = c$, wenn das Design ζ^* (standardisiert Maximin) D_s -optimal für $\beta_1, \dots, \beta_{p-1}$ im Poisson-Gamma-Modell mit bekannten Parametern a und b ist.
- c) Es seien $\tilde{\mathbf{A}}$ eine $p \times s$ -Matrix mit $\text{Rang}(\tilde{\mathbf{A}}) = s$ und $\mathbf{A}^T = (\tilde{\mathbf{A}}^T, \mathbf{0}_s)$. Ein Populationsdesign ζ^* ist genau dann (standardisiert Maximin) L -optimal für $\mathbf{A}^T(\boldsymbol{\beta}^T, a)^T = \tilde{\mathbf{A}}^T \boldsymbol{\beta}$ im Poisson-Gamma-Modell mit unbekanntem Parametern $a/b = c$, wenn das Design ζ^* (standardisiert Maximin) L -optimal für $\tilde{\mathbf{A}}^T \boldsymbol{\beta}$ im Poisson-Gamma-Modell mit bekannten Parametern a und b ist.

Beweis:

a) Es sei $\mathbf{A}^T = (\mathbf{I}_p, \mathbf{0}_p)$. Da die Informationsmatrix $\mathbf{M}(\zeta; \boldsymbol{\beta}, a)$ eine Blockdiagonalmatrix ist, muss $\mathbf{M}(\zeta; \boldsymbol{\beta})$ regulär sein, damit $\mathbf{A}^T(\boldsymbol{\beta}^T, a)^T = \boldsymbol{\beta}$ identifizierbar ist. Mit Satz 3.5 folgt für die Kriteriumsfunktion für D -Optimalität eines Populationsdesigns ζ :

$$\begin{aligned} \det(\mathbf{A}^T \mathbf{M}(\zeta; \boldsymbol{\beta}, a)^{-1} \mathbf{A}) &= \frac{\det(M(\zeta; a))}{\det(\mathbf{M}(\zeta; \boldsymbol{\beta}, a))} = \frac{M(\zeta; a)}{\det(\mathbf{M}(\zeta; \boldsymbol{\beta})) \cdot M(\zeta; a)} \\ &= \frac{1}{\det(\mathbf{M}(\zeta; \boldsymbol{\beta}))}. \end{aligned}$$

Da bei D_s -Optimalität die Kriteriumsfunktion minimiert und bei D -Optimalität die Kriteriumsfunktion maximiert wird, folgt die Gleichheit des optimalen Populationsdesigns.

b) Es sei $\mathbf{A}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1}, \mathbf{0}_{p-1})$. Da die Identifizierbarkeit von $\tilde{\mathbf{A}}^T \boldsymbol{\beta}$ im Poisson-Gamma-Modell mit bekannten und unbekanntem Parametern der Gamma-Verteilung äquivalent ist, ist $\mathbf{M}(\zeta; \boldsymbol{\beta})$ nach Satz 3.4 regulär. Es seien M_{11} das oberste Hauptdiagonalelement von $\mathbf{M}(\zeta; \boldsymbol{\beta})$ und

$$\tilde{\mathbf{M}} = \begin{pmatrix} M_{11} & 0 \\ 0 & M(\zeta; a) \end{pmatrix}$$

eine Teilmatrix von $\mathbf{M}(\zeta; \boldsymbol{\beta}, a)$. Mit Satz 3.5 und $\tilde{\mathbf{A}}^T = (\mathbf{0}_{p-1}, \mathbf{I}_{p-1})$ folgt für die Kriteriumsfunktion für D_s -Optimalität für ein Populationsdesign ζ :

$$\begin{aligned} \det(\mathbf{A}^T \mathbf{M}(\zeta; \boldsymbol{\beta}, a)^{-1} \mathbf{A}) &= \frac{\det(\tilde{\mathbf{M}})}{\det(\mathbf{M}(\zeta; \boldsymbol{\beta}, a))} = \frac{M(\zeta; a) \cdot M_{11}}{\det(\mathbf{M}(\zeta; \boldsymbol{\beta})) \cdot M(\zeta; a)} \\ &= \frac{M_{11}}{\det(\mathbf{M}(\zeta; \boldsymbol{\beta}))} = \det(\tilde{\mathbf{A}}^T \mathbf{M}(\zeta; \boldsymbol{\beta})^{-1} \tilde{\mathbf{A}}). \end{aligned}$$

Somit folgt die Gleichheit der D_s -optimalen Designs.

c) Es seien $\mathbf{B} = \mathbf{A}\mathbf{A}^T$ und $\tilde{\mathbf{B}} = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^T$. Für die Kriteriumsfunktion gilt:

$$\begin{aligned} \text{Spur}(\mathbf{M}(\zeta; \boldsymbol{\beta}, a)^{-1} \mathbf{B}) &= \text{Spur}(\mathbf{A}^T \mathbf{M}(\zeta; \boldsymbol{\beta}, a)^{-1} \mathbf{A}) \\ &= \text{Spur}\left(\begin{pmatrix} \tilde{\mathbf{A}}^T & \mathbf{0}_s \end{pmatrix} \cdot \begin{pmatrix} \mathbf{M}(\zeta; \boldsymbol{\beta})^{-1} & \mathbf{0}_p \\ \mathbf{0}_p^T & M(\zeta; a)^{-1} \end{pmatrix} \cdot \begin{pmatrix} \tilde{\mathbf{A}} \\ \mathbf{0}_s^T \end{pmatrix}\right) \\ &= \text{Spur}(\tilde{\mathbf{A}}^T \mathbf{M}(\zeta; \boldsymbol{\beta})^{-1} \tilde{\mathbf{A}}) = \text{Spur}(\mathbf{M}(\zeta; \boldsymbol{\beta})^{-1} \tilde{\mathbf{B}}). \end{aligned}$$

Aus der Gleichheit der Kriteriumsfunktionen folgt die Gleichheit der L -optimalen Designs.

Da die jeweiligen Kriteriumsfunktionen im Poisson-Gamma-Modell mit bekannten und unbekanntem Parametern der Gamma-Verteilung identisch sind, folgt direkt die Gleichheit der standardisiert Maximin optimalen Designs. \square

Als Spezialfall von L -Optimalität ergibt sich c -Optimalität, wenn \mathbf{A} ein $(p+1)$ -dimensionaler Vektor ist. Somit folgt aus Satz 7.1, dass die (standardisiert Maximin) c -optimalen Populationsdesigns ζ^* für $\mathbf{c} = (\tilde{\mathbf{c}}^T, 0)^T$ mit $\tilde{\mathbf{c}} \in \mathbb{R}^p$ im Poisson-Gamma-Modell mit bekannten und unbekanntem Parametern der Gamma-Verteilung übereinstimmen. Satz 7.1 zeigt, dass sich die Resultate zu lokal optimalen und standardisiert Maximin optimalen Designs für den Fall bekannter Parameter a und b der Gamma-Verteilung auch auf den Fall unbekannter Parameter übertragen lassen, wenn das Verhältnis von a und b bekannt ist.

8 Diskussion

Das Poisson-Gamma-Modell ist eine Verallgemeinerung des Poisson-Modells, welches sich durch Annahme eines Gamma-verteiltern zufälligen Blockeffekts für jede statistische Einheit ergibt. Es ist deutlich komplexer als das Poisson-Modell, wodurch die analytische Untersuchung des Modells erschwert wird. So existieren für das Poisson-Modell bereits viele Resultate zu optimalen Designs für verschiedene Optimalitätskriterien, während das Poisson-Gamma-Modell im Kontext der optimalen Versuchsplanung bisher nur wenig betrachtet wurde.

In dieser Arbeit wurden für das Poisson-Gamma-Modell viele analytische Resultate erzielt. Sie leistet somit einen wichtigen Beitrag zur optimalen Versuchsplanung für Verallgemeinerte Lineare Gemischte Modelle.

Für das Poisson-Gamma-Modell kann die Dichte analytisch berechnet werden, so dass die Fisher-Informationsmatrix hergeleitet werden kann. Sie ist notwendig, um für das Poisson-Gamma-Modell analytische Resultate zu optimalen Designs für verschiedene Optimalitätskriterien erzielen zu können.

Die analytische Berechnung der Fisher-Informationsmatrix ist bei Annahme anderer Verteilungen für den zufälligen Effekt, wie etwa der Normalverteilung, nicht möglich. Dann muss die Fisher-Informationsmatrix approximiert werden und die Güte der so erhaltenen Designs hängt stark von der verwendeten Approximation ab. Eine Beurteilung der optimalen Designs im Vergleich zu den optimalen Designs für das Modell ohne zufälligen Effekt ist somit nur eingeschränkt möglich.

Es wurden individuelle Designs, welche die Versuchseinstellungen und zugehörigen Gewichte für eine einzelne statistische Einheit angeben, und Populationsdesigns, welche aus den individuellen Designs und zugehörigen Gewichten für die statistischen Einheiten des gesamten Experiments bestehen, eingeführt. Da die Informationsmatrizen für das Poisson-Gamma-Modell konkav auf der Menge der individuellen Designs sind, ist das Populationsdesign optimal, welches allen statistischen Einheiten das optimale individuelle Design zuweist. Daher konnte sich auf die Herleitung optimaler individueller Designs beschränkt werden.

Zunächst wurde der Fall bekannter Parameter der Gamma-Verteilung betrachtet. Die Informationsmatrix für das Poisson-Gamma-Modell lässt sich in Abhängigkeit von der Informationsmatrix für das Poisson-Modell darstellen. Weiterhin sind die Ränge der Informationsmatrizen bei Verwendung des gleichen Designs immer identisch und auch die verallgemeinerte Inverse für das Poisson-Gamma-Modell kann unter Verwendung der verallgemeinerten Inverse für das Poisson-Modell dargestellt werden.

Basierend auf dem Zusammenhang zwischen den Informationsmatrizen für beide Modelle konnte das D -Optimalitätskriterium für das Poisson-Gamma-Modell in eine gewichtete Summe aus D -Optimalitätskriterium und D_s -Optimalitätskriterium für die Effektparameter $\beta_1, \dots, \beta_{p-1}$ für das Poisson-Modell zerlegt werden. Daher ergeben sich die D - und D_s -optimalen Designs für das Poisson-Modell als Spezialfall aus dem D -optimalen Design für das Poisson-Gamma-Modell.

Für das multiple Regressionsmodell mit einer beliebigen Anzahl von Kovariablen besitzen die optimalen Designs für alle diese Optimalitätskriterien die gleiche Struktur, insbesondere handelt es sich um Designs mit minimalem Träger. Ein Stützpunkt liegt auf der Ecke des Designraums, bei dem die Intensitätsfunktion maximal wird. Die anderen Stützpunkte befinden sich auf den zu dieser Ecke inzidenten Kanten, wobei sich der Abstand dieser Stützpunkte zu der Ecke zwischen den verschiedenen Optimalitätskriterien unterscheidet. Zusätzlich zu den Unterschieden in Bezug auf die Stützpunkte weichen die D -optimalen Gewichte beim Poisson-Gamma-Modell von der Gleichverteilungsregel für die D -optimalen Gewichte beim Poisson-Modell ab. Weiterhin wurde gezeigt, dass die D_s -optimalen Designs für das Poisson- und Poisson-Gamma-Modell übereinstimmen. Dies führt dazu, dass beim Poisson-Gamma-Modell sich das D_s -optimale Design als Spezialfall aus dem D -optimalen Design ergibt.

Da die D - und D_s -optimalen Designs einen minimalen Träger besitzen, lassen sie sich einfacher zu einem exakten Design runden als Designs mit einer größeren Anzahl an Stützpunkten. Dies ist insbesondere wichtig, falls die Anzahl der Beobachtungen pro statistischer Einheit klein ist.

Würde der zufällige Effekt ignoriert werden und das D -optimale Design des Poisson-Modells für das Poisson-Gamma-Modell verwendet werden, so tritt ein Effizienzverlust auf. Der Effizienzverlust hängt vom Parameter b der Gamma-Verteilung und von der Anzahl m der Beobachtungen pro statistischer Einheit ab. In den berechneten Beispielen trat ein moderater Effizienzverlust auf.

Auch bei linearen Optimalitätskriterien, wie etwa bei L - und c -Optimalität, stimmen die optimalen Designs für das Poisson- und Poisson-Gamma-Modell überein. Somit können bekannte Resultate zu optimalen Designs für das Poisson-Modell auf das Poisson-Gamma-Modell übertragen werden.

Da es sich beim Poisson-Gamma-Modell um ein nichtlineares Modell handelt, hängen die lokal optimalen Designs für alle betrachteten Optimalitätskriterien von den unbekanntem Parametern ab. Um robustere Designs bezüglich Parametermisspezifikation zu erhalten, wurden standardisierte Maximin Optimalitätskriterien betrachtet, welche die minimale Effizienz in einem vorgegebenen Parameterbereich maximieren. Dabei wurde ein Parameterbereich für die Effektparameter vorgegeben, während der Parameter für den Achsenabschnitt als beliebig, aber fest angesehen wurde, da die Missspezifikation der Effektparameter zu einem größeren Effizienzverlust als die Missspezifikation des Parameters für den Achsenabschnitt führt.

Für das Regressionsmodell mit einer Kovariablen und einen binären Designraum konnte gezeigt werden, dass sich das standardisiert Maximin D -optimale Design als Mittelwert der für die Randpunkte des Parameterbereiches lokal D -optimalen Designs ergibt. Da das standardisiert Maximin c -optimale Design für den Steigungsparameter β_1 auch diese Struktur besitzt, wäre eine interessante Erweiterung zu untersuchen, ob sich diese Resultate auf andere Maximin Optimalitätskriterien übertragen lassen.

Bei einem stetigen Designraum müssen sowohl die Stützpunkte als auch die Gewichte optimiert werden, was die Komplexität des Optimierungsproblems erhöht. Unter gewissen Voraussetzungen an den Designraum und den Parameterbereich konnte für eine Kovariable das standardisiert Maximin D -optimale Design in der Klasse der 2-Punkt-Designs bestimmt werden. Diese optimalen 2-Punkt-Designs sind auch in der Klasse aller Designs optimal, solange der Parameterbereich klein genug gewählt wird.

Um auch standardisiert Maximin D -optimale Designs mit mehr als zwei Stützpunkten zu berechnen, wurde ein Algorithmus zur Berechnung der optimalen Designs entworfen. In numerischen Untersuchungen wurde gezeigt, dass mit zunehmender Größe des Parameterbereiches die Anzahl der Stützpunkte des standardisiert Maximin D -optimalen Designs steigt. Ebenso wächst die Anzahl der Parameter, für die die Effizienz minimal wird, mit der Größe des Parameterbereiches. Da die Anzahl der Stützpunkte sehr groß werden kann, werden viele Beobachtungen pro statistischer Einheit benötigt, um diese Designs sinnvoll zu einem exakten Design runden zu können.

Die Effizienz der standardisiert Maximin D -optimalen 2-Punkt-Designs ist unimodal bezüglich des Steigungsparameters. Die maximal erreichte Effizienz ist sehr hoch, zu den Rändern des Parameterbereiches fällt die Effizienz jedoch ab. Sind Designs mit mehr als zwei Stützpunkten standardisiert Maximin D -optimal, besitzen sie eine höhere minimale Effizienz als die optimalen 2-Punkt-Designs, aber ihre maximale Effizienz ist kleiner. Ihre Effizienz variiert deutlich weniger über den Parameterbereich als bei den optimalen 2-Punkt-Designs.

Durch die standardisiert Maximin D -optimalen Designs lässt sich bei kleinem Parame-

terbereich eine gute Effizienz für alle möglichen Parameterwerte erreichen. Aber auch bei großer Unsicherheit über den Parameter und somit großem Parameterbereich lässt sich noch eine akzeptable Effizienz sicherstellen.

Für das multiple Regressionsmodell mit zwei Kovariablen wurden ebenfalls numerisch standardisiert Maximin D -optimale Designs berechnet. Die Anzahl der Stützpunkte auf den Kanten steigt mit der Größe des Parameterbereiches. Wird der Parameterbereich groß genug gewählt, kann das standardisiert Maximin D -optimale Design auch einen Stützpunkt im Inneren des Designraums besitzen. Dies ist ein weiterer Unterschied zu den lokal D -optimalen Designs, bei denen die Stützpunkte immer auf den Ecken und Kanten des Designraums liegen.

Es wurde gezeigt, dass das standardisiert Maximin c -optimale Design für den Steigungsparameter β_1 beim Poisson- und Poisson-Gamma-Modell übereinstimmt. Wie bei den lokal optimalen Designs konnte das standardisiert Maximin c -optimale Design als Spezialfall aus dem standardisiert Maximin D -optimalen Design gefolgert werden. Dabei hängen diese optimalen Designs nicht vom Parameter für den Achsenabschnitt ab.

Es wurde auch der Fall unbekannter Parameter a und b der Gamma-Verteilung betrachtet. Ist ihr Verhältnis $a/b = c$ bekannt, was häufig bei der statistischen Modellierung von Experimenten angenommen wird, dann handelt es sich bei der Fisher-Informationsmatrix um eine Blockdiagonalmatrix. Liegt Interesse nur in der Schätzung des Parametervektors β , können die erzielten Resultate zu den optimalen Designs auch auf diesen Fall angewendet werden. Ist das Verhältnis der Parameter der Gamma-Verteilung nicht bekannt, dann besitzt die Fisher-Informationsmatrix keine Blockdiagonalstruktur und die Resultate können nicht übertragen werden.

Eine mögliche Erweiterung dieser Arbeit ist für weitere Optimalitätskriterien zu untersuchen, ob es ebenfalls Beziehungen zwischen den optimalen Designs für das Poisson- und Poisson-Gamma-Modell gibt. Zudem könnten Bayessche Optimalitätskriterien betrachtet werden, um die Parameterabhängigkeit der lokal optimalen Designs zu reduzieren. Um auch Interaktionen zwischen den verschiedenen Einflussgrößen berücksichtigen zu können, wären optimale Designs für Modelle mit Interaktionsterm von Interesse. Jedoch sind diese auch bei dem im Vergleich zum Poisson-Gamma-Modell weniger komplexen Poisson-Modell bisher nicht analytisch bestimmt worden.

Ein weiterer Forschungsbereich ist die Betrachtung von Verallgemeinerten Linearen Gemischten Modellen basierend auf anderen konjugierten Verteilungen. Falls die Informationsmatrix analytisch berechenbar ist, sind ebenfalls Zusammenhänge zwischen den Informationsmatrizen und optimalen Designs beim Modell mit und ohne zufälligen Effekt zu erwarten.

A Anhang

A.1 Resultate zu Matrizen

Satz A.1 (vgl. Harville (1997, S. 98–99), Rohde (1965))

Für eine symmetrische 2×2 -Blockmatrix

$$M = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} \quad (\text{A.1})$$

mit quadratischen Teilmatrizen \mathbf{A} und \mathbf{C} gelten die folgenden Aussagen:

a) Es sei \mathbf{A} regulär. Die Matrix \mathbf{M} ist genau dann invertierbar, wenn das Schurkomplement $\mathbf{S} = \mathbf{C} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B}$ regulär ist. Für die Inverse von \mathbf{M} gilt dann:

$$M^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{B} \mathbf{S}^{-1} \mathbf{B}^T \mathbf{A}^{-1} & -\mathbf{A}^{-1} \mathbf{B} \mathbf{S}^{-1} \\ -\mathbf{S}^{-1} \mathbf{B}^T \mathbf{A}^{-1} & \mathbf{S}^{-1} \end{pmatrix}. \quad (\text{A.2})$$

b) Ist \mathbf{M} positiv semidefinit, dann ist

$$M^{-} = \begin{pmatrix} \mathbf{A}^{-} + \mathbf{A}^{-} \mathbf{B} \mathbf{S}^{-} \mathbf{B}^T \mathbf{A}^{-} & -\mathbf{A}^{-} \mathbf{B} \mathbf{S}^{-} \\ -\mathbf{S}^{-} \mathbf{B}^T \mathbf{A}^{-} & \mathbf{S}^{-} \end{pmatrix} \quad (\text{A.3})$$

eine verallgemeinerte Inverse von \mathbf{M} , wobei \mathbf{A}^{-} und \mathbf{S}^{-} verallgemeinerte Inversen von \mathbf{A} und dem verallgemeinerten Schurkomplement $\mathbf{S} = \mathbf{C} - \mathbf{B}^T \mathbf{A}^{-} \mathbf{B}$ sind.

Satz A.2 (vgl. Harville, 1997, S. 191)

Die symmetrische 2×2 -Blockmatrix besitze die Darstellung

$$M = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} \quad (\text{A.4})$$

mit quadratischen Teilmatrizen \mathbf{A} und \mathbf{C} . Ist \mathbf{A} regulär, dann gilt für die Determinante von \mathbf{M} :

$$\det(\mathbf{M}) = \det(\mathbf{A}) \cdot \det(\mathbf{C} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B}). \quad (\text{A.5})$$

Satz A.3 (Woodbury-Matrix-Identität (vgl. Harville, 1997, S. 428))

Es sei \mathbf{A} eine reguläre $n \times n$ -Matrix, \mathbf{U} eine $n \times k$ -Matrix, \mathbf{C} eine reguläre $k \times k$ -Matrix und \mathbf{V} eine $k \times n$ -Matrix. Die Matrix $\mathbf{A} + \mathbf{UCV}$ ist genau dann regulär, wenn $\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}$ regulär ist. Dann gilt für die Inverse von $\mathbf{A} + \mathbf{UCV}$:

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1}. \quad (\text{A.6})$$

A.2 Beweise zu Kapitel 2

Beweis von Satz 2.2:

Die Dichte von Y_j ist in Gleichung (2.7) gegeben. Für den Erwartungswert von Y_j gilt:

$$\begin{aligned} E_{\beta,a,b}(Y_j) &= \sum_{y=0}^{\infty} y \cdot \frac{b^a \cdot \Gamma(a+y)}{\Gamma(a) \cdot y!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+y}} \\ &= \sum_{y=1}^{\infty} \frac{b^a \cdot \Gamma(a+y)}{\Gamma(a) \cdot (y-1)!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+y}} \\ &= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \sum_{y=1}^{\infty} \frac{b^{a+1} \cdot \Gamma(a+1+y-1)}{\Gamma(a+1) \cdot (y-1)!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot (y-1)}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+1+y-1}} \\ &= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \sum_{y=0}^{\infty} \frac{b^{a+1} \cdot \Gamma(a+1+y)}{\Gamma(a+1) \cdot y!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+1+y}} \\ &= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}. \end{aligned}$$

Unter Verwendung der obigen Berechnung des Erwartungswertes von Y_j folgt für den Erwartungswert von Y_j^2 :

$$\begin{aligned} E_{\beta,a,b}(Y_j^2) &= \sum_{y=0}^{\infty} y^2 \cdot \frac{b^a \cdot \Gamma(a+y)}{\Gamma(a) \cdot y!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+y}} \\ &= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \sum_{y=1}^{\infty} y \cdot \frac{b^{a+1} \cdot \Gamma(a+1+y-1)}{\Gamma(a+1) \cdot (y-1)!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot (y-1)}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+1+y-1}} \\ &= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \sum_{y=0}^{\infty} (y+1) \cdot \frac{b^{a+1} \cdot \Gamma(a+1+y)}{\Gamma(a+1) \cdot y!} \cdot \frac{e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y}}{(b + e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{a+1+y}} \\ &= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \left(\frac{a+1}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} + 1 \right). \end{aligned}$$

Es folgt

$$\begin{aligned}
\text{Var}(Y_j)_{\beta,a,b} &= E_{\beta,a,b}(Y_j^2) - E_{\beta,a,b}(Y_j)^2 \\
&= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \left(\frac{a+1}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} + 1 \right) - \left(\frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^2 \\
&= \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \left(\frac{1}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} + 1 \right)
\end{aligned}$$

für die Varianz von Y_j . □

Beweis von Satz 2.3:

Der Logarithmus der Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$, welche in Satz 2.1 berechnet wurde, ist durch

$$\begin{aligned}
\ln(f_{\mathbf{Y}}(\mathbf{y})) &= \ln \left(\Gamma \left(a + \sum_{j=1}^m y_j \right) \right) - \ln(\Gamma(a)) - \sum_{j=1}^m \ln(y_j!) + \sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j + a \cdot \ln(b) \\
&\quad - \left(a + \sum_{j=1}^m y_j \right) \cdot \ln \left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right) \tag{A.7}
\end{aligned}$$

gegeben. Die erste und zweite Ableitung der logarithmierten Dichte nach $\boldsymbol{\beta}$ berechnen sich zu:

$$\begin{aligned}
\frac{\partial \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial \boldsymbol{\beta}} &= \sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \cdot y_j - \frac{a + \sum_{j=1}^m y_j}{b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)^T, \\
\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} &= \frac{a + \sum_{j=1}^m y_j}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^2} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)^T \\
&\quad - \frac{a + \sum_{j=1}^m y_j}{b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T.
\end{aligned}$$

Nach Satz 2.2 gilt $E_{\beta}(Y_j) = (a/b) \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta})$. Unter Regularitätsbedingungen, insbesondere der Vertauschbarkeit der ersten und zweiten Ableitungen und des Integrals der Dichte, ist die Fisher-Information gleich dem negativen Erwartungswert der zweiten Ableitungen der Dichte. Somit folgt für die Fisher-Informationsmatrix:

$$\begin{aligned}
\mathbf{I}(\boldsymbol{\beta}) &= -E_{\boldsymbol{\beta}} \left(\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{Y}))}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \right) \\
&= \frac{a + \sum_{j=1}^m \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^2} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)^T \\
&\quad - \frac{a + \sum_{j=1}^m \frac{a}{b} \cdot e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T \\
&= \frac{a}{b} \cdot \left(\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T - \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)^T}{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} + b} \right).
\end{aligned}$$

Bei gleicher Wahl der Versuchseinstellungen lassen sich die Komponenten der Fisher-Informationsmatrix $\mathbf{I}(\boldsymbol{\beta})$ für das Poisson-Gamma-Modell als Komponenten der Fisher-Informationsmatrix $\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) = \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j) \mathbf{f}(\mathbf{x}_j)^T$ für das Poisson-Modell identifizieren. Es folgt die Darstellung

$$\mathbf{I}(\boldsymbol{\beta}) = \frac{a}{b} \cdot \left(\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) - \frac{\mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta})}{\mathbf{e}_{1,p}^T \mathbf{I}_{\text{Po}}(\boldsymbol{\beta}) \mathbf{e}_{1,p} + b} \right)$$

für die Fisher-Informationsmatrix. □

Beweis von Satz 2.5:

Die Matrix $\mathbf{I}_{\beta, \beta}$ folgt direkt aus Satz 2.3 mit $b = 1$. Unter Verwendung der Digamma-Funktion $\psi(z) = \frac{d}{dz} \ln(\Gamma(z))$ ist die erste Ableitung der logarithmierten Dichte aus Gleichung (A.7) nach a gegeben durch:

$$\frac{\partial \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial a} = \psi \left(a + \sum_{j=1}^m y_j \right) - \psi(a) - \ln \left(1 + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right).$$

Mit der Trigamma-Funktion $\psi_1(z) = \frac{d^2}{dz^2} \ln(\Gamma(z))$ gilt für die zweiten Ableitungen:

$$\begin{aligned}
\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial a \partial \boldsymbol{\beta}} &= \frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial \boldsymbol{\beta} \partial a} = - \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{1 + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}, \\
\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial a^2} &= \psi_1 \left(a + \sum_{j=1}^m y_j \right) - \psi_1(a).
\end{aligned}$$

Mit $E_{\beta,a}(Y_j) = a \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta})$ nach Satz 2.2 lassen sich die Einträge der Fisher-Informationsmatrix berechnen. Für $\mathbf{I}_{\beta,a}$ gilt:

$$\mathbf{I}_{\beta,a} = -E_{\beta,a} \left(\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{Y}))}{\partial a \partial \boldsymbol{\beta}} \right) = \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{1 + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}.$$

Für $I_{a,a}$ folgt:

$$I_{a,a} = -E_{\beta,a} \left(\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{Y}))}{\partial a^2} \right) = -E_{\beta,a} \left(\psi_1 \left(a + \sum_{j=1}^m Y_j \right) \right) + \psi_1(a).$$

Zusammen ergeben die Einträge die Fisher-Informationsmatrix. \square

Beweis von Satz 2.6:

Die Matrix $\mathbf{I}_{\beta,\beta}$ ergibt sich direkt aus Satz 2.3 mit $b = a/c$. Der Logarithmus der Dichte von $\mathbf{Y} = (Y_1, \dots, Y_m)$ in Gleichung (2.19) berechnet sich zu:

$$\begin{aligned} \ln(f_{\mathbf{Y}}(\mathbf{y})) &= \ln \left(\Gamma \left(a + \sum_{j=1}^m y_j \right) \right) - \ln(\Gamma(a)) - \sum_{j=1}^m \ln(y_j!) + \sum_{j=1}^m \mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta} \cdot y_j \\ &\quad + a \cdot (\ln(a) - \ln(c)) - \left(a + \sum_{j=1}^m y_j \right) \cdot \ln \left(\frac{a}{c} + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right). \end{aligned}$$

Mit der Digamma-Funktion $\psi(z) = \frac{d}{dz} \ln(\Gamma(z))$ gilt für die erste Ableitung der logarithmierten Dichte nach a :

$$\begin{aligned} \frac{\partial \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial a} &= \psi \left(a + \sum_{j=1}^m y_j \right) - \psi(a) + \ln(a) + 1 - \ln(c) - \ln \left(\frac{a}{c} + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right) \\ &\quad - \frac{a + \sum_{j=1}^m y_j}{a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}. \end{aligned}$$

Mit der Trigamma-Funktion $\psi_1(z) = \frac{d^2}{dz^2} \ln(\Gamma(z))$ berechnen sich die zweiten Ableitungen der logarithmierten Dichte zu:

$$\begin{aligned} \frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial a \partial \boldsymbol{\beta}} &= \frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial \boldsymbol{\beta} \partial a} \\ &= - \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{\frac{a}{c} + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} + \frac{\left(a + \sum_{j=1}^m y_j \right) \cdot c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{\left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \right)^2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{y}))}{\partial a^2} &= \psi_1\left(a + \sum_{j=1}^m y_j\right) - \psi_1(a) + \frac{1}{a} - \frac{2}{a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \\ &\quad + \frac{a + \sum_{j=1}^m y_j}{\left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^2}. \end{aligned}$$

Nach Satz 2.2 gilt $E_{\beta,a}(Y_j) = c \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta})$. Somit folgt für die Teilmatrix $\mathbf{I}_{\beta,a}$:

$$\begin{aligned} \mathbf{I}_{\beta,a} &= -E_{\beta,a}\left(\frac{\partial^2 \ln(f_{\mathbf{Y}}(\mathbf{Y}))}{\partial a \partial \boldsymbol{\beta}}\right) \\ &= \frac{\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{\frac{a}{c} + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} - \frac{\left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right) \cdot c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}} \cdot \mathbf{f}(\mathbf{x}_j)}{\left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^2} = \mathbf{0}_p. \end{aligned}$$

Für $I_{a,a}$ gilt:

$$\begin{aligned} I_{a,a} &= -E_{\beta,a}\left(\psi_1\left(a + \sum_{j=1}^m Y_j\right)\right) + \psi_1(a) - \frac{1}{a} + \frac{2}{a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}} \\ &\quad - \frac{a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{\left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^2} \\ &= -E_{\beta,a}\left(\psi_1\left(a + \sum_{j=1}^m Y_j\right)\right) + \psi_1(a) - \frac{c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}}{a \cdot \left(a + c \cdot \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)}. \end{aligned}$$

Die Teilmatrizen ergeben zusammen die Fisher-Informationsmatrix. □

Beweis von Satz 2.7:

Die Dichte f_Z von $Z = \sum_{j=1}^m Y_j$ ergibt sich aus Summation über die Dichte f_Y von $\mathbf{Y} = (Y_1, \dots, Y_m)$ aus Satz 2.1:

$$\begin{aligned} f_Z(z) &= \sum_{\substack{y_1, \dots, y_m \\ y_1 + \dots + y_m = z}} f_{\mathbf{Y}}(y_1, \dots, y_m) \\ &= \frac{b^a \cdot \Gamma(a + z)}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^{a+z} \cdot \Gamma(a)} \cdot \sum_{\substack{y_1, \dots, y_m \\ y_1 + \dots + y_m = z}} \frac{\prod_{j=1}^m (e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{y_j}}{\prod_{j=1}^m y_j!} \\ &= \frac{b^a \cdot \Gamma(a + z)}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^{a+z} \cdot \Gamma(a) \cdot z!} \cdot \sum_{\substack{y_1, \dots, y_m \\ y_1 + \dots + y_m = z}} \frac{z!}{\prod_{j=1}^m y_j!} \cdot \prod_{j=1}^m (e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}})^{y_j}. \end{aligned}$$

Unter Verwendung des Multinomialtheorems

$$(c_1 + \dots + c_m)^z = \sum_{\substack{y_1, \dots, y_m \\ y_1 + \dots + y_m = z}} \frac{z!}{\prod_{j=1}^m y_j!} \cdot \prod_{j=1}^m c_j^{y_j}$$

und mit $c_j = e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}$ folgt

$$f_Z(z) = \frac{b^a \cdot \Gamma(a+z)}{\left(b + \sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^{a+z} \cdot \Gamma(a) \cdot z!} \cdot \left(\sum_{j=1}^m e^{\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta}}\right)^z$$

und damit die Dichte der Summe von Y_1, \dots, Y_m . □

A.3 Beweise und Hilfsresultat zu Kapitel 4

Beweis von Lemma 4.1:

Es seien $\mathbf{M} = \mathbf{M}(\xi; \boldsymbol{\beta})$, $\mathbf{M}_{\text{Po}} = \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})$ und $\mathbf{L} = \mathbf{L}(\xi; \boldsymbol{\beta})$.

(i) Die Aussage folgt direkt aus Gleichung (4.1) für die Informationsmatrix, wobei für die zweite Gleichung die Symmetrie von \mathbf{M}_{Po} verwendet wird.

(ii) Die Matrix \mathbf{L} ist eine untere Dreiecksmatrix, bei der alle Hauptdiagonalelemente bis auf den Obersten gleich eins sind. Für die Determinante von \mathbf{L} gilt

$$\det(\mathbf{L}) = 1 - \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} = \left(\frac{m}{b} \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + 1\right)^{-1} \neq 0$$

und somit ist die Matrix \mathbf{L} regulär.

(iii) Die Aussage folgt mit Gleichung (4.1):

$$\begin{aligned} \mathbf{I}_p - \frac{m}{a} \cdot \mathbf{M} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T &= \mathbf{I}_p - \frac{m}{b} \cdot \left(\mathbf{M}_{\text{Po}} - \frac{\mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right) \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \\ &= \mathbf{I}_p - \frac{m}{b} \cdot \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \cdot \left(1 - \frac{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right) \\ &= \mathbf{I}_p - \frac{\mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}}. \end{aligned}$$

Hierbei wurde verwendet, dass $\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p}$ eine reelle Zahl ist. □

Lemma A.4

Die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ besitzt die folgende Darstellung:

$$\mathbf{M}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{X}^T \left(\mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} + \frac{m}{b} \cdot \mathbf{X} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}^T \right)^{-1} \mathbf{X}. \quad (\text{A.8})$$

Dabei ist \mathbf{X} die Designmatrix, \mathbf{W} die Gewichtsmatrix und $\boldsymbol{\Lambda}$ die Intensitätsmatrix.

Beweis:

Mit Gleichung (4.1) und $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \mathbf{X}$ gilt für die Informationsmatrix:

$$\begin{aligned} \mathbf{M}(\xi; \boldsymbol{\beta}) &= \frac{a}{b} \cdot \left(\mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \mathbf{X} - \frac{\mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \mathbf{X} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \mathbf{X}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} + \frac{b}{m}} \right) \\ &= \frac{a}{b} \cdot \mathbf{X}^T \left(\mathbf{W} \boldsymbol{\Lambda} - \frac{\mathbf{W} \boldsymbol{\Lambda} \mathbf{X} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} + \frac{b}{m}} \right) \mathbf{X}. \end{aligned}$$

Mit den Matrizen $\mathbf{A} = \mathbf{W} \boldsymbol{\Lambda}$, $\mathbf{U} = \mathbf{W} \boldsymbol{\Lambda} \mathbf{X} \mathbf{e}_{1,p}$, $\mathbf{C} = -(\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} + \frac{b}{m})^{-1}$ und $\mathbf{V} = \mathbf{e}_{1,p}^T \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda}$ gilt:

$$\mathbf{C}^{-1} + \mathbf{V} \mathbf{A}^{-1} \mathbf{U} = -\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} - \frac{b}{m} + \mathbf{e}_{1,p}^T \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \boldsymbol{\Lambda}^{-1} \mathbf{W}^{-1} \mathbf{W} \boldsymbol{\Lambda} \mathbf{X} \mathbf{e}_{1,p} = -\frac{b}{m}.$$

Mit der Woodbury-Matrix-Identität in Satz A.3 folgt:

$$\begin{aligned} &\left(\mathbf{W} \boldsymbol{\Lambda} - \frac{\mathbf{W} \boldsymbol{\Lambda} \mathbf{X} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda}}{\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}} \mathbf{e}_{1,p} + \frac{b}{m}} \right)^{-1} \\ &= (\mathbf{A} + \mathbf{U} \mathbf{C} \mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{C}^{-1} + \mathbf{V} \mathbf{A}^{-1} \mathbf{U})^{-1} \mathbf{V} \mathbf{A}^{-1} \\ &= \boldsymbol{\Lambda}^{-1} \mathbf{W}^{-1} - \boldsymbol{\Lambda}^{-1} \mathbf{W}^{-1} \mathbf{W} \boldsymbol{\Lambda} \mathbf{X} \mathbf{e}_{1,p} \cdot \left(-\frac{m}{b} \right) \cdot \mathbf{e}_{1,p}^T \mathbf{X}^T \mathbf{W} \boldsymbol{\Lambda} \boldsymbol{\Lambda}^{-1} \mathbf{W}^{-1} \\ &= \mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} + \frac{m}{b} \cdot \mathbf{X} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}^T. \end{aligned}$$

Es folgt

$$\mathbf{M}(\xi; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{X}^T \left(\mathbf{W}^{-1} \boldsymbol{\Lambda}^{-1} + \frac{m}{b} \cdot \mathbf{X} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}^T \right)^{-1} \mathbf{X}$$

als Darstellung für die Informationsmatrix. □

Beweis von Satz 4.6:

Der Beweis ist analog zum Beweis von Niaparast und Schwabe (2013) und basiert auf der Darstellung der Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ in Lemma A.4. Für $\alpha \in \{0, 1\}$ ist Ungleichung (4.6) mit Gleichheit erfüllt. Daher sei $\alpha \in (0, 1)$. Für das Design ξ_1 seien \mathbf{X}_1 , \mathbf{W}_1 sowie $\boldsymbol{\Lambda}_1$ und für das Design ξ_2 seien \mathbf{X}_2 , \mathbf{W}_2 sowie $\boldsymbol{\Lambda}_2$ die zugehörigen Designmatrizen, Gewichtsmatrizen und Intensitätsmatrizen. Dabei sei \mathbf{X}_1 eine $l_1 \times p$ -Matrix und \mathbf{X}_2 eine $l_2 \times p$ -Matrix. Es seien $\mathbf{X}_{\text{Ges}} = (\mathbf{X}_1^T, \mathbf{X}_2^T)^T$ und

$$\mathbf{C} = \frac{m}{b} \cdot \mathbf{X}_{\text{Ges}} \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}_{\text{Ges}}^T = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_{1,2} \\ \mathbf{C}_{1,2}^T & \mathbf{C}_2 \end{pmatrix},$$

wobei $\mathbf{C}_1 = \frac{m}{b} \cdot \mathbf{X}_1 \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}_1^T$ und $\mathbf{C}_2 = \frac{m}{b} \cdot \mathbf{X}_2 \mathbf{e}_{1,p} \mathbf{e}_{1,p}^T \mathbf{X}_2^T$ ist. Außerdem wird mit $\mathbf{0}_{r,s}$ die $r \times s$ -Nullmatrix und mit \mathbf{I}_l die $l \times l$ -Einheitsmatrix bezeichnet. Mit der Diagonalmatrix

$$\mathbf{U} = \begin{pmatrix} \alpha \cdot \mathbf{W}_1 \boldsymbol{\Lambda}_1 & \mathbf{0}_{l_1, l_2} \\ \mathbf{0}_{l_2, l_1} & (1 - \alpha) \cdot \mathbf{W}_2 \boldsymbol{\Lambda}_2 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\alpha} \cdot \mathbf{W}_1^{-1} \boldsymbol{\Lambda}_1^{-1} & \mathbf{0}_{l_1, l_2} \\ \mathbf{0}_{l_2, l_1} & \frac{1}{1 - \alpha} \cdot \mathbf{W}_2^{-1} \boldsymbol{\Lambda}_2^{-1} \end{pmatrix}$$

ist nach Lemma A.4 die Informationsmatrix für die Konvexkombination der Designs $\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2$ durch

$$\mathbf{M}(\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2; \boldsymbol{\beta}) = \frac{a}{b} \cdot \mathbf{X}_{\text{Ges}}^T (\mathbf{U} + \mathbf{C})^{-1} \mathbf{X}_{\text{Ges}}$$

gegeben. Es sei $\mathbf{G} = \text{Diag}(\sqrt{(1 - \alpha)/\alpha} \cdot \mathbf{I}_{l_1}, -\sqrt{\alpha/(1 - \alpha)} \cdot \mathbf{I}_{l_2})$. Dann ist die Matrix

$$\mathbf{H} = \begin{pmatrix} \frac{1 - \alpha}{\alpha} \cdot \mathbf{C}_1 & -\mathbf{C}_{1,2} \\ -\mathbf{C}_{1,2}^T & \frac{\alpha}{1 - \alpha} \cdot \mathbf{C}_2 \end{pmatrix} = \mathbf{G} \mathbf{C} \mathbf{G}$$

positiv semidefinit, da \mathbf{C} positiv semidefinit ist. Mit der Diagonalmatrix

$$\mathbf{V} = \begin{pmatrix} \frac{1}{\alpha} \cdot (\mathbf{W}_1^{-1} \boldsymbol{\Lambda}_1^{-1} + \mathbf{C}_1) & \mathbf{0}_{l_1, l_2} \\ \mathbf{0}_{l_2, l_1} & \frac{1}{1 - \alpha} \cdot (\mathbf{W}_2^{-1} \boldsymbol{\Lambda}_2^{-1} + \mathbf{C}_2) \end{pmatrix}$$

gilt $\mathbf{U} + \mathbf{C} + \mathbf{H} = \mathbf{V}$. Da \mathbf{H} positiv semidefinit ist, folgt $\mathbf{U} + \mathbf{C} \leq \mathbf{V}$. Somit gilt $(\mathbf{U} + \mathbf{C}^{-1}) \geq \mathbf{V}^{-1}$. Es folgt

$$\begin{aligned} \mathbf{M}(\alpha \cdot \xi_1 + (1 - \alpha) \cdot \xi_2; \boldsymbol{\beta}) &= \frac{a}{b} \cdot \mathbf{X}_{\text{Ges}}^T (\mathbf{U} + \mathbf{C})^{-1} \mathbf{X}_{\text{Ges}} \\ &\geq \frac{a}{b} \cdot \mathbf{X}_{\text{Ges}}^T \mathbf{V}^{-1} \mathbf{X}_{\text{Ges}} = \alpha \cdot \mathbf{M}(\xi_1; \boldsymbol{\beta}) + (1 - \alpha) \cdot \mathbf{M}(\xi_2; \boldsymbol{\beta}) \end{aligned}$$

und somit die Konvexität der Informationsmatrizen auf der Menge der Designs. \square

A.4 Beweise und Hilfsresultat zu Kapitel 5

Beweis von Satz 5.3:

Nach Satz 5.2 sind die D -optimalen Gewichte für das Poisson-Gamma-Modell gleich den c -optimalen Gewichten für $\mathbf{c} = (0, 1)$ für ein Modell mit zugehöriger Informationsmatrix $\tilde{\mathbf{M}}(\xi; \boldsymbol{\beta}) = \sum_{j=1}^2 w_j \cdot \lambda(\beta_0 + \beta_1 x_j) \cdot \mathbf{f}(x_j) \mathbf{f}(x_j)^T$ mit $\lambda(z) = 1 + \frac{m}{b} \cdot \exp(z)$, da für eine Kovariable c -Optimalität und D_s -Optimalität übereinstimmen. Da die Designmatrix \mathbf{X} und $\mathbf{\Lambda} = \text{Diag}(\lambda(\beta_0 + \beta_1 x_1), \lambda(\beta_0 + \beta_1 x_2))$ regulär sind, folgt für den Vektor $\mathbf{v} = (v_1, v_2)^T$ aus Satz 3.10 c):

$$\begin{aligned} \mathbf{v} &= \left(\mathbf{\Lambda}^{\frac{1}{2}} \mathbf{X} \mathbf{X}^T \mathbf{\Lambda}^{\frac{1}{2}} \right)^{-1} \mathbf{\Lambda}^{\frac{1}{2}} \mathbf{X} \mathbf{c} = \mathbf{\Lambda}^{-\frac{1}{2}} (\mathbf{X}^T)^{-1} \mathbf{c} = \mathbf{\Lambda}^{-\frac{1}{2}} \cdot \frac{1}{\det(\mathbf{X})} \begin{pmatrix} x_2 & -1 \\ -x_1 & 1 \end{pmatrix} \mathbf{c} \\ &= \mathbf{\Lambda}^{-\frac{1}{2}} \cdot \frac{1}{\det(\mathbf{X})} \cdot (-1, 1)^T = \frac{1}{\det(\mathbf{X})} \cdot \left(-\frac{1}{\sqrt{\lambda(\beta_0 + \beta_1 x_1)}}, \frac{1}{\sqrt{\lambda(\beta_0 + \beta_1 x_2)}} \right)^T. \end{aligned}$$

Somit sind nach Satz 3.10 c) die optimalen Gewichte durch

$$\begin{aligned} w_1^* &= \frac{|v_1|}{|v_1| + |v_2|} = \frac{\sqrt{\lambda(\beta_0 + \beta_1 x_2)}}{\sqrt{\lambda(\beta_0 + \beta_1 x_1)} + \sqrt{\lambda(\beta_0 + \beta_1 x_2)}}, \\ w_2^* &= \frac{|v_2|}{|v_1| + |v_2|} = \frac{\sqrt{\lambda(\beta_0 + \beta_1 x_1)}}{\sqrt{\lambda(\beta_0 + \beta_1 x_1)} + \sqrt{\lambda(\beta_0 + \beta_1 x_2)}} \end{aligned}$$

gegeben. □

Beweis von Satz 5.8:

Der Beweis der D -Optimalität der Gewichte kann mit Satz 5.2 erfolgen. Alternativ können die Gewichte auch direkt optimiert werden, wie im Folgenden gezeigt wird.

Es sei $\delta_j = 1 + \frac{m}{b} \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta})$ für $j = 1, \dots, p$. Da $\mathbf{x}_1, \dots, \mathbf{x}_p$ affin unabhängig sind, sind die Regressionsfunktionen $\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_p)$ linear unabhängig und somit sind die Designmatrix \mathbf{X} und die Informationsmatrix $\mathbf{M}(\xi; \boldsymbol{\beta})$ regulär. Nach Satz 5.1 und mit der Matrixzerlegung der Informationsmatrix $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \mathbf{X}^T \mathbf{W} \mathbf{\Lambda} \mathbf{X}$ ist die Kriteriumsfunction durch

$$\begin{aligned} \det(\mathbf{M}(\xi; \boldsymbol{\beta})) &= \left(\frac{a}{b} \right)^p \cdot \frac{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}{1 + \frac{m}{b} \cdot \sum_{j=1}^p w_j \cdot \exp(\mathbf{f}(\mathbf{x}_j)^T \boldsymbol{\beta})} \\ &= \left(\frac{a}{b} \right)^p \cdot \frac{\prod_{j=1}^p w_j \cdot \det(\mathbf{X})^2 \cdot \det(\mathbf{\Lambda})}{\sum_{j=1}^p w_j \cdot \delta_j} \end{aligned}$$

gegeben. Da die Stützpunkte $\mathbf{x}_1, \dots, \mathbf{x}_{p-1}$ die Gleichung $\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} = c$ erfüllen, folgt $\delta_1 = \dots = \delta_{p-1}$. Somit ist die Funktion

$$\frac{\prod_{j=1}^p w_j}{\sum_{j=1}^p w_j \cdot \delta_j} = \frac{\prod_{j=1}^p w_j}{\delta_1 \cdot \sum_{j=1}^{p-1} w_j + w_p \cdot \delta_p} = \frac{\prod_{j=1}^p w_j}{\delta_1 \cdot (1 - w_p) + w_p \cdot \delta_p}$$

bezüglich der Gewichte w_1, \dots, w_p zu maximieren. Für festes w_p wird das Produkt für $w_1 = \dots = w_{p-1}$ maximal. Somit vereinfacht sich das Optimierungsproblem zur Maximierung des Ausdrucks $(w_1^{p-1} \cdot w_p) / [(1 - w_p) \cdot \delta_1 + w_p \cdot \delta_p]$. Unter Verwendung des Zusammenhangs $w_1 = (1 - w_p)/(p - 1)$ ist die folgende Funktion

$$g(w_p) = \frac{1}{(p-1)^{p-1}} \cdot \frac{(1-w_p)^{p-1} \cdot w_p}{(1-w_p) \cdot \delta_1 + w_p \cdot \delta_p}$$

bezüglich w_p zu maximieren. Die erste Ableitung ist durch

$$g'(w_p) = \left[\left(-(p-1) \cdot (1-w_p)^{p-2} \cdot w_p + (1-w_p)^{p-1} \right) \cdot \left((1-w_p) \cdot \delta_1 + w_p \cdot \delta_p \right) - (1-w_p)^{p-1} \cdot w_p \cdot (\delta_p - \delta_1) \right] \cdot \frac{1}{(p-1)^{p-1} \cdot \left((1-w_p) \cdot \delta_1 + w_p \cdot \delta_p \right)^2}$$

gegeben. Für $w_p \neq 1$ ist $g'(w_p) = 0$ äquivalent zu:

$$\begin{aligned} 0 &= \left(-(p-1) \cdot w_p + 1 - w_p \right) \cdot \left((1-w_p) \cdot \delta_1 + w_p \cdot \delta_p \right) - (1-w_p) \cdot w_p \cdot (\delta_p - \delta_1) \\ &= w_p^2 \cdot (p-1) \cdot (\delta_1 - \delta_p) - w_p \cdot p \cdot \delta_1 + \delta_1. \end{aligned} \tag{A.9}$$

Die Lösungen dieser quadratischen Gleichung sind durch

$$w_p = \frac{1}{2} \cdot \frac{p \cdot \delta_1 \pm \sqrt{p^2 \cdot \delta_1^2 - 4 \cdot (p-1) \cdot \delta_1 \cdot (\delta_1 - \delta_p)}}{(p-1) \cdot (\delta_1 - \delta_p)}$$

gegeben. Wegen $\mathbf{f}(\mathbf{x}_p)^T \boldsymbol{\beta} > c$ gilt $\delta_p > \delta_1$. Es folgt:

$$\sqrt{p^2 \cdot \delta_1^2 - 4 \cdot (p-1) \cdot \delta_1 \cdot (\delta_1 - \delta_p)} > p \cdot \delta_1$$

Somit ist die Lösung mit positivem Vorzeichen vor der Wurzel negativ und die Lösung mit negativem Vorzeichen vor der Wurzel positiv. Die positive Lösung wird im Folgenden betrachtet. Mit dem Zusammenhang $\sqrt{s} - \sqrt{t} = (s - t)/(\sqrt{s} + \sqrt{t})$ folgt:

$$w_p^* = \frac{1}{2 \cdot (p-1) \cdot (\delta_1 - \delta_p)} \cdot \frac{p^2 \cdot \delta_1^2 - (p^2 \cdot \delta_1^2 - 4 \cdot (p-1) \cdot \delta_1 \cdot (\delta_1 - \delta_p))}{p \cdot \delta_1 + \sqrt{p^2 \cdot \delta_1^2 - 4 \cdot (p-1) \cdot \delta_1 \cdot (\delta_1 - \delta_p)}}$$

$$\begin{aligned}
&= \frac{2 \cdot \delta_1}{p \cdot \delta_1 + \sqrt{p^2 \cdot \delta_1^2 - 4 \cdot (p-1) \cdot \delta_1 \cdot (\delta_1 - \delta_p)}} = \frac{2}{p + \sqrt{p^2 - 4 \cdot (p-1) \cdot \left(1 - \frac{\delta_p}{\delta_1}\right)}} \\
&= \frac{2}{p + \sqrt{(p-2)^2 + 4 \cdot (p-1) \cdot \frac{\delta_p}{\delta_1}}}.
\end{aligned}$$

Wegen $\delta_p/\delta_1 > 1$ gilt:

$$w_p^* < \frac{2}{p + \sqrt{(p-2)^2 + 4 \cdot (p-1)}} = \frac{2}{p + \sqrt{p^2}} = \frac{1}{p}.$$

Somit folgt $0 < w_p^* < 1/p$. Es gilt $g(0) = g(1) = 0$ und $g(w_p) > 0$ für $w_p \in (0, 1)$. Daher wird die Funktion g im Intervall $(0, 1)$ bei w_p^* maximal. Es folgt $w_1^* = (1 - w_p^*)/(p-1)$. Mit $1 = (p-1) \cdot w_1^* + w_p^* < (p-1) \cdot w_1^* + \frac{1}{p}$ und $1 = (p-1) \cdot w_1^* + w_p^* > (p-1) \cdot w_1^*$ folgt die Ungleichung $\frac{1}{p} < w_1^* < \frac{1}{p-1}$. \square

Lemma A.5

Es sei ξ ein Design mit Stützpunkten $\mathbf{x}_1, \dots, \mathbf{x}_p$ und D -optimalen Gewichten w_1^*, \dots, w_p^* wie in Satz 5.8. Dann gilt die folgende Gleichung:

$$\mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} \cdot w_p^* - w_1^* \cdot e^c + \frac{b}{m} \cdot (1 - p \cdot w_1^*) = 0. \quad (\text{A.10})$$

Beweis:

Es seien $\lambda_1 = \exp(c)$ und $\lambda_p = \exp(\mathbf{f}(\mathbf{x}_p)^T \boldsymbol{\beta})$ sowie $\delta_1 = 1 + \frac{m}{b} \cdot \lambda_1$ und $\delta_p = 1 + \frac{m}{b} \cdot \lambda_p$. Nach Gleichung (A.9) aus dem Beweis von Satz 5.8 erfüllen die D -optimalen Gewichte die Gleichung $0 = (w_p^*)^2 \cdot (p-1) \cdot (\delta_1 - \delta_p) - w_p^* \cdot p \cdot \delta_1 + \delta_1$. Es folgt:

$$\begin{aligned}
0 &= \frac{m}{b} \cdot (w_p^*)^2 \cdot (p-1) \cdot (\lambda_1 - \lambda_p) + \left(1 + \frac{m}{b} \cdot \lambda_1\right) \cdot (1 - p \cdot w_p^*) \\
&= \frac{m}{b} \cdot \left[(w_p^*)^2 \cdot (p-1) \cdot \lambda_1 - (w_p^*)^2 \cdot (p-1) \cdot \lambda_p + \left(\frac{b}{m} + \lambda_1\right) \cdot (1 - p \cdot w_p^*) \right].
\end{aligned}$$

Es gilt $w_p^* = 1 - (p-1) \cdot w_1^*$. Multiplikation mit b/m und Division durch $p-1$ ergibt:

$$\begin{aligned}
0 &= w_p^* \cdot (1 - (p-1) \cdot w_1^*) \cdot \lambda_1 - (w_p^*)^2 \cdot \lambda_p + \left(\frac{b}{m} + \lambda_1\right) \cdot \frac{1 - p \cdot w_p^*}{p-1} \\
&= -w_p^* \cdot (p-1) \cdot w_1^* \cdot \lambda_1 - (w_p^*)^2 \cdot \lambda_p + w_p^* \cdot \lambda_1 + \lambda_1 \cdot \frac{1 - p \cdot w_p^*}{p-1} + \frac{b}{m} \cdot \frac{1 - p \cdot w_p^*}{p-1} \\
&= -w_p^* \cdot ((p-1) \cdot w_1^* \cdot \lambda_1 + w_p^* \cdot \lambda_p) + \left(\frac{1 - w_p^*}{p-1}\right) \cdot \lambda_1 + \frac{b}{m} \cdot \frac{1 - p \cdot w_p^*}{p-1}
\end{aligned}$$

$$= -w_p^* \cdot \mathbf{e}_{1,p}^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_{1,p} + w_1^* \cdot \lambda_1 + \frac{b}{m} \cdot \frac{1 - p \cdot w_p^*}{p - 1}.$$

Mit dem Zusammenhang

$$\frac{1 - p \cdot w_p^*}{p - 1} = \frac{1 - p \cdot (1 - (p - 1) \cdot w_1^*)}{p - 1} = \frac{1 - p + p \cdot (p - 1) \cdot w_1^*}{p - 1} = -1 + p \cdot w_1^*$$

und Multiplikation mit -1 ergibt sich die im Lemma angegebene Gleichung. \square

A.5 Beweise und Hilfsresultate zu Kapitel 6

Beweis von Satz 6.2:

Es sei $\lambda(z) = \exp(\beta_0 + \beta_1 z)$ und $\delta(z) = 1 + \frac{m}{b} \cdot \lambda(z)$. Die Lösung $z^* > 0$ von Gleichung (6.4) hängt nicht von β_1 ab. Die Bedingung $z^* \leq \min\{|b_0|, |b_1|\} \cdot (v - u)$ sichert, dass für alle $\beta_1 \in [b_0, b_1]$ der Punkt $-z^*/\beta_1$ im Designraum $\mathcal{X} = [u, v]$ liegt. Nach Satz 5.6 besitzt das lokal D -optimale Design $\xi_{\boldsymbol{\beta}}^*$ die Stützpunkte $x_1^* = 0$ und $x_{2,\boldsymbol{\beta}}^* = -z^*/\beta_1$ mit Gewichten $w_{1,\boldsymbol{\beta}}^*$ und $w_{2,\boldsymbol{\beta}}^*$. Nach Gleichung (5.13) ist die Determinante der Informationsmatrix für $\xi_{\boldsymbol{\beta}}^*$ durch

$$\det(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*; \boldsymbol{\beta})) = \left(\frac{a}{b}\right)^2 \cdot \frac{\lambda(0) \cdot \lambda(-\frac{z^*}{\beta_1})}{(\sqrt{\delta(0)} + \sqrt{\delta(-z^*/\beta_1)})^2} \cdot \left(\frac{z^*}{\beta_1}\right)^2$$

gegeben. Als Kriteriumsfunktion wird $\text{eff}_D(\xi; \boldsymbol{\beta})^2 = \det(\mathbf{M}(\xi; \boldsymbol{\beta})) / \det(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*; \boldsymbol{\beta}))$ betrachtet. Für ein Design ξ mit Stützpunkten x_1, \dots, x_l und zugehörigen Gewichten w_1, \dots, w_l ergibt sie sich zu:

$$\frac{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi_{\boldsymbol{\beta}}^*; \boldsymbol{\beta}))} = \frac{\det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}))}{1 + \frac{m}{b} \cdot \mathbf{e}_1^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_1} \cdot \frac{(\sqrt{\delta(0)} + \sqrt{\delta(-\frac{z^*}{\beta_1})})^2 \cdot \beta_1^2}{\lambda(0) \cdot \lambda(-\frac{z^*}{\beta_1}) \cdot (z^*)^2}.$$

Die Terme $\lambda(0) = \exp(\beta_0)$, $\lambda(-z^*/\beta_1) = \exp(\beta_0 - z^*)$, $\delta(0) = 1 + (m/b) \cdot \exp(\beta_0)$ und $\delta(-z^*/\beta_1) = 1 + (m/b) \cdot \exp(\beta_0 - z^*)$ hängen nicht von dem Parameter β_1 ab. Es gilt:

$$\mathbf{e}_1^T \mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) \mathbf{e}_1 = \sum_{i=1}^l w_i \cdot \exp(\beta_0 + \beta_1 x_i).$$

Es gilt $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \mathbf{X}^T \boldsymbol{\Lambda} \mathbf{W} \mathbf{X}$ mit Designmatrix $\mathbf{X} = (\mathbf{1}_l, \mathbf{x})$, wobei $\mathbf{x} = (x_1, \dots, x_l)^T$ ist, sowie $\boldsymbol{\Lambda} = \text{Diag}(\exp(\beta_0 + \beta_1 x_1), \dots, \exp(\beta_0 + \beta_1 x_l))$ und $\mathbf{W} = \text{Diag}(w_1, \dots, w_l)$. Mit $\tilde{\mathbf{X}} = \boldsymbol{\Lambda}^{\frac{1}{2}} \mathbf{W}^{\frac{1}{2}} \mathbf{X}$ folgt $\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta}) = \tilde{\mathbf{X}}^T \tilde{\mathbf{X}}$. Für $S \subseteq \{1, \dots, l\}$ sei $\tilde{\mathbf{X}}_S$ die Matrix, welche nur aus den Zeilen von $\tilde{\mathbf{X}}$ besteht, deren Indices in S liegen. Für $S = \{i, j\}$ mit $i < j$

gilt somit $\tilde{\mathbf{X}}_S = \text{Diag}(\exp(\beta_0 + \beta_1 x_i), \exp(\beta_0 + \beta_1 x_j))^{\frac{1}{2}} \cdot \text{Diag}(w_i, w_j)^{\frac{1}{2}} \cdot (\mathbf{1}_2, (x_i, x_j)^T)$. Mit dem Satz von Binet-Cauchy für die Determinante des Produkts von zwei Matrizen folgt (vgl. Harville, 1997, S. 203–205):

$$\begin{aligned} \det(\mathbf{M}_{\text{Po}}(\xi; \boldsymbol{\beta})) &= \det(\tilde{\mathbf{X}}^T \tilde{\mathbf{X}}) = \sum_{\substack{S \subseteq \{1, \dots, l\} \\ |S|=2}} \det((\tilde{\mathbf{X}}_S)^T) \cdot \det(\tilde{\mathbf{X}}_S) = \sum_{\substack{S \subseteq \{1, \dots, l\} \\ |S|=2}} \det(\tilde{\mathbf{X}}_S)^2 \\ &= \sum_{i < j} w_i \cdot w_j \cdot \exp(\beta_0 + \beta_1 x_i) \cdot \exp(\beta_0 + \beta_1 x_j) \cdot (x_j - x_i)^2. \end{aligned}$$

Somit treten der Parameter β_1 und die Stützpunkte x_i in der Kriteriumsfunktion

$$\begin{aligned} \frac{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi^*; \boldsymbol{\beta}))} &= \frac{(\sqrt{\delta(0)} + \sqrt{\delta(-\frac{z^*}{\beta_1})})^2}{\lambda(0) \cdot \lambda(-\frac{z^*}{\beta_1}) \cdot (z^*)^2} \\ &\cdot \frac{\sum_{i < j} w_i \cdot w_j \cdot \exp(\beta_0 + \beta_1 x_i) \cdot \exp(\beta_0 + \beta_1 x_j) \cdot (\beta_1 x_j - \beta_1 x_i)^2}{1 + \frac{m}{b} \cdot \sum_{i=1}^l w_i \cdot \exp(\beta_0 + \beta_1 x_i)} \end{aligned}$$

nur als Produkt $\beta_1 \cdot x_i$ auf. Sie lässt sich somit als Funktion $g(\beta_1 \cdot x_1, \dots, \beta_1 \cdot x_l)$ darstellen. Für $\mathcal{X} = [u, v]$ und $\beta_1 \in [b_0, b_1]$ sei ξ^* mit Stützpunkten x_1^*, \dots, x_l^* und Gewichten w_1^*, \dots, w_l^* standardisiert Maximin D -optimal. Für $c \neq 0$ wird die Transformation $h: \mathcal{X} \rightarrow \mathcal{X}_c = \frac{1}{c} \cdot [u, v]$, $x \mapsto z = x/c$ betrachtet. Mit $z_i^* = h(x_i^*) = x_i^*/c$ gilt $g(\beta_1 \cdot x_1^*, \dots, \beta_1 \cdot x_l^*) = g(\beta_1 \cdot c \cdot z_1^*, \dots, \beta_1 \cdot c \cdot z_l^*)$. Somit ist das Design ξ_c^* mit Stützpunkten z_1^*, \dots, z_l^* und Gewichten w_1^*, \dots, w_l^* für $\beta_1 \in c \cdot [b_0, b_1]$ und $\mathcal{X}_c = \frac{1}{c} \cdot [u, v]$ standardisiert Maximin D -optimal und besitzt die gleiche minimale Effizienz wie ξ^* . \square

Lemma A.6

Es sei $g(x) = (1+x) \cdot x^{\frac{x}{1-x}}$. Dann gilt $0 < g(x) < 1$ für $x \in (0, 1)$.

Beweis:

Da $x \in (0, 1)$ ist, gilt $g(x) > 0$. Da $\lim_{x \searrow 0} x \cdot \ln(x) = 0$ ist, folgt für $g(x)$:

$$\lim_{x \searrow 0} g(x) = \lim_{x \searrow 0} (1+x) \cdot \exp\left(\frac{\ln(x) \cdot x}{1-x}\right) = 1.$$

Es gilt mit dem Satz von L'Hospital:

$$\lim_{x \rightarrow 1} \frac{\ln(x)}{1-x} = \lim_{x \rightarrow 1} \frac{\frac{1}{x}}{-1} = -1.$$

Somit folgt:

$$\lim_{x \rightarrow 1} g(x) = \lim_{x \rightarrow 1} (1+x) \cdot \exp\left(\frac{\ln(x) \cdot x}{1-x}\right) = 2 \cdot \exp(-1) < 1.$$

Für die Ableitung von $g(x)$ gilt unter Verwendung von $x^{\frac{x}{1-x}} = \exp(\ln(x) \cdot x/(1-x))$:

$$\begin{aligned} g'(x) &= x^{\frac{x}{1-x}} + (1+x) \cdot x^{\frac{x}{1-x}} \cdot \left(\frac{\left(\frac{1}{x} \cdot x + \ln(x)\right) \cdot (1-x) + \ln(x) \cdot x}{(1-x)^2} \right) \\ &= x^{\frac{x}{1-x}} \cdot \left(1 + (1+x) \cdot \frac{1-x + \ln(x)}{(1-x)^2} \right) = x^{\frac{x}{1-x}} \cdot \left(\frac{x \cdot \ln(x) + \ln(x) - 2 \cdot x + 2}{(1-x)^2} \right). \end{aligned}$$

Es gilt genau dann $g'(x) < 0$ für $x \in (0, 1)$, wenn $h(x) = x \cdot \ln(x) + \ln(x) - 2 \cdot x + 2 < 0$ für $x \in (0, 1)$ ist. Es gilt $\lim_{x \searrow 0} h(x) = -\infty$ und $\lim_{x \rightarrow 1} h(x) = 0$. Die Ableitung von $h(x)$ berechnet sich zu $h'(x) = \ln(x) + \frac{1}{x} - 1$. Für die Grenzwerte der Ableitung gilt $\lim_{x \searrow 0} h'(x) = \lim_{x \searrow 0} \frac{1}{x} \cdot (x \cdot \ln(x) + 1) - 1 = \infty$ und $\lim_{x \rightarrow 1} h'(x) = 0$. Die zweite Ableitung von $h(x)$ ist durch $h''(x) = \frac{1}{x} - \frac{1}{x^2} = \frac{x-1}{x^2}$ gegeben. Für $x \in (0, 1)$ gilt somit:

$$h''(x) < 0 \Rightarrow h'(x) > 0 \Rightarrow h(x) < 0 \Rightarrow g'(x) < 0.$$

Daher folgt unter Verwendung der Grenzwerte $2 \cdot \exp(-1) < g(x) < 1$ für $x \in (0, 1)$. \square

Lemma A.7

Es sei $\lambda_{\beta_1}(x) = \exp(\beta_0 + \beta_1 x)$, $\delta_{\beta_1}(x) = 1 + \frac{m}{b} \cdot \lambda_{\beta_1}(x)$ und $\delta(0) = 1 + \frac{m}{b} \cdot \exp(\beta_0)$. Weiterhin sei

$$w_{b_0, b_1}(x) := \frac{\lambda_{b_1}(x) \cdot b_1^2 \cdot \delta_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \delta_{b_1}(x)}{\lambda_{b_0}(x) \cdot b_0^2 \cdot (\delta(0) - \delta_{b_1}(x)) - \lambda_{b_1}(x) \cdot b_1^2 \cdot (\delta(0) - \delta_{b_0}(x))}. \quad (\text{A.11})$$

Für $b_0 < b_1 < 0$ und $\tilde{x} = 2 \cdot \ln(b_1/b_0)/(b_0 - b_1)$ besitzt $w_{b_0, b_1}(x)$ im Intervall $(0, \infty)$ eine eindeutige Nullstelle $x_0 > \tilde{x}$ und es gilt $w_{b_0, b_1}(\tilde{x}) = 1$. Außerdem gilt $0 < w_{b_0, b_1}(x) < 1$ nur für $x \in (\tilde{x}, x_0)$. Auf dem Intervall (\tilde{x}, x_0) ist $w_{b_0, b_1}(x)$ streng monoton fallend und besitzt die folgende Darstellung:

$$w_{b_0, b_1}(x) = \frac{1}{1 - \frac{\delta(0)}{1 + \frac{m}{b} \cdot \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\lambda_{b_1}(x)} - \frac{b_1^2}{\lambda_{b_0}(x)}}}}. \quad (\text{A.12})$$

Für $0 < b_0 < b_1$ gilt $w_{b_0, b_1}(x) = w_{-b_1, -b_0}(-x)$.

Beweis:

Es sei $b_0 < b_1 < 0$. Die Indices von $w_{b_0, b_1}(x)$ werden zur Vereinfachung der Notation zunächst weggelassen.

Zuerst wird gezeigt, dass Zähler und Nenner von $w(x)$ keine gemeinsamen Nullstellen besitzen. Es wird angenommen, dass der Zähler gleich 0 ist. Dann vereinfacht sich der Nenner zu $\delta(0) \cdot (\lambda_{b_0}(x) \cdot b_0^2 - \lambda_{b_1}(x) \cdot b_1^2)$. Wäre der Nenner auch gleich 0, dann muss $\lambda_{b_0}(x) \cdot b_0^2 = \lambda_{b_1}(x) \cdot b_1^2$ gelten. Für den Zähler würde dann gelten:

$$0 = \lambda_{b_1}(x) \cdot b_1^2 \cdot \delta_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \delta_{b_1}(x) = \delta_{b_0}(x) - \delta_{b_1}(x)$$

Es würde $\delta_{b_0}(x) = \delta_{b_1}(x)$ und somit $b_0 = b_1$ folgen, was einen Widerspruch ergibt. Für $\lambda_{b_1}(x) \cdot b_1^2 \cdot \delta_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \delta_{b_1}(x) \neq 0$ lässt sich $w(x)$ umschreiben als:

$$w(x) = \frac{1}{1 + \frac{\lambda_{b_0}(x) \cdot b_0^2 \cdot \delta(0) - \lambda_{b_1}(x) \cdot b_1^2 \cdot \delta(0)}{\lambda_{b_1}(x) \cdot b_1^2 \cdot \delta_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \delta_{b_1}(x)}}. \quad (\text{A.13})$$

Weiteres Umformen von $w(x)$ ergibt:

$$w(x) = \frac{1}{1 + \frac{\delta(0) \cdot (\lambda_{b_0}(x) \cdot b_0^2 - \lambda_{b_1}(x) \cdot b_1^2)}{\lambda_{b_1}(x) \cdot b_1^2 - \lambda_{b_0}(x) \cdot b_0^2 + \frac{m}{b} \cdot (\lambda_{b_1}(x) \cdot b_1^2 \cdot \lambda_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \lambda_{b_1}(x))}}.$$

Für $\lambda_{b_0}(x) \cdot b_0^2 - \lambda_{b_1}(x) \cdot b_1^2 \neq 0$ folgt für $w(x)$:

$$\begin{aligned} w(x) &= \frac{1}{1 + \frac{\delta(0)}{-1 + \frac{m}{b} \cdot \frac{\lambda_{b_1}(x) \cdot b_1^2 \cdot \lambda_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \lambda_{b_1}(x)}{\lambda_{b_0}(x) \cdot b_0^2 - \lambda_{b_1}(x) \cdot b_1^2}}} \\ &= \frac{1}{1 - \frac{\delta(0)}{1 + \frac{m}{b} \cdot \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\lambda_{b_1}(x)} - \frac{b_1^2}{\lambda_{b_0}(x)}}}}. \end{aligned} \quad (\text{A.14})$$

Im Folgenden wird der Verlauf der Funktion $w(x)$ charakterisiert. Für Zähler und Nenner von $w(x)$ in Gleichung (A.11) gelten die folgenden Grenzwerte:

$$\lim_{x \searrow 0} \lambda_{b_1}(x) \cdot b_1^2 \cdot \delta_{b_0}(x) - \lambda_{b_0}(x) \cdot b_0^2 \cdot \delta_{b_1}(x) = \lambda(0) \cdot \delta(0) \cdot (b_1^2 - b_0^2) < 0,$$

$$\begin{aligned} & \lim_{x \searrow 0} \lambda_{b_0}(x) \cdot b_0^2 \cdot (\delta(0) - \delta_{b_1}(x)) - \lambda_{b_1}(x) \cdot b_1^2 \cdot (\delta(0) - \delta_{b_0}(x)) \\ &= \lambda(0) \cdot b_0^2 \cdot (\delta(0) - \delta(0)) - \lambda(0) \cdot b_1^2 \cdot (\delta(0) - \delta(0)) = 0. \end{aligned}$$

Somit folgt $\lim_{x \searrow 0} w(x) = -\infty$. Mit $b_0 - b_1 < 0$ gilt:

$$\lim_{x \rightarrow \infty} \frac{b_0^2}{\lambda_{b_1}(x)} - \frac{b_1^2}{\lambda_{b_0}(x)} = \lim_{x \rightarrow \infty} \exp(-\beta_0 - b_0 \cdot x) \cdot (b_0^2 \cdot \exp((b_0 - b_1) \cdot x) - b_1^2) = -\infty.$$

Damit folgt der Grenzwert von $w(x)$ für $x \rightarrow \infty$ unter Verwendung der Darstellung in Gleichung (A.14):

$$\lim_{x \rightarrow \infty} w(x) = \frac{1}{1 - \delta(0)} = -\frac{b}{m} \cdot \exp(-\beta_0) < 0. \quad (\text{A.15})$$

Aus Gleichung (A.13) folgt, dass genau dann $w(x) = 1$ gilt, wenn $\lambda_{b_0}(x) \cdot b_0^2 - \lambda_{b_1}(x) \cdot b_1^2 = 0$ erfüllt ist. Es gilt äquivalent:

$$\exp((b_0 - b_1) \cdot x) = \frac{b_1^2}{b_0^2} \Leftrightarrow x = \frac{2 \cdot \ln\left(\frac{b_1}{b_0}\right)}{b_0 - b_1}.$$

Insbesondere nimmt die Funktion $w(x)$ den Wert 1 nur an der Stelle $x = \tilde{x}$ an, wobei $\tilde{x} = 2 \cdot \ln(b_1/b_0)/(b_0 - b_1) > 0$ ist.

Es sei $g(x)$ der Zähler und $h(x)$ der Nenner von $w(x)$ in Gleichung (A.11). Die Ableitungen von $g(x)$ und $h(x)$ nach x berechnen sich zu:

$$\begin{aligned} g'(x) &= \lambda_{b_1}(x) \cdot b_1^2 \cdot \left(\delta_{b_0}(x) \cdot b_1 + \frac{m}{b} \cdot b_0 \cdot \lambda_{b_0}(x) \right) \\ &\quad - \lambda_{b_0}(x) \cdot b_0^2 \cdot \left(\delta_{b_1}(x) \cdot b_0 + \frac{m}{b} \cdot b_1 \cdot \lambda_{b_1}(x) \right), \\ h'(x) &= \lambda_{b_0}(x) \cdot b_0^2 \cdot \left((\delta(0) - \delta_{b_1}(x)) \cdot b_0 - \frac{m}{b} \cdot b_1 \cdot \lambda_{b_1}(x) \right) \\ &\quad - \lambda_{b_1}(x) \cdot b_1^2 \cdot \left((\delta(0) - \delta_{b_0}(x)) \cdot b_1 - \frac{m}{b} \cdot b_0 \cdot \lambda_{b_0}(x) \right). \end{aligned}$$

Mit der Quotientenregel gilt für die Ableitung von $w(x)$ nach x :

$$w'(x) = \frac{g'(x) \cdot h(x) - g(x) \cdot h'(x)}{h(x)^2}.$$

Der Zähler vereinfacht sich zu:

$$\begin{aligned}
& g'(x) \cdot h(x) - g(x) \cdot h'(x) \\
&= b_0^2 \cdot b_1^3 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta_{b_0}(x) \cdot \delta(0) - b_0^2 \cdot b_1^3 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta_{b_1}(x) \cdot \delta(0) \\
&+ b_0^3 \cdot b_1^2 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta_{b_1}(x) \cdot \delta(0) - b_0^3 \cdot b_1^2 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta_{b_0}(x) \cdot \delta(0) \\
&+ \frac{m}{b} \cdot b_0^3 \cdot b_1^2 \cdot \lambda_{b_0}(x)^2 \cdot \lambda_{b_1}(x) \cdot \delta(0) + \frac{m}{b} \cdot b_0^2 \cdot b_1^3 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x)^2 \cdot \delta(0) \\
&- \frac{m}{b} \cdot b_0^4 \cdot b_1 \cdot \lambda_{b_0}(x)^2 \cdot \lambda_{b_1}(x) \cdot \delta(0) - \frac{m}{b} \cdot b_0 \cdot b_1^4 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x)^2 \cdot \delta(0) \\
&= \frac{m}{b} \cdot b_0^2 \cdot b_1^3 \cdot \lambda_{b_0}(x)^2 \cdot \lambda_{b_1}(x) \cdot \delta(0) + \frac{m}{b} \cdot b_0^3 \cdot b_1^2 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x)^2 \cdot \delta(0) \\
&\quad - \frac{m}{b} \cdot b_0^4 \cdot b_1 \cdot \lambda_{b_0}(x)^2 \cdot \lambda_{b_1}(x) \cdot \delta(0) - \frac{m}{b} \cdot b_0 \cdot b_1^4 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x)^2 \cdot \delta(0) \\
&= \frac{m}{b} \cdot b_0 \cdot b_1 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta(0) \cdot (b_0 \cdot b_1^2 \cdot \lambda_{b_0}(x) + b_0^2 \cdot b_1 \cdot \lambda_{b_1}(x) - b_0^3 \cdot \lambda_{b_0}(x) \\
&\quad - b_1^3 \cdot \lambda_{b_1}(x)) \\
&= \frac{m}{b} \cdot b_0 \cdot b_1 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta(0) \cdot (b_0^2 - b_1^2) \cdot (b_1 \cdot \lambda_{b_1}(x) - b_0 \cdot \lambda_{b_0}(x)) \\
&= c_3(x) \cdot (b_1 \cdot \lambda_{b_1}(x) - b_0 \cdot \lambda_{b_0}(x)).
\end{aligned}$$

Dabei gilt $c_3(x) = \frac{m}{b} \cdot b_0 \cdot b_1 \cdot \lambda_{b_0}(x) \cdot \lambda_{b_1}(x) \cdot \delta(0) \cdot (b_0^2 - b_1^2) > 0$. Es gilt daher genau dann $w'(x) = 0$, wenn $b_1 \cdot \lambda_{b_1}(x) - b_0 \cdot \lambda_{b_0}(x) = 0$ gilt. Als eindeutige Lösung ergibt sich $x = \ln(b_1/b_0)/(b_0 - b_1) = \tilde{x}/2$. Die Nullstelle von $w'(x)$ liegt links von der Stelle $x = \tilde{x}$, bei der die Funktion $w(x)$ den Wert 1 annimmt.

Nun wird gezeigt, dass $w(\tilde{x}/2) < 0$ gilt. Mit der Darstellung von $w(x)$ in Gleichung (A.14) folgt, dass $w(\tilde{x}/2) < 0$ äquivalent ist zu:

$$\begin{aligned}
& \frac{\delta(0)}{1 + \frac{m}{b} \cdot \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\lambda_{b_1}(\tilde{x}/2)} - \frac{b_1^2}{\lambda_{b_0}(\tilde{x}/2)}}} > 1 \\
\Leftrightarrow \delta(0) > 1 + \frac{m}{b} \cdot \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\lambda_{b_1}(\tilde{x}/2)} - \frac{b_1^2}{\lambda_{b_0}(\tilde{x}/2)}} > 0 \\
\Leftrightarrow 1 > \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\exp(b_1 \cdot \tilde{x}/2)} - \frac{b_1^2}{\exp(b_0 \cdot \tilde{x}/2)}} > -\frac{b}{m} \cdot \exp(-\beta_0).
\end{aligned}$$

Es gilt:

$$\begin{aligned}
\frac{\frac{b_0^2 - b_1^2}{b_0^2} - \frac{b_0^2 - b_1^2}{b_1^2}}{\exp(b_1 \cdot \tilde{x}/2) - \exp(b_0 \cdot \tilde{x}/2)} &= \frac{\frac{b_0^2 - b_1^2}{b_0^2} - \frac{b_0^2 - b_1^2}{b_1^2}}{\left(\frac{b_1}{b_0}\right)^{\frac{b_1}{b_0 - b_1}} - \left(\frac{b_1}{b_0}\right)^{\frac{b_0}{b_0 - b_1}}} = \left(\frac{b_1}{b_0}\right)^{\frac{b_1}{b_0 - b_1}} \cdot \frac{b_0^2 - b_1^2}{b_0^2 - b_0 \cdot b_1} \\
&= \left(\frac{b_1}{b_0}\right)^{\frac{b_1}{b_0 - b_1}} \cdot \frac{(b_0 + b_1) \cdot (b_0 - b_1)}{b_0 \cdot (b_0 - b_1)} \\
&= \left(\frac{b_1}{b_0}\right)^{\frac{b_1/b_0}{1 - b_1/b_0}} \cdot \left(1 + \frac{b_1}{b_0}\right).
\end{aligned}$$

Da $0 < b_1/b_0 < 1$ gilt, folgt mit Lemma A.6, dass $w(\tilde{x}/2) < 0$ ist.

Es sei $n \in \mathbb{N}$ und $n \geq 2$. Es wird nun gezeigt, dass $w'(n \cdot \tilde{x}/2) < 0$ gilt. Dazu wird zunächst der Nenner betrachtet. Es gilt:

$$\begin{aligned}
h(n \cdot \tilde{x}/2) &= h\left(n \cdot \frac{\ln\left(\frac{b_1}{b_0}\right)}{b_0 - b_1}\right) \\
&= \frac{m}{b} \cdot e^{2 \cdot \beta_0} \cdot \left[\left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_0}{b_0 - b_1}} \cdot b_0^2 \cdot \left(1 - \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}}\right) - \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}} \cdot b_1^2 \cdot \left(1 - \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_0}{b_0 - b_1}}\right) \right] \\
&= \frac{m}{b} \cdot e^{2 \cdot \beta_0} \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_0}{b_0 - b_1}} \cdot \left[b_0^2 - b_0^2 \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}} - \left(\frac{b_1}{b_0}\right)^{-n} \cdot b_1^2 + b_1^2 \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}} \right] \\
&= \frac{m}{b} \cdot e^{2 \cdot \beta_0} \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_0}{b_0 - b_1}} \cdot \left[b_0^2 \cdot \left(1 - \left(\frac{b_0}{b_1}\right)^{n-2}\right) + (b_1^2 - b_0^2) \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}} \right].
\end{aligned}$$

Da $n \geq 2$ gilt, folgt $1 - (b_0/b_1)^{n-2} \leq 0$. Außerdem gilt $b_1^2 - b_0^2 < 0$. Es folgt somit $h(n \cdot \tilde{x}/2) < 0$ und $h(n \cdot \tilde{x}/2)^2 > 0$. Nun wird gezeigt, dass der Zähler negativ ist:

$$\begin{aligned}
&g'(n \cdot \tilde{x}/2) \cdot h(n \cdot \tilde{x}/2) - g(n \cdot \tilde{x}/2) \cdot h'(n \cdot \tilde{x}/2) \\
&= c_3(n \cdot \tilde{x}/2) \cdot e^{\beta_0} \cdot \left[b_1 \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}} - b_0 \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_0}{b_0 - b_1}} \right] \\
&= c_3(n \cdot \tilde{x}/2) \cdot e^{\beta_0} \cdot b_1 \cdot \left(\frac{b_1}{b_0}\right)^{\frac{n \cdot b_1}{b_0 - b_1}} \cdot \left[1 - \left(\frac{b_1}{b_0}\right)^{n-1} \right].
\end{aligned}$$

Da $b_1 < 0$ und $b_1/b_0 < 1$ gilt, ist der Zähler von $w'(n \cdot \tilde{x}/2)$ negativ und es folgt

$w'(n \cdot \tilde{x}/2) < 0$ für $n \geq 2$. Insbesondere ergibt der Fall $n = 2$, dass $w(x)$ bei $x = \tilde{x}$ fallend ist.

Aus den gezeigten Eigenschaften der Funktion $w(x)$ lässt sich ihr Verlauf herleiten. Die Funktion $w(x)$ ist streng monoton wachsend auf dem Intervall $(0, \tilde{x}/2)$. Bei dem einzigen Extremum $\tilde{x}/2$ gilt $w(\tilde{x}/2) < 0$. Es liegt bei $\tilde{x}/2$ kein Sattelpunkt vor, da sonst wegen $w'(\tilde{x}) < 0$ ein $\hat{x} \in (\tilde{x}/2, \tilde{x})$ mit $w(\hat{x}) = 1$ existieren müsste. Dies ergibt einen Widerspruch, da $w(x) = 1$ nur für $x = \tilde{x}$ gilt. Somit handelt es sich bei $\tilde{x}/2$ um ein Maximum.

Da $w'(\tilde{x}) < 0$ und $w(x) = 1$ nur für $x = \tilde{x}$ gilt, muss sich im Intervall $(\tilde{x}/2, \tilde{x})$ genau eine Polstelle x_{Pol} befinden, für die $\lim_{x \nearrow x_{\text{Pol}}} w(x) = -\infty$ und $\lim_{x \searrow x_{\text{Pol}}} w(x) = \infty$ gilt. Es kann im Intervall $[\tilde{x}, \infty)$ keine Polstellen geben. Angenommen es gibt eine Polstelle $\hat{x}_{\text{Pol}} \in [\tilde{x}, \infty)$. Der Fall $\lim_{x \searrow \hat{x}_{\text{Pol}}} w(x) = -\infty$ führt zum Widerspruch, da $w'(n \cdot \tilde{x}/2) < 0$ für $n \geq 2$ gilt. Der Fall $\lim_{x \searrow \hat{x}_{\text{Pol}}} w(x) = \infty$ führt zum Widerspruch, da für $x \rightarrow \infty$ die Funktion $w(x)$ gegen einen negativen Wert konvergiert und somit ein $\hat{x} \neq \tilde{x}$ mit $w(\hat{x}) = 1$ existieren müsste.

Somit ist $w(x)$ auf den Intervallen $(\tilde{x}/2, x_{\text{Pol}})$ und (x_{Pol}, ∞) streng monoton fallend. Insbesondere besitzt $w(x)$ im Intervall $(0, \infty)$ eine eindeutige Nullstelle $x_0 > \tilde{x}$. Es gilt $0 < w(x) < 1$ nur für $x \in (\tilde{x}, x_0)$. Auf diesem Intervall ist $w(x)$ streng monoton fallend und besitzt die Darstellung in Gleichung (A.14).

Der beschriebene Verlauf von $w(x)$ wird in Abbildung A.1 grafisch dargestellt, wobei der Teil des Graphen grün eingezeichnet ist, für den $0 \leq w(x) \leq 1$ gilt.

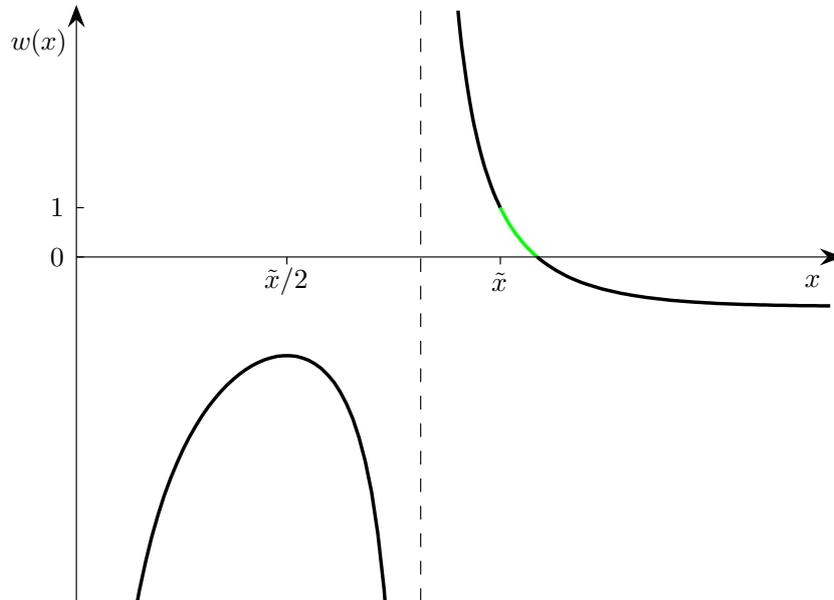


Abbildung A.1: Graph der Funktion $w_{b_0, b_1}(x)$ für $b_0 < b_1 < 0$

Für $0 < b_0 < b_1$ folgt $w_{b_0, b_1}(x) = w_{-b_1, -b_0}(-x)$, da $\lambda_{b_i}(x) = \lambda_{-b_i}(-x)$ für $i = 1, 2$ gilt. \square

Beweis von Satz 6.3:

Es wird der Fall $b_1 < 0$ und $u = 0$ betrachtet. Der Fall $b_0 > 0$ und $v = 0$ folgt analog. Es seien $\lambda(z) = \lambda_{\beta_1}(z) = \exp(\beta_0 + \beta_1 z)$ und $\delta(z) = \delta_{\beta_1}(z) = 1 + \frac{m}{b} \cdot \lambda(z)$. Die Lösung $z^* > 0$ von Gleichung (6.4) hängt nicht von dem Parameter β_1 ab. Die Bedingung $z^* \leq \min\{|b_0|, |b_1|\} \cdot (v - u) = |b_1| \cdot v$ sichert, dass für alle $\beta_1 \in [b_0, b_1]$ der Punkt $-z^*/\beta_1$ im Designraum $\mathcal{X} = [0, v]$ liegt. Nach Satz 5.6 besitzt das lokal D -optimale Design ξ_{β}^* die Stützpunkte $x_1^* = 0$ und $x_{2,\beta}^* = -z^*/\beta_1$ mit Gewichten $w_{1,\beta}^*$ und $w_{2,\beta}^*$. Für die Determinante der Informationsmatrix für ξ_{β}^* gilt nach Gleichung (5.13):

$$\det(\mathbf{M}(\xi_{\beta}^*; \boldsymbol{\beta})) = \left(\frac{a}{b}\right)^2 \cdot \frac{\lambda(0) \cdot \lambda\left(-\frac{z^*}{\beta_1}\right)}{\left(\sqrt{\delta(0)} + \sqrt{\delta\left(-\frac{z^*}{\beta_1}\right)}\right)^2} \cdot \left(\frac{z^*}{\beta_1}\right)^2.$$

Für ein Design ξ mit Stützpunkten $x_1 < x_2$ und zugehörigen Gewichten w und $1 - w$ gilt für die Determinante der Informationsmatrix nach Gleichung (5.12):

$$\det(\mathbf{M}(\xi; \boldsymbol{\beta})) = \frac{w \cdot (1 - w) \cdot \lambda(x_1) \cdot \lambda(x_2) \cdot (x_2 - x_1)^2}{w \cdot \delta(x_1) + (1 - w) \cdot \delta(x_2)}.$$

Als Kriteriumsfunktion wird $\text{eff}_D(\xi; \boldsymbol{\beta})^2 = \det(\mathbf{M}(\xi; \boldsymbol{\beta})) / \det(\mathbf{M}(\xi_{\beta}^*; \boldsymbol{\beta}))$ betrachtet. Für ein Design ξ mit Stützpunkten $x_1 < x_2$ und zugehörigen Gewichten w und $1 - w$ ergibt sie sich zu:

$$\begin{aligned} g(w, \beta_1, x_1, x_2) &= \frac{\det(\mathbf{M}(\xi; \boldsymbol{\beta}))}{\det(\mathbf{M}(\xi_{\beta}^*; \boldsymbol{\beta}))} \\ &= \frac{w \cdot (1 - w) \cdot \lambda(x_1) \cdot \lambda(x_2) \cdot (x_2 - x_1)^2}{w \cdot \delta(x_1) + (1 - w) \cdot \delta(x_2)} \cdot \frac{\left(\sqrt{\delta(0)} + \sqrt{\delta\left(-\frac{z^*}{\beta_1}\right)}\right)^2 \cdot \beta_1^2}{\lambda(0) \cdot \lambda\left(-\frac{z^*}{\beta_1}\right) \cdot (z^*)^2}. \end{aligned}$$

Die Kriteriumsfunktion lässt sich schreiben als

$$g(w, \beta_1, x_1, x_2) = c_1(w, \beta_1, x_2) \cdot (x_2 - x_1)^2 \cdot \left(w \cdot \left(\frac{1}{\lambda(x_1)} + \frac{m}{b} \right) + (1 - w) \cdot \frac{\delta(x_2)}{\lambda(x_1)} \right)^{-1}$$

mit einer nicht von x_1 abhängenden Funktion $c_1(w, \beta_1, x_2)$. Da $\beta_1 < 0$ gilt, ist $\lambda(x_1)$ streng monoton fallend in x_1 . Es folgt, dass $g(w, \beta_1, x_1, x_2)$ ebenfalls streng monoton fallend in x_1 ist und daher auf dem Intervall $[0, v]$ bei $x_1 = 0$ maximiert wird. Somit lässt sich die Kriteriumsfunktion zu

$$\tilde{g}(w, \beta_1, x_2) = g(w, \beta_1, 0, x_2) = \frac{w \cdot (1 - w) \cdot \lambda(x_2) \cdot x_2^2}{w \cdot \delta(0) + (1 - w) \cdot \delta(x_2)} \cdot \frac{\left(\sqrt{\delta(0)} + \sqrt{\delta\left(-\frac{z^*}{\beta_1}\right)}\right)^2 \cdot \beta_1^2}{\lambda\left(-\frac{z^*}{\beta_1}\right) \cdot (z^*)^2}$$

vereinfachen.

Es wird nun gezeigt, dass die Funktion $\tilde{g}(w, \beta_1, x_2)$ bezüglich β_1 unimodal auf dem Intervall $[b_0, b_1]$ ist. Die Terme $\lambda(-z^*/\beta_1) = \exp(\beta_0 - z^*)$, $\delta(0) = 1 + (m/b) \cdot \exp(\beta_0)$ und $\delta(-z^*/\beta_1) = 1 + (m/b) \cdot \exp(\beta_0 - z^*)$ hängen nicht von dem Parameter β_1 ab. Mit $\partial\delta(x_2)/\partial\beta_1 = (m/b) \cdot x_2 \cdot \lambda(x_2)$ folgt für die Ableitung der Kriteriumsfunktion $\tilde{g}(w, \beta_1, x_2)$ nach β_1 :

$$\frac{\partial\tilde{g}(w, \beta_1, x_2)}{\partial\beta_1} = \frac{w \cdot (1-w) \cdot x_2^2 \cdot (\sqrt{\delta(0)} + \sqrt{\delta(-\frac{z^*}{\beta_1})})^2}{\lambda(-\frac{z^*}{\beta_1}) \cdot (z^*)^2} \cdot \left(\frac{x_2 \cdot \lambda(x_2) \cdot \beta_1^2 + \lambda(x_2) \cdot 2 \cdot \beta_1}{w \cdot \delta(0) + (1-w) \cdot \delta(x_2)} - \frac{\lambda(x_2)^2 \cdot \beta_1^2 \cdot (1-w) \cdot \frac{m}{b} \cdot x_2}{(w \cdot \delta(0) + (1-w) \cdot \delta(x_2))^2} \right).$$

Es gilt $\partial\tilde{g}(w, \beta_1, x_2)/\partial\beta_1 = 0$, genau dann wenn $\beta_1 = 0$ oder

$$\begin{aligned} 0 &= (x_2 \cdot \beta_1 + 2) \cdot (w \cdot \delta(0) + (1-w) \cdot \delta(x_2)) - \lambda(x_2) \cdot \beta_1 \cdot (1-w) \cdot \frac{m}{b} \cdot x_2 \\ &= (x_2 \cdot \beta_1 + 2) \cdot (w \cdot \delta(0) + 1-w) + 2 \cdot (1-w) \cdot \frac{m}{b} \cdot \lambda(x_2) \end{aligned}$$

gilt. Die Gleichung ist äquivalent zu:

$$\begin{aligned} \frac{-x_2 \cdot \beta_1 - 2}{\exp(\beta_0 + \beta_1 \cdot x_2)} &= \frac{m}{b} \cdot \frac{2 \cdot (1-w)}{w \cdot \delta(0) + 1-w} \\ \Leftrightarrow (-x_2 \cdot \beta_1 - 2) \cdot \exp(-x_2 \cdot \beta_1 - 2) &= \frac{m}{b} \cdot \frac{2 \cdot (1-w)}{w \cdot \delta(0) + 1-w} \cdot \exp(\beta_0 - 2). \end{aligned}$$

Anwenden der Lambertschen W-Funktion und Auflösen nach β_1 liefert die Lösung:

$$\beta_1^* = -\frac{1}{x_2} \cdot \left(W \left(\frac{m}{b} \cdot \frac{2 \cdot (1-w)}{w \cdot \delta(0) + 1-w} \cdot \exp(\beta_0 - 2) \right) + 2 \right).$$

Für $\beta_1 = 0$ gilt $\tilde{g}(w, 0, x_2) = 0$ und für $\beta_1 < 0$ gilt $\tilde{g}(w, \beta_1, x_2) > 0$. Da $x_2 > 0$ ist, folgt mit einer nicht von β_1 abhängenden Funktion $c_2(w, x_2)$ für $\beta_1 \rightarrow -\infty$:

$$\begin{aligned} \lim_{\beta_1 \rightarrow -\infty} \tilde{g}(w, \beta_1, x_2) &= c_2(w, x_2) \cdot \lim_{\beta_1 \rightarrow -\infty} \frac{\exp(\beta_0 + \beta_1 \cdot x_2) \cdot \beta_1^2}{w \cdot \delta(0) + (1-w) \cdot \left(1 + \frac{m}{b} \cdot \exp(\beta_0 + \beta_1 \cdot x_2)\right)} \\ &= c_2(w, x_2) \cdot \lim_{\beta_1 \rightarrow -\infty} \frac{\beta_1^2}{w \cdot \delta(0) \cdot \exp(-\beta_0 - \beta_1 \cdot x_2) + (1-w) \cdot \left(\exp(-\beta_0 - \beta_1 \cdot x_2) + \frac{m}{b}\right)} \\ &= 0. \end{aligned}$$

Somit besitzt $\tilde{g}(w, \beta_1, x_2)$ ein eindeutiges Maximum bei $\beta_1 = \beta_1^*$ und für $\beta_1 < 0$ existieren keine weiteren Extrema. Daher ist $\tilde{g}(w, \beta_1, x_2)$ unimodal bezüglich $\beta_1 \leq 0$ und wird im Intervall $[b_0, b_1]$ auf dem Rand minimal.

Um die minimale Effizienz zu maximieren, wird $\min \{\tilde{g}(w, b_0, x_2), \tilde{g}(w, b_1, x_2)\}$ bezüglich w und x_2 maximiert. Da bei $\tilde{g}(w, \beta_1, x_2)$ die Variablen β_1 und x_2 immer als Produkt $\beta_1 \cdot x_2$ auftreten, besitzt $\tilde{g}(w, b_i, x_2)$, $i = 0, 1$, bezüglich x_2 ein eindeutiges Maximum bei

$$x_{2,b_i} = -\frac{1}{b_i} \cdot \left(W \left(\frac{m}{b} \cdot \frac{2 \cdot (1-w)}{w \cdot \delta(0) + 1 - w} \cdot \exp(\beta_0 - 2) \right) + 2 \right)$$

und ist unimodal bezüglich $x_2 \geq 0$. Da $b_0 < b_1$ ist, gilt $x_{2,b_0} < x_{2,b_1}$. Für $x_2 \leq x_{2,b_0}$ sind die Funktionen $\tilde{g}(w, b_0, x_2)$ und $\tilde{g}(w, b_1, x_2)$ streng monoton wachsend in x_2 und für $x_2 \geq x_{2,b_1}$ sind die Funktionen $\tilde{g}(w, b_0, x_2)$ und $\tilde{g}(w, b_1, x_2)$ streng monoton fallend in x_2 . Somit wird $\min \{\tilde{g}(w, b_0, x_2), \tilde{g}(w, b_1, x_2)\}$ bezüglich x_2 im Intervall $[x_{2,b_0}, x_{2,b_1}]$ maximal. In diesem Intervall ist $\tilde{g}(w, b_0, x_2)$ streng monoton fallend und $\tilde{g}(w, b_1, x_2)$ streng monoton wachsend bezüglich x_2 . Da $b_0 \cdot x_{2,b_0} = b_1 \cdot x_{2,b_1}$ gilt, folgt $\tilde{g}(w, b_0, x_{2,b_0}) = \tilde{g}(w, b_1, x_{2,b_1})$. Somit existiert genau ein $x_{2,w}^* \in [x_{2,b_0}, x_{2,b_1}]$ mit $\tilde{g}(w, b_0, x_{2,w}^*) = \tilde{g}(w, b_1, x_{2,w}^*)$.

Für jedes $w \in (0, 1)$ wird das Minimum von $\tilde{g}(w, b_0, x_2)$ und $\tilde{g}(w, b_1, x_2)$ bei $x_{2,w}^*$ maximal. Insbesondere muss für ein Design mit Stützpunkten 0 und x_2 und zugehörigen Gewichten w und $1 - w$ die Gleichung $\tilde{g}(w, b_0, x_2) = \tilde{g}(w, b_1, x_2)$ gelten, damit es standardisiert Maximin D -optimal sein kann.

Die Gleichung $\tilde{g}(w, b_0, x_2) = \tilde{g}(w, b_1, x_2)$ wird nun nach w aufgelöst. Für $w \neq 0$ und $w \neq 1$ gilt äquivalent:

$$\begin{aligned} \frac{\lambda_{b_0}(x_2) \cdot b_0^2}{w \cdot \delta(0) + (1-w) \cdot \delta_{b_0}(x_2)} &= \frac{\lambda_{b_1}(x_2) \cdot b_1^2}{w \cdot \delta(0) + (1-w) \cdot \delta_{b_1}(x_2)} \\ \Leftrightarrow w \cdot \lambda_{b_0}(x_2) \cdot b_0^2 \cdot (\delta(0) - \delta_{b_1}(x_2)) + \lambda_{b_0}(x_2) \cdot b_0^2 \cdot \delta_{b_1}(x_2) \\ &= w \cdot \lambda_{b_1}(x_2) \cdot b_1^2 \cdot (\delta(0) - \delta_{b_0}(x_2)) + \lambda_{b_1}(x_2) \cdot b_1^2 \cdot \delta_{b_0}(x_2). \end{aligned}$$

Für $\lambda_{b_0}(x_2) \cdot b_0^2 \cdot (\delta(0) - \delta_{b_1}(x_2)) - \lambda_{b_1}(x_2) \cdot b_1^2 \cdot (\delta(0) - \delta_{b_0}(x_2)) \neq 0$ ergibt sich als Lösung:

$$w(x_2) = \frac{\lambda_{b_1}(x_2) \cdot b_1^2 \cdot \delta_{b_0}(x_2) - \lambda_{b_0}(x_2) \cdot b_0^2 \cdot \delta_{b_1}(x_2)}{\lambda_{b_0}(x_2) \cdot b_0^2 \cdot (\delta(0) - \delta_{b_1}(x_2)) - \lambda_{b_1}(x_2) \cdot b_1^2 \cdot (\delta(0) - \delta_{b_0}(x_2))}.$$

Nach Lemma A.7 nimmt $w(x_2)$ Werte im Intervall $(0, 1)$ nur für $\tilde{x} < x_2 < x_0$ an, wobei $\tilde{x} = 2 \cdot \ln(b_1/b_0)/(b_0 - b_1)$ und $x_0 > \tilde{x}$ die eindeutige Nullstelle von $w(x_2)$ im Intervall $(0, \infty)$ ist. Es gilt genau dann $w(x_2) = 0$, wenn $\lambda_{b_1}(x_2) \cdot b_1^2 \cdot \delta_{b_0}(x_2) - \lambda_{b_0}(x_2) \cdot b_0^2 \cdot \delta_{b_1}(x_2) = 0$

ist. Umformen und Division durch $\exp(\beta_0)$ liefert die äquivalente Bedingung:

$$\frac{m}{b} \cdot \exp(\beta_0 + (b_0 + b_1)x) \cdot (b_1^2 - b_0^2) + \exp(b_1x) \cdot b_1^2 - \exp(b_0x) \cdot b_0^2 = 0.$$

Für $x_2 \in (\tilde{x}, x_0)$ ist $w(x_2)$ nach Lemma A.7 streng monoton fallend und besitzt die Darstellung in Gleichung (6.7).

Die Funktion $\tilde{g}(w(x_2), b_0, x_2) = \tilde{g}(w(x_2), b_1, x_2)$ ist bezüglich $x_2 \in (\tilde{x}, x_0)$ zu maximieren, um das standardisiert Maximin D -optimale Design zu erhalten. Umformen ergibt, dass somit die Funktion

$$\begin{aligned} & \frac{w(x_2) \cdot (1 - w(x_2)) \cdot \lambda_{b_0}(x_2) \cdot x_2^2}{w(x_2) \cdot \delta(0) + (1 - w(x_2)) \cdot \delta_{b_0}(x_2)} \\ &= \frac{w(x_2) \cdot (1 - w(x_2)) \cdot \lambda_{b_0}(x_2) \cdot x_2^2}{w(x_2) \cdot \left(1 + \frac{m}{b} \cdot \lambda(0)\right) + (1 - w(x_2)) \cdot \left(1 + \frac{m}{b} \cdot \lambda_{b_0}(x_2)\right)} \\ &= \frac{w(x_2) \cdot (1 - w(x_2)) \cdot \lambda_{b_0}(x_2) \cdot x_2^2}{\frac{m}{b} \cdot w(x_2) \cdot (\lambda(0) - \lambda_{b_0}(x_2)) + 1 + \frac{m}{b} \cdot \lambda_{b_0}(x_2)} \\ &= \frac{w(x_2) \cdot (1 - w(x_2)) \cdot x_2^2}{\frac{m}{b} \cdot w(x_2) \cdot (\exp(-b_0 \cdot x_2) - 1) + \exp(-\beta_0 - b_0 \cdot x_2) + \frac{m}{b}} \end{aligned}$$

zu maximieren ist. Da $w(\tilde{x}) = 1$ und $w(x_0) = 0$ gilt, ist die zu maximierende Funktion für $x_2 = \tilde{x}$ und $x_2 = x_0$ gleich null. Somit existiert auch im offenen Intervall (\tilde{x}, x_0) immer ein Maximum x^* . Gilt $x^* \in \mathcal{X}$, dann ist das Design mit Stützpunkten 0 und x^* sowie zugehörigen Gewichten $w(x^*)$ und $1 - w(x^*)$ standardisiert Maximin D -optimal in der Klasse der 2-Punkt-Designs. \square

Beweis von Satz 6.10:

Es wird der Fall $b_1 < 0$ und $u = 0$ betrachtet. Der Fall $b_0 > 0$ und $v = 0$ folgt analog. Die Funktion $w(x)$ ergibt sich aus Gleichung (A.12) für $b \rightarrow 0$:

$$\begin{aligned} & \lim_{b \rightarrow 0} \frac{1}{1 - \frac{1 + \frac{m}{b} \cdot \exp(\beta_0)}{1 + \frac{m}{b} \cdot \frac{b_0^2 - b_1^2}{\frac{b_0^2}{\exp(\beta_0 + b_1x)} - \frac{b_1^2}{\exp(\beta_0 + b_0x)}}}} \\ &= \frac{1}{1 - \frac{1}{b_0^2 - b_1^2} \cdot \left(\frac{b_0^2}{\exp(b_1x)} - \frac{b_1^2}{\exp(b_0x)} \right)}. \end{aligned}$$

Die Eigenschaften von $w(x)$ ergeben sich analog zum Beweis von Lemma A.7 für $b \rightarrow 0$,

wobei zu beachten ist, dass $w(x)$ keine Nullstelle im Intervall $(0, \infty)$ besitzt. Im Gegensatz zu Gleichung (A.15) gilt für den Grenzwert $\lim_{x \rightarrow \infty} w(x) = 0$. Da $w(x)$ auf dem Intervall $I = (\tilde{x}, \infty)$ streng monoton fallend ist, gilt $0 < w(x) < 1$ für $x \in I$. Das standardisiert Maximin c -optimale Design für $\mathbf{c} = (0, 1)^T$ ergibt sich analog zum Beweis von Satz 6.3. Es ist die Funktion $\tilde{g}(w(x), b_0, x) = \tilde{g}(w(x), b_1, x)$ aus dem Beweis bezüglich $x \in I$ für $b \rightarrow 0$ zu maximieren. Es gilt für den Grenzwert:

$$\begin{aligned} & \lim_{b \rightarrow 0} \frac{w(x) \cdot (1 - w(x)) \cdot \exp(\beta_0 + b_0 x) \cdot x^2}{w(x) \cdot \left(1 + \frac{m}{b} \cdot \exp(\beta_0)\right) + (1 - w(x)) \cdot \left(1 + \frac{m}{b} \cdot \exp(\beta_0 + b_0 x)\right)} \\ & \quad \cdot \frac{\left(\sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0)} + \sqrt{1 + \frac{m}{b} \cdot \exp(\beta_0 - z^*)}\right)^2 \cdot b_0^2}{\exp(\beta_0 - z^*) \cdot (z^*)^2} \\ & = \frac{w(x) \cdot (1 - w(x)) \cdot x^2}{w(x) \cdot (\exp(-b_0 x) - 1) + 1} \cdot \frac{(1 + \sqrt{\exp(-z^*)})^2 \cdot b_0^2}{\exp(-z^*) \cdot (z^*)^2}. \end{aligned}$$

Da der zweite Bruch nicht von x abhängt, kann er bei der Maximierung weggelassen werden. □

Abbildungsverzeichnis

5.1	Grafische Darstellung des D -optimalen Designs für das Poisson-Gamma-Modell für den Fall von drei Kovariablen	47
6.1	D -Effizienz des für $\beta = (0, -1)^T$ lokal D -optimalen Designs ξ_{β}^* bei Variation von β_0 (links) und β_1 (rechts)	53
6.2	Kriteriumsfunktion bei $\beta_1 = b_0$ und $\beta_1 = b_1$ in Abhängigkeit vom Gewicht w für den Stützpunkt $x_1 = 0$	58
6.3	Minimale Effizienz für ein Design mit Stützpunkten 0 und x sowie Gewichten $w(x)$ und $1 - w(x)$, welche bei $b_0 = -3$ und $b_1 = -1$ angenommen wird	62
6.4	Effizienz (links) und Sensitivitätsfunktion (rechts) für das Design ξ^* . . .	63
6.5	Standardisiert Maximin D -optimale Designs in Abhängigkeit von der Intervalllänge des Parameterbereichs $[b_0, b_1]$ mit $b_1 = -1$	65
6.6	Effizienz (links) und Sensitivitätsfunktion $d(\xi^*, x)$ (rechts) für die standardisiert Maximin D -optimalen Designs in Tabelle 6.1 für die Parameterbereiche $[-4, -1]$, $[-6, -1]$ und $[-12, -1]$ (von oben nach unten) . . .	67
6.7	Effizienz (links) und Sensitivitätsfunktion $d(\xi^*, x)$ (rechts) für die standardisiert Maximin D -optimalen Designs in Tabelle 6.1 für die Parameterbereiche $[-18, -1]$, $[-35, -1]$ und $[-41, -1]$ (von oben nach unten) . .	68
6.8	Minimale Effizienz der standardisiert Maximin D -optimalen Designs in Abhängigkeit von der Intervalllänge des Parameterbereichs $[b_0, b_1]$ mit $b_1 = -1$	69
6.9	Effizienzen der standardisiert Maximin D -optimalen Designs in der Klasse der 2-Punkt-Designs (blau) und in der Klasse aller Designs (schwarz) für die Parameterbereiche $[-4, -1]$ (links) und $[-12, -1]$ (rechts)	70
6.10	D -Effizienz (links) und Sensitivitätsfunktion (rechts) für die standardisiert Maximin D -optimalen Designs in Tabelle 6.2 für die Parameterbereiche $[-2, -1]^2$, $[-3.4, -1]^2$, $[-4, -1]^2$ und $[-4, -1] \times [-2, -1]$ (von oben nach unten)	72

A.1 Graph der Funktion $w_{b_0, b_1}(x)$ für $b_0 < b_1 < 0$ 104

Tabellenverzeichnis

5.1	Optimale Designs in Beispiel 5.7 und Vergleich der Effizienzen für das Poisson-Gamma-Modell (P-G) und das Poisson-Modell (Po) im Fall einer Kovariablen	40
5.2	Optimale Designs in Beispiel 5.14 und Vergleich der Effizienzen für das Poisson-Gamma-Modell (P-G) und das Poisson-Modell (Po)	50
6.1	Standardisiert Maximin D -optimale Designs für verschiedene Parameterbereiche $B = [b_0, b_1]$	66
6.2	Standardisiert Maximin D -optimale Designs für verschiedene Parameterbereiche $B = [\mathbf{b}_0, \mathbf{b}_1]$ für multiples Regressionsmodell in Abschnitt 6.1.3	71
6.3	Verteilung π zum Nachweis der Maximin D -Optimalität für die Designs in Tabelle 6.2 mit dem Äquivalenzsatz	71

Symbolverzeichnis

$\mathbf{0}_s$	s -dimensionaler Nullvektor
$\mathbf{0}_{r,s}$	$r \times s$ -Nullmatrix
B	Parameterbereich bei standardisierter Maximin Optimalität
$[b_0, b_1]$	Parameterintervall für den Steigungsparameter bei standardisierter Maximin Optimalität
$\boldsymbol{\beta} = (\beta_0, \dots, \beta_{p-1})^T$	unbekannter Parametervektor
$d(\xi, \boldsymbol{x})$	Sensitivitätsfunktion
$\text{Diag}(z_1, \dots, z_s)$	Diagonalmatrix mit Hauptdiagonaleinträgen z_1, \dots, z_s
$\mathbf{e}_{i,p-1}$	i -te Standard-Einheitsvektor im \mathbb{R}^{p-1}
$\mathbf{e}_{i,p}$	i -te Standard-Einheitsvektor im \mathbb{R}^p
$\text{eff}_c(\xi, \boldsymbol{\beta})$	c -Effizienz des Designs ξ
$\text{eff}_D(\xi, \boldsymbol{\beta})$	D -Effizienz des Designs ξ
$\text{eff}_{D_A}(\xi, \boldsymbol{\beta})$	D_A -Effizienz des Designs ξ
\mathbf{f}	Vektor der Regressionsfunktionen
$F_{\Phi_{\boldsymbol{\beta}}}(\xi_1, \xi_2)$	Fréchet-Ableitung von $\Phi_{\boldsymbol{\beta}}$ an der Stelle ξ_1 in Richtung ξ_2
$f_{\Theta}(\cdot)$	Dichte des zufälligen Effekts Θ
$f_{\mathbf{Y} \Theta=\theta}(\cdot)$	Dichte von \mathbf{Y} gegeben $\Theta = \theta$
$f_{\mathbf{Y}}(\cdot)$	Dichte von \mathbf{Y}
$\Gamma(\cdot)$	Gamma-Funktion
$\gamma(a, b)$	Gamma-Verteilung mit Parametern a und b
$\mathbf{I}(\boldsymbol{\beta})$	Fisher-Informationsmatrix für $\boldsymbol{\beta}$
$\mathbf{I}(\boldsymbol{\beta}, a)$	Fisher-Informationsmatrix für $\boldsymbol{\beta}$ und a
\mathbf{I}_s	$s \times s$ -Einheitsmatrix
$\mathbf{I}_{\text{Po}}(\boldsymbol{\beta})$	Fisher-Informationsmatrix für Poisson-Modell für $\boldsymbol{\beta}$
Λ	Intensitätsmatrix
$\lambda(\cdot)$	Intensitätsfunktion
m	Anzahl der Beobachtungen pro statistischer Einheit
$\mathbf{M}(\xi; \boldsymbol{\beta})$	Informationsmatrix für das individuelle Design ξ und den Parameter $\boldsymbol{\beta}$

$M(\xi; \boldsymbol{\beta}, a)$	Informationsmatrix für das individuelle Design ξ und die Parameter $\boldsymbol{\beta}$ und a
$M(\zeta; \boldsymbol{\beta})$	Informationsmatrix für das Populationsdesign ζ und den Parameter $\boldsymbol{\beta}$
$M(\zeta; \boldsymbol{\beta}, a)$	Informationsmatrix für das Populationsdesign ζ und die Parameter $\boldsymbol{\beta}$ und a
M^{-}	Verallgemeinerte Inverse der Matrix M
$\mathcal{M}_{\boldsymbol{\beta}, A}$	Menge der Informationsmatrizen, für die $A^T \boldsymbol{\beta}$ identifizierbar ist
$\mathcal{M}_{\boldsymbol{\beta}}$	Menge der Informationsmatrizen für gegebenes $\boldsymbol{\beta}$
$M_{\text{Po}}(\xi; \boldsymbol{\beta})$	Informationsmatrix für das Poisson-Modell für das individuelle Design ξ und den Parameter $\boldsymbol{\beta}$
n	Anzahl der statistischen Einheiten
$\mathcal{N}(\xi)$	Menge der $\boldsymbol{\beta}$, für welche die Effizienz des Designs ξ minimal ist
p	Dimension des Parametervektors $\boldsymbol{\beta}$
$\phi(M(\cdot; \boldsymbol{\beta}))$	Kriteriumsfunktion definiert auf $\mathcal{M}_{\boldsymbol{\beta}}$
$\Phi_{\boldsymbol{\beta}}(\cdot)$	Kriteriumsfunktion für gegebenes $\boldsymbol{\beta}$ definiert auf Ξ
$\text{Po}(\lambda)$	Poisson-Verteilung mit Parameter λ
$\psi(\cdot)$	Digamma-Funktion
$\psi_1(\cdot)$	Trigamma-Funktion
q_i	Gewicht für das Design ξ_i
Θ_i	Gamma-verteilter zufälliger Blockeffekt für die i -te statistische Einheit
\mathbf{W}	Gewichtsmatrix
$W(\cdot)$	Hauptzweig der Lambertschen W-Funktion
w_j	Gewicht für die Versuchseinstellung \boldsymbol{x}_j
\mathbf{X}	Designmatrix
\mathcal{X}	Designraum
\boldsymbol{x}_j	Kovariablenvektor, Versuchseinstellung
Ξ	Menge der Designs ξ auf \mathcal{X}
ξ	Individuelles Design
ξ^*	Optimales individuelles Design
Y_{ij}	Zufallsvariable, zufällige Beobachtung
ζ	Populationsdesign
ζ^*	Optimales Populationsdesign

Literaturverzeichnis

- [1] Arbous, A. G. und Kerrich, J. E. (1951). *Accident Statistics and the Concept of Accident-Proneness*. Biometrics, 7, 340–432.
- [2] Atkinson, A. C., Donev, A. N. und Tobias, R. D. (2007). *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford.
- [3] Bates, G. E. und Neyman, J. (1952). *Contributions to the theory of accident proneness: An optimistic model of the correlation between light and severe accidents*. University of California Publications in Statistics, 1, 215–254.
- [4] Braess, D. und Dette, H. (2007). *On the number of support points of maximin and Bayesian optimal designs*. Ann. Statist., 35, 772–792.
- [5] Breslow, N. E. und Clayton, D. G. (1993). *Approximate Inference in Generalized Linear Mixed Models*. J. Amer. Statist. Assoc., 88, 9–25.
- [6] Chernoff, H. (1953). *Locally optimal designs for estimating parameters*. Ann. Math. Statist., 24, 586–602.
- [7] Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J. und Knuth, D. E. (1996). *On The Lambert W Function*. Advances in Computational Mathematics, 5, 329–359.
- [8] Dette, H. (1997). *Designing Experiments with Respect to ‘Standardized’ Optimality Criteria*. J. R. Statist. Soc. B, 59, 97–110.
- [9] Dette, H., Melas, V. B. und Pepelyshev, A. (2006). *Local c- and E-optimal designs for exponential regression models*. Ann. Inst. Stat. Math., 58, 407–426.
- [10] Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.
- [11] Fedorov, V. V. und Hackl, P. (1997). *Model-Oriented Design of Experiments*. Springer, New York.

- [12] Ford, I., Torsney, B. und Wu, C. F. J. (1992). *The Use of a Canonical Form in the Construction of Locally Optimal Designs for Non-Linear Problems*. J. R. Statist. Soc. B, 54, 569–583.
- [13] Graßhoff, U., Holling, H. und Schwabe, R. (2013). *Optimal Design for Count Data with Binary Predictors in Item Response Theory*. In mODa 10 – Advances in Model-Oriented Design and Analysis (Hrsg.: Ucinski, D., Atkinson, A. und Patan, M.), Springer, 117–124.
- [14] Graßhoff, U., Holling, H. und Schwabe, R. (2015). *Poisson Model with Three Binary Predictors: When are Saturated Designs Optimal?*. In Stochastic Models, Statistics and Their Applications (Hrsg.: Steland, A., Rafajłowicz E. und Szajowski K.), Springer Proceedings in Mathematics & Statistics, 75–81.
- [15] Graßhoff, U., Holling, H. und Schwabe, R. (2016). *Optimal Design for the Rasch Poisson-Gamma Model*. In mODa11 - Advances in Model-Oriented Design and Analysis (Hrsg.: Kunert, J., Müller C. H. und Atkinson, A. C.), Springer, 133–141.
- [16] Graßhoff, U., Holling, H. und Schwabe, R. (2018). *D-Optimal Design for the Rasch Counts Model with Multiple Binary Predictors*. arXiv:1810.03893 [stat.ME].
- [17] Harville, D. A. (1997). *Matrix Algebra From a Statisticians's Perspective*. Springer, New York.
- [18] Kiefer, J. und Wolfowitz, J. (1960). *The Equivalence of Two Extremum Problems*. Canadian Journal of Mathematics, 12, 363–366.
- [19] Konstantinou, M., Biedermann, S. und Kimber, A. (2014). *Optimal designs for two-parameter nonlinear models with application to survival models*. Statistica Sinica, 24, 415–428.
- [20] Liski, E. P., Mandal, N. K., Shah, K. R. und Sinha, B. K. (2002). *Topics in Optimal Design*. Springer, New York.
- [21] Longford, N. T. (1994). *Logistic regression with random coefficients*. Computational Statistics & Data Analysis, 17, 1–15.
- [22] McCulloch, C. E. und Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*. Wiley, New York.
- [23] Molenberghs, G., Verbeke, G., Demétrio, C. G. B. und Vieira, A. M. C. (2010). *A Family of Generalized Linear Models for Repeated Measures with Normal and Conjugate Random Effects*. Statistical Science, 25, 325–347.

- [24] Müller, C. H. (1995). *Maximin efficient designs for estimating nonlinear aspects in linear models*. J. Statist. Plann. Inference, 44, 117–132.
- [25] Niaparast, M. (2009). *On optimal design for a Poisson regression model with random intercept*. Statistics & Probability Letters, 79, 741–747.
- [26] Niaparast, M. und Schwabe, R. (2013). *Optimal design for quasi-likelihood estimation in Poisson regression with random coefficients*. J. Statist. Plann. Inference, 143, 296–306.
- [27] Ogungbenro, K. und Aarons, L. (2011). *Population Fisher information matrix and optimal design of discrete data responses in population pharmacodynamic experiments*. J. Pharmacokinet. Pharmacodyn., 38, 449–469.
- [28] Pronzato, L. und Pázman, A. (2013). *Design of Experiments in Nonlinear Models*. Springer, New York.
- [29] Pukelsheim, F. (1993). *Optimal Design of Experiments*. Wiley, New York.
- [30] Pukelsheim, F. und Rieder, B. (1992). *Efficient rounding of approximate designs*. Biometrika, 79, 763–770.
- [31] Pukelsheim, F. und Torsney B. (1991). *Optimal Weights for Experimental Designs on Linearly Independent Support Points*. Ann. Statist., 19, 1614–1625.
- [32] Rodríguez-Torreblanca, C. und Rodríguez-Díaz, J. M. (2007). *Locally D- and c-optimal designs for Poisson and negative binomial regression models*. Metrika, 66, 161–172.
- [33] Rohde, C. A. (1965). *Generalized Inverses of Partitioned Matrices*. Journal of the Society for Industrial and Applied Mathematics, 13, 1033–1035.
- [34] Russell, K. G., Woods, D. C., Lewis, S. M. und Eccleston, E. C. (2009). *D-optimal designs for Poisson regression models*. Statistica Sinica, 19, 721–730.
- [35] Schmelter, T. (2007). *The optimality of single-group designs for certain mixed models*. Metrika, 65, 183–193.
- [36] Schmidt, D. und Schwabe, R. (2017). *Optimal Design for Multiple Regression with Information Driven by the Linear Predictor*. Statistica Sinica, 27, 1371–1384.
- [37] Schmidt, D. (2019). *Characterization of c-, L- and ϕ_k -optimal designs for a class of non-linear multiple-regression models*. J. R. Statist. Soc. B, 81, 101–120.

- [38] Schmidt, M. und Schwabe, R. (2020). *Optimal designs for Poisson count data with Gamma block effects*. J. Statist. Plann. Inference, 204, 128–140.
- [39] Schwabe, R. (1996). *Optimum Designs for Multi-Factor Models*. Springer, New York.
- [40] Silvey, S. D. (1980). *Optimal Design*. Chapman and Hall, London.
- [41] Solis-Trapala, I. L. und Farewell, V. T. (2005). *Regression analysis of overdispersed correlated count data with subject specific covariates*. Statist. Med., 24, 2557–2575.
- [42] Waller, L. A. und Zelterman, D. (1997). *Log-Linear Modeling with the Negative Multinomial Distribution*. Biometrics, 53, 971–982.
- [43] Wang, Y., Myers, R. H., Smith, E. P. und Ye, K. (2006). *D-optimal designs for Poisson regression models*. J. Statist. Plann. Inference, 136, 2831–2845.