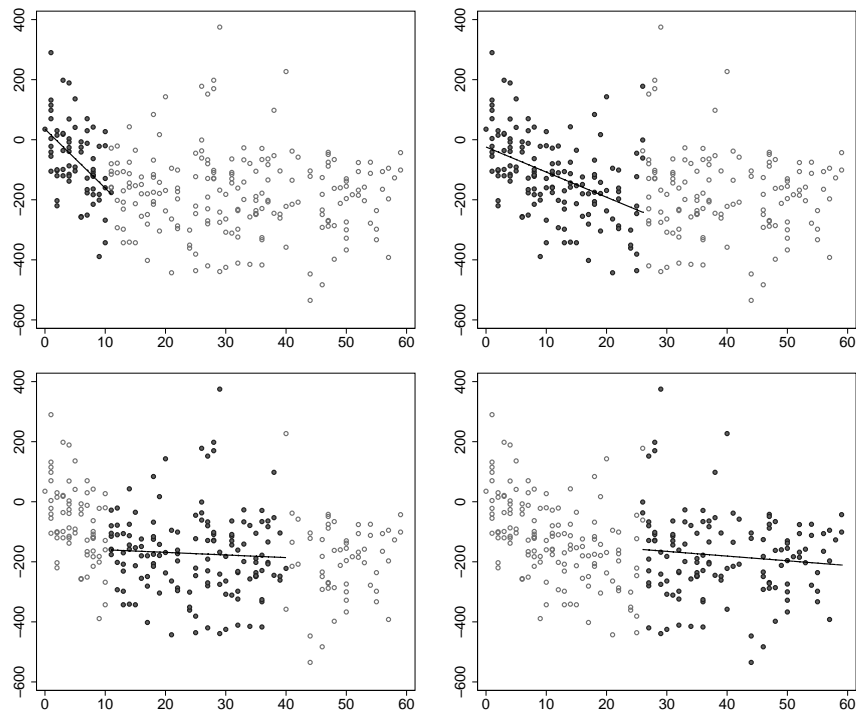


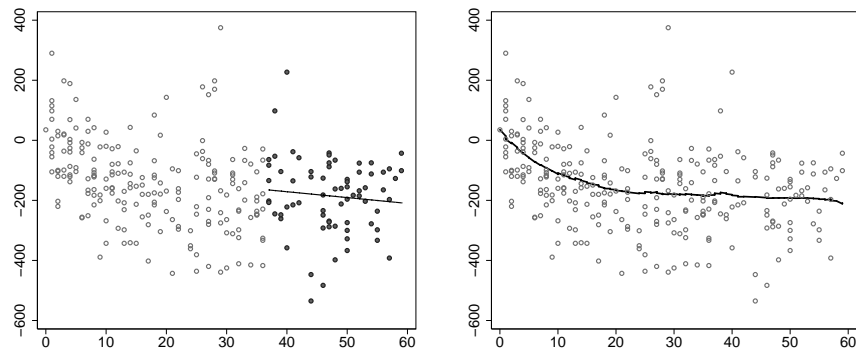
Errata (1. Auflage)

- Kapitel 1, S. 10: Fehlerhafte Variablenbezeichnung in Tabelle 1.5. Um mit späteren Teilen des Buchs und dem auf der Internet-Seite bereitgestellten Datensatz konsistent zu sein, sollte die Variable *art* als *artkat* bezeichnet werden, also

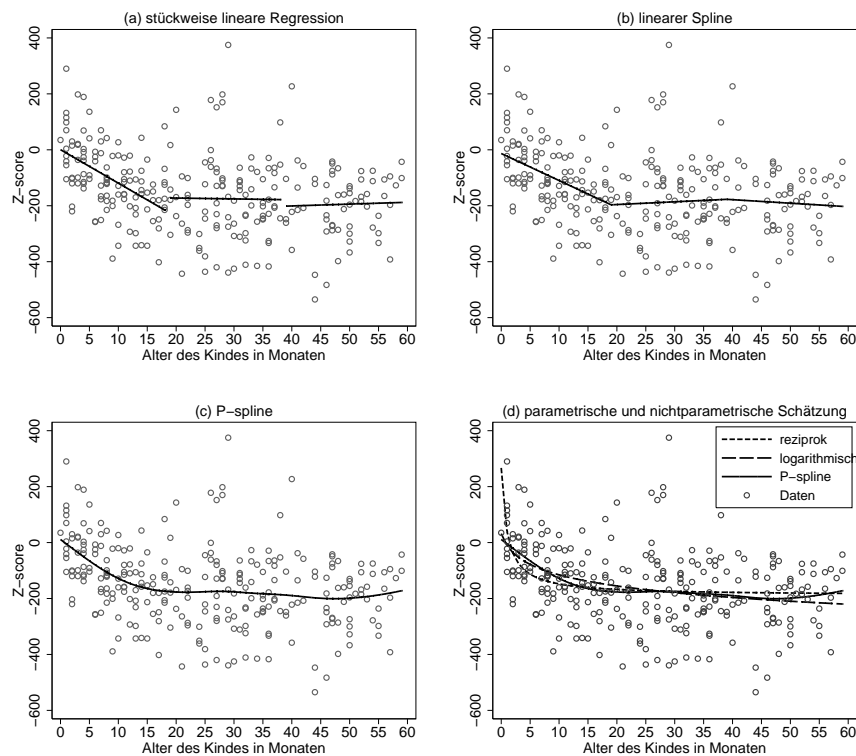
Variable	Beschreibung	Mittelwert/ Häufigkeit in %	Std- abw. Min/Max
<i>artkat</i>	Art des Waldes		
	0 = Mischwald	49.78	
	1 = Laubwald	50.22	

- Kapitel 2, S. 26: Falsche Referenz, richtig wäre:
Dieser beträgt in Beispiel 2.3 111.66 DM und hier 1.75 DM pro Quadratmeter.
- Kapitel 2, S. 32: Fehlende Klammerung bei der Definition der Chance:
... eines multiplikativen Modells für die Chance $\pi_i / (1 - \pi_i)$.
- Kapitel 2, S. 33: Falsche Beschriftung in Abb. 2.8: Die Variable heißt *ansp*, nicht *anspr*.
- Kapitel 2, S. 43: Abb. 2.13 basiert links unten auf einem falschen Datensatz:





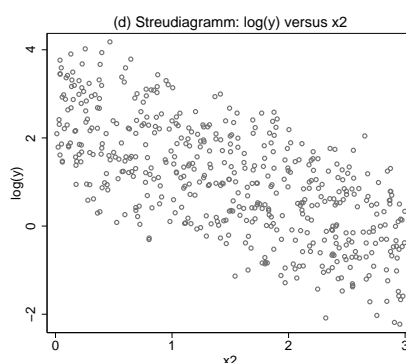
- Kapitel 2, S.44: Abb. 2.14 basiert rechts unten auf einem falschen Datensatz:



- Kapitel 3, S. 61: Ungenaue Formulierung der notwendigen Bedingung für den vollen Spaltenrang der Designmatrix \mathbf{X} :
Eine notwendige Bedingung hierfür ist, dass die Anzahl n der Beobachtungen größer oder gleich als die Zahl p der Regressionskoeffizienten ist.
- Kapitel 3, Seite 71: Falscher Titel in Panel d) von Abbildung 3.6:

Modell	Gleichung	R^2
M1	$\widehat{mieteqm} = 18.46 - 0.068 \cdot flaeche$	0.116
M2	$\widehat{mieteqm} = 9.23 + 273.35 \cdot 1/flaeche$	0.154
M3	$\widehat{mieteqm} = 23.07 - 0.206 \cdot flaeche + 0.00092 \cdot flaeche^2$	0.143
M4	$\widehat{mieteqm} = 27.84 - 0.42 \cdot flaeche + 0.0039 \cdot flaeche^2 - 0.000012 \cdot flaeche^3$	0.150

Tabelle 3.1. Vergleich verschiedener Modelle zur Modellierung des Zusammenhangs zwischen Nettomiete pro qm ($mieteqm$) und Wohnfläche ($flaeche$).



- Kapitel 3, S. 79: Im Text wird ein Regressionsmodell mit $flaecheinvm$ als Kovariable angekündigt, die Modellgleichung enthält dann allerdings die Variable $flaecheinvm$. Tatsächlich berechnet wird das Modell mit $flaecheinvm$, so dass der entsprechende Satz ersetzt werden muss durch:

Mit Hilfe der orthogonalen Variablen $bjahro$, $bjahro2$ und $bjahro3$ und der zentrierten Variable $flaecheinvm$ erhalten wir die Schätzung

$$\widehat{mieteqm}_i = 13.87 + 252.66 \cdot flaecheinvm_i - 1.54 \cdot bjahro_i + 0.97 \cdot bjahro2_i - 0.062 \cdot bjahro3_i.$$

und die in Abbildung 3.11 abgedruckten Effekte $\hat{f}_1(flache)$ und $\hat{f}_2(bjahr)$.

- Kapitel 3, Seite 86: Tippfehler in der Interpretation des Differenzeffekts von Wohnungen in guter Lage im Vergleich zu Wohnungen in normaler Lage:
Der Unterschied wird umso größer, je größer die Wohnfläche ist. Für jeden zusätzlichen Quadratmeter Wohnfläche erhöht sich der Unterschied zu Wohnungen in normaler Wohnlage um 1.6 DM.
- Kapitel 3, Seite 101: Vorzeichenfehler beim Effekt von $1/flache$ in Modell M2:
- Kapitel 3, Seite 105: Falsche Formatierung der 1 in der Definition des Vektors \mathbf{x} :
Diese ist insbesondere dann erfüllt, wenn die beobachteten Regressorvektoren \mathbf{x}_i , $i = 1, \dots, n$ unabhängige und identisch verteilte Realisierungen stochastischer Kovariablen $\mathbf{x} = (1, x_1, \dots, x_k)'$ sind, also wenn (y_i, \mathbf{x}_i) aus einer zufälligen Stichprobe bezüglich (y, \mathbf{x}) stammen.
- Kapitel 3, Seite 109: In der Formel für $\hat{\sigma}_{(i)}^2$ fehlen zusätzliche Klammern und Quadrate:

$$\hat{\sigma}_{(i)}^2 = \frac{1}{n-p-1} \left[\left(y_1 - \mathbf{x}'_1 \hat{\boldsymbol{\beta}}_{(i)} \right)^2 + \dots + \left(y_{i-1} - \mathbf{x}'_{i-1} \hat{\boldsymbol{\beta}}_{(i)} \right)^2 + \right. \\ \left. + \left(y_{i+1} - \mathbf{x}'_{i+1} \hat{\boldsymbol{\beta}}_{(i)} \right)^2 + \dots + \left(y_n - \mathbf{x}'_n \hat{\boldsymbol{\beta}}_{(i)} \right)^2 \right]$$

- Kapitel 3, S.115: Es wird natürlich ein Zusammenhang zum Wald-Test aufgezeigt:
Wir zeigen jetzt noch eine interessante Verbindung zum Wald-Test auf
- Kapitel 3, Seite 119: Verwechslung von Zähler und Nenner:
Da dieser Schätzer konsistent ist, konvergiert der Nenner asymptotisch gegen den wahren, festen Wert, so dass der Quotient zweier χ^2 -Verteilungen gegen die χ_r^2/r -Verteilung des Zählers konvergiert.
- Kapitel 3, Seite 127: Fehlendes \mathbf{P}' in der Definition der Transformation:
Der Vektor der Zielgröße, die Designmatrix und der Vektor der Störgrößen können dann mit der Matrix

$$\mathbf{W}^{-1/2} = \mathbf{P} \operatorname{diag} \left(\frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_n}} \right) \mathbf{P}'$$

transformiert werden.

- Kapitel 3, Seite 131: In Beispiel 3.17 ist die Teststatistik falsch angegeben:
Basierend auf diesem Modell erhalten wir $T = 1164.37$ als Teststatistik für den Breusch-Pagan-Test.
- Kapitel 3, Seite 133: Falsche Definition der Gewichte in der zweistufigen Schätzung:

$$\hat{w}_i = \mathbf{z}'_i \hat{\boldsymbol{\alpha}}.$$

- Kapitel 3, Seite 134: Falsche Definition der Gewichte in der zweistufigen Schätzung:

$$\hat{w}_i = \exp(\mathbf{z}'_i \hat{\boldsymbol{\alpha}}).$$

und

$$\hat{w}_i = \exp(\hat{\eta}_i)$$

Die geschätzten Standardfehler basieren auf der geschätzten Kovarianzmatrix

$$\widehat{\operatorname{Cov}} \hat{\boldsymbol{\beta}} = \hat{\sigma}^2 (\mathbf{X}' \operatorname{diag}(1/\hat{w}_1, \dots, 1/\hat{w}_n) \mathbf{X})^{-1}.$$

- Kapitel 3, S. 142ff: Die mit Hilfe der R-Funktion `dwtest` berechneten (exakten) p-Werte zum Durbin-Watson-Test sind teilweise falsch bzw. zweifelhaft. Neue Berechnungen mit Matlab lieferten:

Beispiel 3.21 Simulierte Daten – Durbin-Watson Test

Für die im letzten Beispiel 3.20 illustrativ verwendeten simulierten Daten erhalten wir als Teststatistik $d = 0.2619$ und mit der Matlab-Funktion `'dwtest'` einen sehr kleinen p-Wert. Damit liefert auch der Durbin-Watson Test zusätzlich zu den vorher betrachteten grafischen Hilfsmitteln einen klaren Hinweis auf Autokorrelation erster Ordnung. \triangle

Beispiel 3.22 Supermarkt Scanner-Daten

In Beispiel 3.2 (Seite 72) haben wir als Illustration für multiplikative Störgrößen Scanner-Daten, die beim Bezahlen an der Supermarktkasse gewonnen werden, untersucht. Ziel ist die Modellierung des Zusammenhangs zwischen dem wöchentlichen Absatz eines Produktes (hier einer Kaffeemarke) und dem eigenen Preis bzw. dem Preis von Konkurrenzprodukten, vergleiche hierzu die Streudiagramme zwischen Absatz und Preis bzw. Konkurrenzpreis in Abbildung 3.6 (Seite 71). In diesem Beispiel analysieren wir die Daten zur Illustration des Vorgehens bei korrelierten Störungen. Der Einfachheit halber benutzen wir lediglich die Daten für ein Geschäft (in Abbildung 3.6 auf Seite 71 sind die Daten für fünf Geschäfte enthalten). Wegen der multiplikativen Störgrößen verwenden wir den logarithmierten Absatz anstelle des untransformierten Absatzes als Zielgröße. Abbildung 3.22 a) zeigt das Streudiagramm zwischen logarithmiertem Absatz und dem eigenen Preis. Zusätzlich eingezeichnet ist der mit Hilfe der gewöhnlichen KQ-Methode geschätzte logarithmierte Absatz wenn der Einfluss des Preises linear modelliert wird. Die dazu gehörenden Residuen sind in Abbildung b) im Zeitablauf dargestellt. Offensichtlich sind die Residuen deutlich korreliert. Die Teststatistik des Durbin-Watson-Tests ist $d = 1.1751$ (p -Wert $< 10^{-4}$). Der berechnete p -Wert basiert dabei auf der Matlab Funktion `dwtest`. Die Nullhypothese $H_0 : \rho = 0$ wird also sowohl zum Niveau $\alpha = 0.01$ als auch zum Niveau $\alpha = 0.05$ verworfen. Zur weiteren Absicherung untersuchen wir in Abbildung 3.23 noch die empirischen Autokorrelationsfunktionen bzw. partiellen Autokorrelationsfunktionen der Residuen. Die Grafiken enthalten zusätzlich punktweise 95% Konfidenzintervalle, auf deren Herleitung wir an dieser Stelle jedoch nicht eingehen. Die Korrelogramme geben deutliche Hinweise auf Autokorrelation. Die Autokorrelationsfunktionen nehmen mit zunehmender Verzögerung nur langsam ab. Die partiellen Autokorrelationsfunktionen sind ab einer Verzögerung von 3 relativ klein. Die Korrelogramme deuten damit nicht unbedingt auf Autokorrelation erster Ordnung hin, eventuell sollte eine komplizierte Korrelationsstruktur (Autokorrelation zweiter Ordnung) zugrunde gelegt werden. Hier zeigt sich ein Vorteil grafischer Hilfsmittel im Vergleich zu formalen Tests. Der Durbin-Watson Test gibt zwar Hinweise auf Autokorrelation, die Art der Korrelation kann aber nicht aufgedeckt werden.

Nachdem wir autokorrelierte Störgrößen mit einiger Sicherheit diagnostiziert haben, liegt folgende Strategie nahe: Zunächst sollte untersucht werden, ob durch verbesserte Modellspezifikation die Korrelationen verschwinden. Erst als letzter Ausweg sollte ein Modell mit autokorrelierten Fehlern geschätzt werden. Im vorliegenden Fall liefert ein Blick auf die Preise im Zeitablauf (Abbildung 3.22 c) den entscheidenden Hinweis, wie die Modellspezifikation verbessert werden kann. Offenbar lassen sich drei Perioden unterscheiden, in denen der „übliche“ Verkaufspreis der Kaffeemarke (in der Marketingliteratur spricht man von regulären Preisen) jeweils unterschiedlich war. Zunächst war der reguläre Preis bei 8.5 DM, anschließend wurde der reguläre Preis dauerhaft auf 8 DM und schließlich sogar auf 7.5 DM gesenkt. In jeder der Perioden wurden regelmäßige Preisaktionen mit erniedrigten Preisen durchgeführt. Es ist relativ klar, dass eine Nichtbeachtung der zeitlichen Entwicklung des regulären Preises problematisch ist. Ein Vorschlag in der Marketingliteratur, vergleiche z.B. Leeflank, Wittink, Wedel & Naert (2000), ist, den Preis als erklärende Variable im Modell zu ersetzen durch den Quotienten aus dem aktuellen und momentan regulären Preis. Verwende also die neue Variable

$$preisq = \frac{\text{aktueller Preis}}{\text{regulärer Preis}}.$$

Dieser Quotient liegt zwischen Null und Eins. Falls $preisq = 1$, ist der aktuelle Preis gleich dem regulären Preis, es gab also in dieser Woche keine Preissenkung. Abbildung 3.24 a) zeigt das Streudiagramm zwischen Log-Absatz und dem Quotienten aus dem aktuellen und regulären Preis. Zusätzlich eingezeichnet ist wieder der geschätzte Log-Absatz bei

linearer Modellierung. Die Residuen im Zeitablauf in Abbildung b) sind bereits deutlich weniger korreliert als bei den Ursprungsregressionen mit dem aktuellen Preis als Einflussgröße (siehe nochmal die Abbildung 3.22 b). Diese Beobachtung wird auch durch den Durbin-Watson Test gestützt. Wir erhalten $d = 1.5946$ mit p -Wert = 0.0348. Die Nullhypothese wird also nur noch bei einem Niveau $\alpha = 0.05$ abgelehnt, auf dem Niveau $\alpha = 0.01$ wird die Nullhypothese beibehalten. Eine weitere Verringerung der Autokorrelation wird erreicht durch Hinzunahme weiterer Kovariablen, etwa der Konkurrenzpreise, oder durch nichtlineare Modellierung. \triangle

- Kapitel 3, S. 148: Fehler in Formel (3.38). Auf der linken Seite der Gleichung sollte σ^2 stehen statt σ , also

$$p(\boldsymbol{\beta}, \sigma^2) \propto \frac{1}{(\sigma^2)^{\frac{p}{2}}} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{\beta} - \mathbf{m})' \mathbf{M}^{-1}(\boldsymbol{\beta} - \mathbf{m})\right) \frac{1}{(\sigma^2)^{a+1}} \exp\left(-\frac{b}{\sigma^2}\right). \quad (3.38)$$

- Kapitel 3, S. 149: $P(\boldsymbol{\beta}, \sigma^2)$ statt $p(\boldsymbol{\beta}, \sigma^2)$:

$$\begin{aligned} p(\sigma^2 | \boldsymbol{\beta}) &\propto p(\boldsymbol{\beta}, \sigma^2) \\ &\propto \frac{1}{(\sigma^2)^{a+\frac{p}{2}+1}} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{\beta} - \mathbf{m})' \mathbf{M}^{-1}(\boldsymbol{\beta} - \mathbf{m})\right) \exp\left(-\frac{b}{\sigma^2}\right). \end{aligned}$$

- Kapitel 3, S. 151: Fehler in der Formel für die Posteriori-Kovarianzmatrix. Seite 150 ist das korrekte Ergebnis zu entnehmen:

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = \left(\frac{1}{\sigma^2} \mathbf{X}' \mathbf{X} + \frac{1}{\sigma^2} \mathbf{M}^{-1}\right)^{-1} \quad \boldsymbol{\mu}_{\boldsymbol{\beta}} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \left(\frac{1}{\sigma^2} \mathbf{X}' \mathbf{y} + \frac{1}{\sigma^2} \mathbf{M}^{-1} \mathbf{m}\right).$$

- Kapitel 3, S. 151:
 - $\hat{\boldsymbol{\beta}}_1$ ist verzerrt, außer für $\mathbf{X}'_1 \mathbf{X}_2 = \mathbf{0}$. Die Bedingung $\mathbf{X}'_1 \mathbf{X}_2 = \mathbf{0}$ bedeutet, dass jede Variable in \mathbf{X}_1 unkorreliert zu jeder Variable in \mathbf{X}_2 ist.
 - Es lässt sich zeigen, dass Schätzungen für die Komponenten in $\hat{\boldsymbol{\beta}}_1$ basierend auf $\hat{\boldsymbol{\beta}}$ variabler sind als Schätzungen basierend auf $\tilde{\boldsymbol{\beta}}_1$
- Kapitel 3, S. 152: Fehlender Fettdruck und fehlerhafte Bezeichnung:
Die Parameter $\tilde{\mathbf{m}}$, $\tilde{\mathbf{M}}$, \tilde{a} und \tilde{b} der Verteilung findet man im nachfolgenden Kasten...
- Kapitel 3, Seite 161: Falsche Verwendung des Betrags in der Definition des AIC:

$$\text{AIC} = -2 \cdot l(\hat{\boldsymbol{\beta}}_M, \hat{\sigma}^2) + 2(|M| + 1)$$

$$\text{AIC} = n \cdot \log(\hat{\sigma}^2) + 2(|M| + 1).$$

- Kapitel 3, Seite 161: Im Text fehlen die Betragsstriche um M :
Man beachte, dass die Gesamtzahl an Parametern $|M| + 1$ ist.
- Kapitel 3, Seite 162: Fehlender Index M in der Definition der Prädiktionsmatrix:
Es gilt

$$\text{CV} = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_{iM}}{1 - h_{iiM}} \right)^2,$$

wobei h_{iiM} die Diagonalelemente der Prädiktionsmatrix $\mathbf{H}_M = \mathbf{X}_M(\mathbf{X}'_M \mathbf{X}_M)^{-1} \mathbf{X}'_M$ sind.

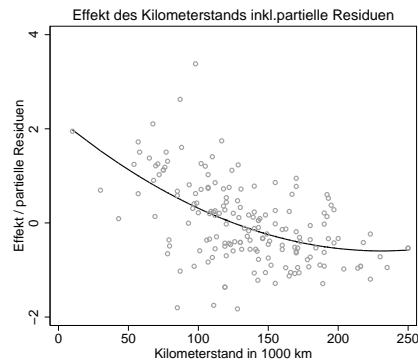
- Kapitel 3, Seite 162: Falsche Parameteranzahl in der Definition des BIC:

$$\text{BIC} = -2 \cdot l(\hat{\beta}_M, \hat{\sigma}^2) + \log(n) (|M| + 1)$$

und

$$\text{BIC} = n \cdot \log(\hat{\sigma}^2) + \log(n) (|M| + 1).$$

- Kapitel 3, Seite 167: Fehlerhafte Achsenbeschriftung der rechten Grafik in Abbildung 3.30:



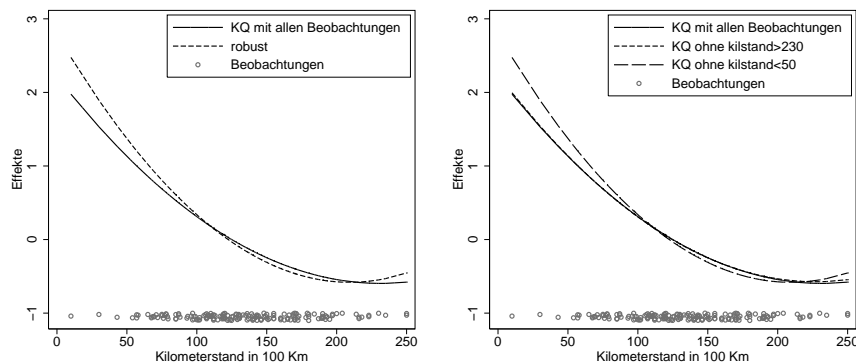
- Kapitel 3, Seite 174: Verwechslung der gestrichelten und der durchgezogenen Linie in der Unterschrift zu Abbildung 3.33:

Die jeweils linken Grafiken zeigen Streudiagramme zwischen y und x inklusive geschätzter Regressionsgerade (durchgezogene Linie) und geschätzter Regressionsgerade bei nicht Berücksichtigung der Ausreißer (gestrichelte Linie).

- Kapitel 3, Seite 178: Fehlendes Quadrat im Zähler bei der Formel für die Hebelwerte:

$$h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_j (x_j - \bar{x})^2}.$$

- Kapitel 3, Seite 180: Fehlerhafte Achsenbeschriftung in Abbildung 3.36:



- Kapitel 3, Seite 186: Der Nullvektor sollte fett formatiert sein:
In der vorletzten Zeile haben wir ausgenutzt, dass die Residuen des vollen Modells und die Spalten der Designmatrix orthogonal sind, d.h. $\hat{\epsilon}'\mathbf{X} = \mathbf{0}$.
- Kapitel 3, S. 186: Missverständliche Formulierung. Es sollte heißen
Wir wenden Satz B.1.6 auf $\mathbf{Z} = \mathbf{C}\hat{\beta} - \mathbf{d}$ und $\mathbf{A} = (\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}$ an und erhalten...
- Kapitel 3, S. 188: Fehler in Zeile 3 der Herleitung der Posteriori-Verteilung (\mathbf{V}^{-1} statt \mathbf{M}^{-1}):

$$\mathbf{y}'\mathbf{y} + \beta'(\mathbf{X}'\mathbf{X} + \mathbf{M}^{-1})\beta - 2\beta'(\mathbf{X}'\mathbf{y} + \mathbf{M}^{-1}\mathbf{m}) + \mathbf{m}'\mathbf{M}^{-1}\mathbf{m}$$

- Kapitel 4, S. 194: Für die Interpretation der Koeffizienten des Logit-Modells müssen Zähler und Nenner des Chancen-Verhältnisses vertauscht werden:
Wird z.B. x_{i1} um 1 auf $x_{i1} + 1$ erhöht, so gilt für das Verhältnis der Chancen

$$\frac{P(y_i = 1 | x_{i1} + 1, \dots)}{P(y_i = 0 | x_{i1} + 1, \dots)} / \frac{P(y_i = 1 | x_{i1}, \dots)}{P(y_i = 0 | x_{i1}, \dots)} = \exp(\beta_1).$$

- Kapitel 4, S. 200: Bei der Bestimmung der beobachteten Fisher-Information fehlen partielle Ableitungs-Zeichen. Es müsste eigentlich heißen:
Die beobachtete Informationsmatrix $\mathbf{H}(\beta) = -\partial^2 l(\beta) / \partial \beta \partial \beta' = -\partial \mathbf{s}(\beta) / \partial \beta'$ erhält man durch nochmaliges Differenzieren von

$$-\mathbf{s}(\beta) = \sum_{i=1}^n \mathbf{x}_i (\pi_i(\beta) - y_i).$$

Mit Hilfe der Kettenregel erhält man nach kurzer Rechnung

$$\mathbf{H}(\beta) = -\frac{\partial \mathbf{s}(\beta)}{\partial \beta'} = \sum_{i=1}^n \mathbf{x}_i \frac{\partial \pi_i(\beta)}{\partial \beta'} = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i' \frac{\partial \pi_i(\beta)}{\partial \eta_i} = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i' \pi_i(\beta) (1 - \pi_i(\beta)).$$

- Kapitel 4, S. 194: Fehlender Bruchstrich in der Gleichung für erwartete und beobachtete Fisher-Information:

$$\mathbf{H}(\beta) = -\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'} = \mathbb{E} \left(-\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'} \right) = \mathbf{F}(\beta).$$

- Kapitel 4, S. 194: Falscher Laufindex in der Verteilung der relativen Häufigkeiten:

$$\bar{y}_i \sim \text{B}(n_i, \pi_i) / n_i \quad i = 1, \dots, G,$$

- Kapitel 4, S. 202: Falsches Transponiertzeichen in der Formel für die Informationsmatrix:

$$\mathbf{F}(\beta) = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i' / \sigma^2 = \frac{1}{\sigma^2} \mathbf{X}'\mathbf{X}.$$

- Kapitel 4, Beispiel 4.5, Seiten 208/09: Unvollständige Erklärung der Variablen *hoehe* (Tabelle 4.6) und falsche Standardabweichung für den Effekte von *zweck* (Tabelle 4.7):

Variable	Beschreibung	Mittelwert/ rel. Häufigkeit
<i>konto1</i>	1 = kein Konto 0 = gutes oder schwaches Konto	0.274
<i>konto2</i>	1 = gutes Konto 0 = kein oder schwaches Konto	0.394
<i>laufz</i>	Laufzeit in Monaten	20.903
<i>hoehe</i>	Kredithöhe in 1000 DM,	3.271
<i>moral</i>	frühere Zahlungsmoral 1 = gut 0 = schlecht	0.911
<i>zweck</i>	Verwendungszweck 1 = privat 0 = geschäftlich	0.657

Tabelle 4.6. Kredit-Scoring: Beschreibung der Variablen.

Variable	Koeffizient	Standard- abweichung	t-Wert	p-Wert
Konstante	-0.284	0.303	-0.94	0.347
<i>konto1</i>	0.618	0.175	3.53	<0.001
<i>konto2</i>	-1.338	0.201	-6.65	<0.001
<i>laufz</i>	0.033	0.008	4.29	<0.001
<i>hoehe</i>	0.023	0.033	0.72	0.474
<i>moral</i>	-0.986	0.251	-3.93	<0.001
<i>zweck</i>	-0.426	0.158	-2.69	0.007

Tabelle 4.7. Kredit-Scoring: Ergebnisse des Logit-Modells.

Variable	Koeffizient	Standard- abweichung	t-Wert	p-Wert
Konstante	0.165	0.032	5.21	<0.001
<i>jahr</i>	-0.075	0.003	-25.86	<0.001
<i>alandc</i>	-0.025	0.004	-6.12	<0.001
<i>anspc</i>	0.020	0.001	17.12	<0.001
<i>biopharm</i>	0.294	0.031	9.41	<0.001
<i>uszw</i>	-0.043	0.025	-1.77	0.077
<i>patus</i>	0.018	0.026	0.69	0.491
<i>patdsg</i>	-0.230	0.031	-7.31	<0.001
<i>einspruch</i>	0.404	0.024	16.55	<0.001

Tabelle 4.9. Anzahl der Zitate von Patenten: Modell mit linearen Effekten, $\phi = 1$.

- Kapitel 4, Seiten 212: Fehler in der Normalisierungskonstante der Likelihood:
Somit ist die *Log-Likelihood* bis auf die additive, von β unabhängige Konstante
– $\sum_i \log(y_i!)$ durch ...

Variable	Koeffizient	Standard- abweichung	t-Wert	p-Wert
Konstante	0.165	0.060	2.74	0.006
<i>jahrc</i>	-0.075	0.005	-13.60	<0.001
<i>alandc</i>	-0.025	0.008	-3.22	0.001
<i>anspc</i>	0.020	0.002	9.01	<0.001
<i>biopharm</i>	0.294	0.059	4.95	<0.001
<i>uszw</i>	-0.043	0.047	-0.93	0.353
<i>patus</i>	0.018	0.050	0.36	0.717
<i>patdsg</i>	-0.230	0.060	-3.84	<0.001
<i>einspruch</i>	0.404	0.046	8.70	<0.001

Tabelle 4.10. Anzahl der Zitate von Patenten: Modell mit linearen Effekten und Überdispersion.

Variable	Koeffizient	Standard- abweichung	t-Wert	p-Wert
Konstante	0.181	0.062	2.91	0.004
<i>jahrc</i>	39.081	5.750	6.80	<0.001
<i>jahrc</i> ²	-0.010	0.001	-6.81	<0.001
<i>alandc</i>	-0.244	0.098	-2.49	0.013
<i>alandc</i> ²	0.037	0.013	2.92	0.004
<i>alandc</i> ³	-0.002	<0.001	-3.42	0.001
<i>anspc</i>	0.040	0.007	5.59	<0.001
<i>anspc</i> ²	-0.001	<0.001	-2.91	0.004
<i>biopharm</i>	0.200	0.062	3.22	0.001
<i>uszw</i>	-0.048	0.048	-1.00	0.320
<i>patus</i>	-0.003	0.052	-0.06	0.952
<i>patdsg</i>	-0.234	0.062	-3.77	<0.001
<i>einspruch</i>	0.401	0.047	8.45	<0.001

Tabelle 4.11. Anzahl der Zitate von Patenten: Erweitertes Modell mit Überdispersion.

- Kapitel 4, Beispiel 4.7, Seiten 214/15: Durch fehlerhaftes Zentrieren sind die Konstanten in den Tabellen 4.9, 4.10 und 4.11 falsch berechnet worden:
- Kapitel 4, S. 214: Fehlende Fett-Formatierung für Vektoren und Matrizen:
Die geschätzte Kovarianzmatrix ist dann mit $\hat{\phi}$ zu multiplizieren, d.h. es gilt $\widehat{\text{Cov}}(\hat{\beta}) = \hat{\phi} \mathbf{F}^{-1}(\hat{\beta})$.
- Kapitel 4, S. 217: Fehlendes Vorzeichen in der reziproken Verknüpfung:

$$\mu_i = -\frac{1}{\eta_i} = -\frac{1}{\mathbf{x}_i' \boldsymbol{\beta}}.$$

- Kapitel 4, S. 219, Tabelle 4.12: Fehler im natürlichen Parameter der Bernoulli-Verteilung:

$$\theta(\pi) = \log(\pi/(1-\pi)) = \log\left(\frac{\pi}{1-\pi}\right).$$

- Kapitel 4, S. 219, Tabelle 4.12: Fehler im natürlichen Parameter der inversen Gauß-Verteilung:

$$\theta(\mu) = -1/(2\mu^2).$$

- Kapitel 4, S. 219, Tabelle 4.12: Fehler im Erwartungswert der Gammaverteilung:

$$\mu = -1/\theta.$$

- Kapitel 4, S. 220: θ anstelle von ϕ in der Darstellung der Normalverteilung als Exponentialfamilie:
... in die Form einer Exponentialfamilie mit $\theta = \mu$, $\phi = \sigma^2$, $b(\theta) = \mu^2/2 = \theta^2/2$ und $c(y, \phi) = -y^2/(2\sigma^2) - 0.5 \log(2\pi\sigma^2)$ bringen.
- Kapitel 4, S. 224: Fehler im Index der Varianz-Schätzung:
Das Diagonalelement a_{jj} ist ein Schätzer für die Varianz $\sigma_j^2 = \text{Var}(\hat{\beta}_j)$ der j -ten Komponente und $\sqrt{a_{jj}}$ für die Standardabweichung σ_j .
- Kapitel 4, S. 225: Falsche Freiheitsgrade in der approximativen χ^2 -Verteilung der Teststatistiken.

Für großes n gilt unter H_0 approximativ:

$$lq, w, u \stackrel{a}{\sim} \chi_r^2.$$

- Kapitel 4, S. 225: Im Fließtext wird die Score-Statistik einmal fälschlicherweise als s bezeichnet:
Unter ähnlichen Voraussetzungen wie für die ML-Schätzung gilt $lq, w, u \stackrel{a}{\sim} \chi_r^2$.
- Kapitel 4, S. 225: Fehlerhafte Freiheitsgrade in der asymptotischen χ^2 -Verteilung:
Falls n_i , $i = 1, \dots, G$, in allen Gruppen hinreichend groß ist, so sind beide Statistiken approximativ (bzw. für $n \rightarrow \infty$ asymptotisch) $\phi \chi^2(G - p)$ -verteilt, wobei p die Anzahl zu schätzender Parameter ist.
- Kapitel 4, S. 229: $p(\beta | \mathbf{y})$ statt $E(\beta | \mathbf{y})$:
Für $n \rightarrow \infty$ ist $\hat{\beta}_p$ asymptotisch (bzw. approximativ) normalverteilt mit

$$\hat{\beta}_p \sim N\left(\beta, \mathbf{F}_p^{-1}(\hat{\beta}_p)\right),$$

so dass der Posteriori-Modus $\hat{\beta}_p$ und die (erwartete) Krümmung $\mathbf{F}_p^{-1}(\hat{\beta}_p)$ gute Approximationen des Posteriori-Erwartungswerts $E(\beta | \mathbf{y})$ und der Posteriori-Kovarianzmatrix $\text{Cov}(\beta | \mathbf{y})$ sind.

- Kapitel 4, S. 230: Fehlende Fett-Formatierung des Nullvektors und der Nullmatrix:
Für den Grenzfall $\mathbf{m} = \mathbf{0}$, $\mathbf{M}^{-1} = \mathbf{0}$ einer nichtinformativen Priori...
- Kapitel 4, S. 232: Fehlende Fett-Formatierung von Vektoren:
Ein Vorteil ist, dass neben den Punktschätzern und Konfidenzintervallen mit Hilfe der gezogenen Zufallszahlen auch die gesamte Posteriori-Dichte $p(\beta | \mathbf{y})$ interessierender Effekte geschätzt werden kann. Abbildung ?? zeigt die Kerndichteschätzungen der Posteriori-Dichten $p(\beta_j | \mathbf{y})$ der Effekte der Kovariablen...
- Kapitel 4, S. 233: Fehlerhafte Beschriftung der Spalten in Tabelle 4.15:
- Kapitel 5, S. 237: Fehler m in der Wahrscheinlichkeitsfunktion bei unabhängigen Wiederholungen:

$$f(\mathbf{y} | \boldsymbol{\pi}) = \frac{m!}{y_1! \cdot \dots \cdot y_q! (m - y_1 - \dots - y_q)!} \pi_1^{y_1} \cdot \dots \cdot \pi_q^{y_q} (1 - \pi_1 - \dots - \pi_q)^{m - y_1 - \dots - y_q}$$

Variable	Koeffizient	Standard- abweichung	2.5%- Quantil	97.5%- Quantil
Konstante	-0.282	0.299	-0.871	0.299
<i>konto1</i>	0.633	0.176	0.301	0.981
<i>konto2</i>	-1.332	0.198	-1.709	-0.932
<i>laufz</i>	0.034	0.008	0.018	0.048
<i>hoehe</i>	0.024	0.033	-0.039	0.090
<i>moral</i>	-1.005	0.257	-1.503	-0.503
<i>zweck</i>	-0.433	0.153	-0.735	-0.144

Tabelle 4.15. Kredit-Scoring: Bayes-Modell.

- Kapitel 5, S. 237: Fehlendes m in der Kovarianzmatrix:

$$E(\mathbf{y}) = m\boldsymbol{\pi} = \begin{pmatrix} m\pi_1 \\ \vdots \\ m\pi_q \end{pmatrix}, \quad \text{Cov}(\mathbf{y}) = \begin{pmatrix} m\pi_1(1-\pi_1) & \cdots & -m\pi_1\pi_q \\ \vdots & \ddots & \vdots \\ -m\pi_q\pi_1 & \cdots & m\pi_q(1-\pi_q) \end{pmatrix}. \quad (5.4)$$

- Kapitel 5, S. 239: Die Berücksichtigung der Referenzkategorie in der äquivalenten Darstellung des multinomialen Logit-Modells ist nicht explizit ausgeführt:

Eine äquivalente Darstellung ist

$$\log \frac{\pi_{ir}}{\pi_{ic}} = \mathbf{x}'_i \boldsymbol{\beta}_r \quad \text{bzw.} \quad \frac{\pi_{ir}}{\pi_{ic}} = \exp(\mathbf{x}'_i \boldsymbol{\beta}_r), \quad r = 1, \dots, c, \quad (5.6)$$

wobei aus $\pi_{ic}/\pi_{ic} = 1$ automatisch $\boldsymbol{\beta}_c = \mathbf{0}$ folgt.

- Kapitel 5, S. 239: Index nicht richtig formatiert:
Besitzt beispielsweise Kategorie s einen größeren Regressionskoeffizienten als Kategorie r , das heißt gilt $\beta_{sj} > \beta_{rj}$, so wächst die Chance für Kategorie s im Vergleich zur Referenzkategorie stärker als die Chance für Kategorie r , so dass sich insgesamt die Wahrscheinlichkeit für Kategorie r verkleinern kann, trotz positivem Koeffizienten β_{rj} .
- Kapitel 5, S. 245: Die Bedingung ist das Auftreten von r oder einer größeren Kategorie:
Auf der linken Seite findet sich hier die Wahrscheinlichkeit für die Zielvariable in Kategorie r , gegeben es tritt die Kategorie r oder eine größere Kategorie auf.
- Kapitel 5, S. 247: Fehlende Fett-Formatierung der Score-Funktion:
Daraus erhält man die Score-Funktion $\mathbf{s}(\boldsymbol{\beta}) = \partial l(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}$ als Vektor der 1. Ableitungen
...
- Kapitel 5, S. 248: Fehlendes Transponiertzeichen beim Prädiktorvektor:

$$\boldsymbol{\eta}_i = (\eta_{i1}, \dots, \eta_{iq})'$$

- Kapitel 5, S. 248: Fehlender Index bei der Fisher-Matrix:

$$\mathbf{F}(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{X}'_i \mathbf{W}_i \mathbf{X}_i$$

- Kapitel 5, S. 248: Fehlende Definition von $\boldsymbol{\mu}$:

wobei $\boldsymbol{\mu}' = (n_1 \boldsymbol{\pi}'_1, \dots, n_n \boldsymbol{\pi}'_n)$ gilt

- Kapitel 6, S. 258: Die Dimension der Kovarianzmatrix \mathbf{D} sollte $(q+1) \times (q+1)$ sein.
- Kapitel 6, S. 260: In Formel (6.13) sind $\mathbf{D}_1, \dots, \mathbf{D}_i, \dots, \mathbf{D}_m$ nicht definiert:

$$\mathbf{R} = \text{diag}(\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_i, \dots, \boldsymbol{\Sigma}_m), \quad \mathbf{G} = \text{diag}(\mathbf{D}_1, \dots, \mathbf{D}_i, \dots, \mathbf{D}_m), \quad (6.13)$$

wobei $\mathbf{D}_i \equiv \mathbf{D}$.

- Kapitel 6, S. 264: In Formel (6.30) muss es korrekt

$$\text{Cov} \begin{pmatrix} \tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \end{pmatrix} = (\mathbf{C}'\mathbf{R}^{-1}\mathbf{C} + \mathbf{B})^{-1}$$

heißen. Für $\hat{\boldsymbol{\beta}}$ bleibt die Änderung ohne Folge, da $\boldsymbol{\beta}$ ein fester Parameter ist. Im Gegensatz dazu ist $\boldsymbol{\gamma}$ zufällig, so dass $\boldsymbol{\gamma}$ in der Berechnung der Kovarianzmatrix eine Rolle spielt. Tatsächlich gilt

$$\text{Cov}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\gamma}}) = 0.$$

Ähnlich müssen Formel (6.31) sowie die entsprechenden Formeln auf den Seiten 273 und 286 geändert werden.

- Kapitel 6, S. 267: Fehlerhafte Transponiertzeichen in der Formel für die Wald-Statistik:

$$\mathbf{W} = (\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{d})'(\mathbf{C}\mathbf{A}_{11}\mathbf{C}')^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{d})$$

- Kapitel 6, S. 269: $\boldsymbol{\beta}$ statt $\hat{\boldsymbol{\beta}}$ in der Formel für die restringierte Log-Likelihood:

$$l_R(\boldsymbol{\vartheta}) = -\frac{1}{2} \sum_{i=1}^m \left\{ \log |\mathbf{V}_i(\boldsymbol{\vartheta})| + \log |\mathbf{X}_i' \mathbf{V}_i(\boldsymbol{\vartheta})^{-1} \mathbf{X}_i| \right. \\ \left. + (\mathbf{y}_i - \mathbf{X}_i' \hat{\boldsymbol{\beta}})' \mathbf{V}_i(\boldsymbol{\vartheta})^{-1} (\mathbf{y}_i - \mathbf{X}_i' \hat{\boldsymbol{\beta}}) \right\}.$$

- Kapitel 6, S. 274: In den Formeln (6.37) und (6.38) fehlen teilweise die Varianzparameter $\boldsymbol{\vartheta}$ in der Bedingung:

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta} | \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta}) p(\boldsymbol{\beta}) p(\boldsymbol{\gamma} | \boldsymbol{\vartheta}) p(\boldsymbol{\vartheta}),$$

$$p(\mathbf{y}) = \int p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta}) p(\boldsymbol{\beta}) p(\boldsymbol{\gamma} | \boldsymbol{\vartheta}) p(\boldsymbol{\vartheta}) d\boldsymbol{\beta} d\boldsymbol{\gamma} d\boldsymbol{\vartheta} < \infty$$

- Kapitel 6, S. 275: In der vollständig bedingten Dichte für $\boldsymbol{\vartheta}$ fehlen die Determinanten der Kovarianzmatrizen, es müsste also heißen

$$p(\boldsymbol{\vartheta} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta}) p(\boldsymbol{\gamma}) p(\boldsymbol{\vartheta}) \propto |\mathbf{R}(\boldsymbol{\vartheta})|^{-1/2} |\mathbf{G}(\boldsymbol{\vartheta})|^{-1/2} \\ \cdot \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{U}\boldsymbol{\gamma})' \mathbf{R}(\boldsymbol{\vartheta})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{U}\boldsymbol{\gamma}) - \frac{1}{2} \boldsymbol{\gamma}' \mathbf{G}(\boldsymbol{\vartheta})^{-1} \boldsymbol{\gamma} \right\} p(\boldsymbol{\vartheta}).$$

- Kapitel 6, S. 275: Mehrmals müsste es $\mathbf{G}(\boldsymbol{\vartheta})$ statt \mathbf{G} heißen und zwar in der vollständig bedingten Dichte für $\boldsymbol{\gamma}$

$$p(\boldsymbol{\gamma} | \boldsymbol{\beta}, \boldsymbol{\vartheta}, \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta}) p(\boldsymbol{\gamma}) \\ \propto \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{U}\boldsymbol{\gamma})' \mathbf{R}(\boldsymbol{\vartheta})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{U}\boldsymbol{\gamma}) - \frac{1}{2} \boldsymbol{\gamma}' \mathbf{G}(\boldsymbol{\vartheta})^{-1} \boldsymbol{\gamma} \right\},$$

sowie in den Momenten der resultierenden Normalverteilung

$$\begin{aligned}\boldsymbol{\mu}_\gamma &= (\mathbf{U}'\mathbf{R}(\boldsymbol{\vartheta})^{-1}\mathbf{U} + \mathbf{G}(\boldsymbol{\vartheta})^{-1})^{-1}(\mathbf{U}'\mathbf{R}(\boldsymbol{\vartheta})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})) \\ \boldsymbol{\Sigma}_\gamma &= (\mathbf{U}'\mathbf{R}(\boldsymbol{\vartheta})^{-1}\mathbf{U} + \mathbf{G}(\boldsymbol{\vartheta})^{-1})^{-1}.\end{aligned}$$

- Kapitel 6, S. 276: Fehler in der vollständig bedingten Dichte für γ_i :

$$\begin{aligned}\gamma_i | \cdot &\sim \text{N}(\boldsymbol{\mu}_{\gamma_i}, \boldsymbol{\Sigma}_{\gamma_i}), \quad i = 1, \dots, m \\ \text{mit } \boldsymbol{\mu}_{\gamma_i} &= (\mathbf{U}'_i \mathbf{U}_i + \sigma^2 \mathbf{D}^{-1})^{-1} \mathbf{U}'_i (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}), \quad \boldsymbol{\Sigma}_{\gamma_i} = \sigma^2 (\mathbf{U}'_i \mathbf{U}_i + \sigma^2 \mathbf{D}^{-1})^{-1},\end{aligned}$$

- Kapitel 6, S. 276: In der vollständig bedingten Dichte für σ^2 müsste \mathbf{R}^{-1} wegfallen. Außerdem müsste für \tilde{a}_0 1/2 durch $n/2$ ersetzt werden, also

$$\begin{aligned}\sigma^2 | \cdot &\sim \text{IG}(\tilde{a}_0, \tilde{b}_0) \\ \text{mit } \tilde{a}_0 &= a_0 + \frac{n}{2}, \quad \tilde{b}_0 = b_0 + \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{U}\boldsymbol{\gamma})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{U}\boldsymbol{\gamma})\end{aligned}$$

- Kapitel 6, S. 276/7: Der Index für die zufälligen Effekte läuft jeweils von 0 bis $q + 1$ (und nicht von 1 bis r), also

$$p(\sigma^2) \propto \sigma^{-2}, \quad p(\mathbf{D}) \propto |\mathbf{D}|^{-\frac{q+1}{2}}$$

mit $q + 1 = \dim(\gamma_i)$.

$$\begin{aligned}\tau_j^2 &\sim \text{IG}(a_j, b_j), \quad j = 0, \dots, q. \\ \tau_j^2 | \cdot &\sim \text{IG}(\tilde{a}_j, \tilde{b}_j), \quad j = 0, \dots, q.\end{aligned}$$

- Kapitel 6, S. 277: In der vollständig bedingten Dichte für τ^2 muss $n/2$ durch $m/2$ ersetzt werden, also

$$\begin{aligned}\tau_j^2 | \cdot &\sim \text{IG}(\tilde{a}_j, \tilde{b}_j), \quad j = 1, \dots, q, \\ \text{mit } \tilde{a}_j &= a_j + \frac{m}{2}, \quad \tilde{b}_j = b_j + \frac{1}{2} \sum_{i=1}^m \gamma_{ij}^2.\end{aligned}$$

- Kapitel 6, S. 283: Fehlendes Transponiertzeichen:

$$\eta_{ir} = \mathbf{x}'_i \boldsymbol{\beta}_r + \mathbf{u}'_i \boldsymbol{\gamma}_r$$

- Kapitel 6, S. 287: Skalare Größen sind fälschlicherweise fett gesetzt:

$$\text{E}(y_{ij} | \gamma_i) = \mu_{ij} = h(\mathbf{x}'_{ij} \boldsymbol{\beta} + \mathbf{u}'_{ij} \boldsymbol{\gamma}_i), \quad i = 1, \dots, m, \quad j = 1, \dots, n_i,$$

- Kapitel 7, Abb. 7.2: In der Bildunterschrift ist die Funktion $f(x)$ falsch spezifiziert. Die richtige Gleichung lautet

$$f(x) = \sin(2(4x - 2)) + 2 \exp(-(16^2)(x - 0.5)^2).$$

- Kapitel 7, S. 296: In der Modellgleichung fehlt (teilweise) der Index i :

$$y_i = \gamma_1 + \gamma_2 z_i + \dots + \gamma_{l+1} z_i^l + \gamma_{l+2} (z_i - \kappa_2)_+^l + \dots + \gamma_{l+m-1} (z_i - \kappa_{m-1})_+^l + \varepsilon_i.$$

- Kapitel 7, S. 297: Falsche Indizierung der Knoten bei der Diskussion der TP-Basis:
Ab dem Knoten $\kappa_2 = 0.1$ wird diese Steigung überlagert durch γ_3 .
- Kapitel 7, S. 305: Fehlerhafte Indizierung in den Formeln für B-Spline-Basisfunktionen und deren Ableitung:

$$\begin{aligned} B_j^1(z) &= \frac{z - \kappa_{j-1}}{\kappa_j - \kappa_{j-1}} I(\kappa_{j-1} \leq z < \kappa_j) + \frac{\kappa_{j+1} - z}{\kappa_{j+1} - \kappa_j} I(\kappa_j \leq z < \kappa_{j+1}) \\ B_j^l(z) &= \frac{z - \kappa_{j-l}}{\kappa_j - \kappa_{j-l}} B_{j-1}^{l-1}(z) + \frac{\kappa_{j+1} - z}{\kappa_{j+1} - \kappa_{j+1-l}} B_j^{l-1}(z) \\ \frac{\partial}{\partial z} B_j^l(z) &= l \cdot \left(\frac{1}{\kappa_j - \kappa_{j-l}} B_{j-1}^{l-1}(z) - \frac{1}{\kappa_{j+1} - \kappa_{j+1-l}} B_j^{l-1}(z) \right) \end{aligned}$$

- Kapitel 7, S. 313: In der Umformung des Integral-Strafterms fällt in der zweiten Zeile das Quadrat weg:

$$\begin{aligned} \int (f''(z))^2 dz &= \int \left(\sum_{j=1}^d \gamma_j B_j''(z) \right)^2 dz \\ &= \int \left(\sum_{i=1}^d \sum_{j=1}^d \gamma_i \gamma_j B_i''(z) B_j''(z) \right) dz \\ &= \sum_{i=1}^d \sum_{j=1}^d \gamma_i \gamma_j \int B_i''(z) B_j''(z) dz \\ &= \boldsymbol{\gamma}' \mathbf{K} \boldsymbol{\gamma}. \end{aligned}$$

- Kapitel 7, S. 313: Falsches Argument in den Funktionsauswertungen:
Der Vektor der beobachteten Funktionswerte $\hat{\mathbf{f}} = (\hat{f}(z_1), \dots, \hat{f}(z_n))'$ berechnet sich damit als...
- Kapitel 7, S.331: Bei dem penalisierten KQ-Kriterium fehlt der Glättungsparameter:

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}}) + \frac{\sigma^2}{\tau^2} \tilde{\boldsymbol{\gamma}}' \mathbf{R} \tilde{\boldsymbol{\gamma}}$$

- Kapitel 7, Seite 332: \mathbf{Z} und $\boldsymbol{\gamma}$ sind jeweils durch $\tilde{\mathbf{Z}}$ und $\tilde{\boldsymbol{\gamma}}$ zu ersetzen:

Kriging als Glättungsverfahren der nichtparametrischen Regression

Die Funktion f wird in Basisfunktionen entwickelt, die sich aus parametrischen Korrelationsfunktionen $\rho(h)$ ergeben. Damit erhält man das penalisierte KQ-Kriterium

$$\text{PKQ}(\lambda) = (\mathbf{y} - \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}})'(\mathbf{y} - \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}}) + \lambda \tilde{\boldsymbol{\gamma}}' \mathbf{K} \tilde{\boldsymbol{\gamma}}$$

mit $\tilde{\mathbf{Z}}[i, j] = \rho(|z_i - z_{(j)}|)$ und $\mathbf{K}[j, k] = \rho(|z_{(j)} - z_{(k)}|)$. Glattheits-Eigenschaften der gewählten Korrelationsfunktion übertragen sich auf die Schätzung $\hat{f}(z)$.

Damit erhält man die Darstellung

$$\mathbf{y} = \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}$$

mit $\tilde{\mathbf{Z}}[i, j] = \rho(|z_i - z_{(j)}|)$, $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_d)'$ und als zu optimierendes Kriterium die penalisierte Residuenquadratsumme

$$\text{PKQ}(\lambda) = (\mathbf{y} - \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}})'(\mathbf{y} - \tilde{\mathbf{Z}}\tilde{\boldsymbol{\gamma}}) + \lambda \tilde{\boldsymbol{\gamma}}' \mathbf{K} \tilde{\boldsymbol{\gamma}}$$

mit $\mathbf{K}[j, k] = \rho(|z_{(j)} - z_{(k)}|)$ und Glättungsparameter $\lambda = \sigma^2/\tau^2$.

- Kapitel 7, Seite 332: Fehlendes Argument in der Korrelationsfunktion:
Dabei ist $c > 0$ eine Konstante, die so bestimmt wird, dass $\rho(c; \phi = 1)$ klein ist ...
- Kapitel 7, S. 336: Falsche Potenz in der Taylorentwicklung (p statt l):

$$f(z_i) \approx f(z) + (z_i - z)f'(z) + (z_i - z)^2 \frac{f''(z)}{2!} + \dots + (z_i - z)^l \frac{f^{(l)}(z)}{l!}.$$

- Kapitel 7, S. 343: Die Definition von A_j bezieht sich auf die wahre Funktion und nicht auf die Funktionsschätzung:
Im vorliegenden Fall bezeichnet A_j das Ereignis „Die Funktion $f(z_j)$ ist nicht im Intervall $[L(z_j), U(z_j)]$ enthalten“.
- Kapitel 7, S. 347: Fehler in der varianzbasierten Definition der äquivalenten Freiheitsgrade:

$$\text{df}_{\text{var}}(\mathbf{S}) = \frac{1}{\sigma^2} \sum_{i=1}^n \text{Var}(\hat{f}(z_i)) = \text{sp}(\mathbf{S}\mathbf{S}')$$

- Kapitel 7, S. 355: Neben den beiden genannten Bedingungen muss noch eine weitere Bedingung an den Rang der beiden Matrizen $\tilde{\mathbf{U}}$ und $\tilde{\mathbf{X}}$ gestellt werden:
3. Die zusammengesetzte Matrix $[\tilde{\mathbf{U}}, \tilde{\mathbf{X}}]$ muss vollen Rang besitzen, damit die Reparametrisierung in $\boldsymbol{\beta}$ und $\tilde{\boldsymbol{\gamma}}$ eindeutig ist.
- Kapitel 7, S. 359: Die vollständig bedingten Dichten für die Varianzparameter hängen von (a_0, b_0) bzw. (a_1, b_1) ab, nicht von (a, b) :

$$\begin{aligned} \tau^2 | \mathbf{y}, \boldsymbol{\gamma}, \sigma^2 &\sim \text{IG}(a_1 + 0.5 \text{rg}(\mathbf{K}), b_1 + 0.5 \boldsymbol{\gamma}' \mathbf{K} \boldsymbol{\gamma}), \\ \sigma^2 | \mathbf{y}, \boldsymbol{\gamma}, \tau^2 &\sim \text{IG}(a_0 + 0.5n, b_0 + 0.5(\mathbf{y} - \mathbf{Z}\boldsymbol{\gamma})'(\mathbf{y} - \mathbf{Z}\boldsymbol{\gamma})). \end{aligned}$$

- Kapitel 7, Seite 361: Der Wertebereich von z wird in links offene und rechte abgeschlossene Intervalle zerlegt:
Der Definitionsbereich von z wird durch die Werte $\kappa_0 < \kappa_1 < \dots < \kappa_d$ zerlegt in Intervalle $R_j = (\kappa_{j-1}, \kappa_j]$, auf denen die Funktion f den Wert γ_j annimmt.
- Kapitel 7, Seite 380: Das penalisierte KQ-Kriterium muss minimiert werden und nicht maximiert, also

$$\sum_{i=1}^n (y_i - f(z_i))^2 + \lambda \int \int \left[\left(\frac{\partial^2}{\partial^2 z_1} + 2 \frac{\partial^2}{\partial z_1 \partial z_2} + \frac{\partial^2}{\partial^2 z_2} \right) f(z_1, z_2) \right]^2 dz_1 dz_2 \rightarrow \min_f$$

- Kapitel 7, Seite 391: Im gewichteten Markov-Zufallsfeld sollten die Gewichte innerhalb der Summe stehen:

$$\gamma_s | \gamma_r, r \in N(s) \sim N \left(\sum_{r: r \sim s} \frac{w_{sr}}{w_{s+}} \gamma_r, \frac{\tau^2}{w_{s+}} \right)$$

- Kapitel 7, S. 396: Falscher Index in der additiven Struktur (q statt k):

$$f(z_1, \dots, z_q) = f_1(z_1) + \dots + f_q(z_q).$$

- Kapitel 8, S. 406: Fehler in der Bildunterschrift zu Abbildung 8.3. Gestrichelte und durchgezogene Linie sind vertauscht, es sollte also heißen

Abb. 8.3. *Mietspiegel München: Kerndichteschätzer für den räumlichen Effekt. Gestrichelte Linie: Modell ohne Expertenurteil, durchgezogene Linie: Modell mit Expertenurteil.*

- Kapitel 8, S. 415: Fehlerhafter Index bei der Interpretation des variierenden Koeffizienten:

Abbildung 8.8 (links) zeigt den Effekt \hat{f}_1 des Alters des Versicherungsnehmers für Frauen und den Alterseffekt $\hat{f}_1 + \hat{f}_6 + \hat{\beta}_1$ für Männer ($geschl = 1$). Die rechte Grafik von Abbildung 8.8 zeigt den in Abhängigkeit des Alters variierenden Effekt $\hat{f}_6 + \hat{\beta}_1$ der Männer im Vergleich zu den Frauen.

- Kapitel 8, S. 418: Fehlerhafte Variablenbezeichnung. In der Aufzählung sollte die Variable *art* durch *artkat* ersetzt werden, also
 - Die Effekte der Gründigkeit (*grund*), des Ph-Wertes (*ph*), der Art des Waldes (*artkat*), des Alkali-Anteils (*alkali*), der Hangneigung (*hang*) und der Höhenlage (*hoehe*) verschwinden oder sind zumindest stark verringert. Bei all diesen Variablen handelt es sich um *standortspezifische Kovariablen*, deren Werte sich über die Zeit nicht oder kaum verändern.
- Kapitel 8, S. 422: Fehler in der Score-Funktion:

$$\begin{aligned} s_{pen,0}(\gamma_1, \dots, \gamma_q, \beta) &= s_0(\gamma_1, \dots, \gamma_q, \beta) = \mathbf{0} \\ s_{pen,j}(\gamma_1, \dots, \gamma_q, \beta) &= s_j(\gamma_1, \dots, \gamma_q, \beta) - \lambda_j \mathbf{K}_j \gamma_j = \mathbf{0}. \end{aligned}$$

- Kapitel 8, S. 426: Fehlender Index j in Formel (8.12) (τ^2 statt τ_j^2):

$$p(\gamma_j | \tau_j^2) \propto \left(\frac{1}{\tau_j^2} \right)^{\text{rg}(\mathbf{K}_j)/2} \exp \left(-\frac{1}{2\tau_j^2} \gamma_j' \mathbf{K}_j \gamma_j \right) \quad (8.12)$$

- Kapitel 8, S. 426: Fehlender Fettdruck in der Formel für die Posteriori-Verteilung:

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{y}) &\propto \frac{1}{(\sigma^2)^{\frac{q}{2}}} \exp \left(-\frac{1}{2\sigma^2} (\mathbf{y} - \boldsymbol{\eta}^{strukt})' (\mathbf{y} - \boldsymbol{\eta}^{strukt}) \right) \\ &\prod_{j=1}^q \frac{1}{(\tau_j^2)^{\text{rg}(\mathbf{K}_j)/2}} \exp \left(-\frac{1}{2\tau_j^2} \gamma_j' \mathbf{K}_j \gamma_j \right) \prod_{j=1}^q (\tau_j^2)^{-a_j-1} \exp \left(-\frac{b_j}{\tau_j^2} \right) \\ &(\sigma^2)^{-a_0-1} \exp \left(-\frac{b_0}{\sigma^2} \right), \end{aligned}$$

- Kapitel 8, Seite 427: Fehler in der Kovarianzmatrix:

$$\Sigma_{\beta} = \text{Cov}(\beta | \cdot) = \sigma^2 (\mathbf{X}' \mathbf{X})^{-1}$$

- Kapitel 8, S. 429: Fehler im additiven Prädiktor (*ansp* statt *anspr*):

$$\eta_i^{add} = f(aland_i) + \beta_0 + \beta_1 jahr_i + \beta_2 azit_i + \beta_3 ansp_i + \beta_4 uszw_i + \beta_5 patus_i + \beta_6 patdsg_i.$$

- Kapitel 8, Seite 441: Fehler in der Interpretation des Effekts des BMI der Mutter:
Basierend auf diesen Ergebnissen sinkt also der durchschnittliche Ernährungszustand bei stark übergewichtigen Müttern nicht wie eigentlich mit einem invers U-förmigen Effekt postuliert.
- Anhang A, Seite 450: Fehler in der Definition von Spalten- und Zeilenraum:
Der Spaltenraum $Z(\mathbf{A})$ einer $n \times p$ Matrix \mathbf{A} ist der durch die Spalten von \mathbf{A} aufgespannte Unterraum des \mathbb{R}^n :

$$Z(\mathbf{A}) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{A}\mathbf{y} \text{ für ein } \mathbf{y} \in \mathbb{R}^p\}.$$

Analog lässt sich der Zeilenraum (als Teilraum des \mathbb{R}^p) definieren.

- Anhang A, S. 451: Fehlender Fettdruck in der Formel für die Determinante der Inversen: Für die Determinante einer $n \times n$ Matrix \mathbf{A} gilt:

$$4. |\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}.$$

- Anhang B, Seite 459: Fehler in der Dichte der Lognormalverteilung:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{x} \exp\left(-(\log(x) - \mu)^2 / 2\sigma^2\right), \quad x > 0.$$

- Anhang B, Seite 460: Fehler in der Dichte der Gammaverteilung:

In Abhängigkeit vom Erwartungswert μ und dem Skalenparameter ν ist die Dichte gegeben durch

$$f(x) = \frac{1}{\Gamma(\nu)} \left(\frac{\nu}{\mu}\right)^\nu x^{\nu-1} \exp\left(-\frac{\nu}{\mu}x\right), \quad x > 0.$$

- Anhang B, S. 461: Fehler in der Varianz der inversen Gammaverteilung:
Man erhält $E(X) = b/(a-1)$ und $\text{Var}(X) = b^2/((a-1)^2(a-2))$.
- Anhang B, S. 466: Für den deterministischen Anteil sollte die Kovarianzmatrix betrachtet werden:

\mathbf{Y}_2 ist der deterministische Anteil von \mathbf{X} mit

$$E(\mathbf{Y}_2) = \mathbf{H}'\boldsymbol{\mu} \quad \text{und} \quad \text{Cov}(\mathbf{Y}_2) = \mathbf{0}.$$

- Anhang B, S. 479: Fehler in der Normierungskonstante der Gammaverteilung:
Damit hat λ die Dichte

$$p(\lambda) = k \lambda^{a-1} \exp(-b\lambda)$$

$$\text{mit } k = \frac{b^a}{\Gamma(a)}.$$

- Anhang B, S. 480: Der Posteriori-Modus sollte definiert sein als

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta} | \mathbf{y}) = \arg \max_{\boldsymbol{\theta}} p(\mathbf{y} | \boldsymbol{\theta}) p(\boldsymbol{\theta}).$$

- Anhang B, S. 479: In der Überdeckungswahrscheinlichkeit sollte das Teilmengenzeichen durch ein Elementzeichen ersetzt werden:

$$P(\boldsymbol{\theta} \in C | \mathbf{y}) \geq 1 - \alpha$$

- Index: Ab Kapitel 7 (Seite 291) sind die Index-Einträge um zwei Seiten verschoben (z.B. Äquivalente Freiheitsgrade: Seite 347 statt 345). Eine korrigierte Version des Index ist von der Errata-Seite erhältlich.