

Übung 4: Quantil-Regression & Verallgemeinerte lineare Modelle 1

Aufgabe 1

Bestimmen Sie die Quantilsfunktion einer $\text{Par}(\alpha, \lambda)$ -Verteilung mit der Verteilungsfunktion

$$F_X(x) = 1 - \left(\frac{\lambda}{\lambda + x} \right)^\alpha \quad \text{für } x > 0$$

mit $\alpha, \lambda > 0$.

Aufgabe 2

Bestimmen Sie für die folgenden Zufallsvariablen jeweils den Median als Lösungen des Minimierungsproblems

$$q_{1/2} = \arg \min_{a \in \mathbb{R}} \mathbb{E}[\rho_{1/2}(X - a)].$$

- Die Zufallsvariable X sei eine auf der Menge $\{1, 2, \dots, 9\}$ gleichverteilt.
- Die Zufallsvariable X sei $\text{Exp}(\lambda)$ -verteilt.

Aufgabe 3

Betrachtet wird das lineare Quantil-Regressionsmodell

$$y = \mathbf{x}^T \boldsymbol{\beta}_\tau + \varepsilon_\tau,$$

wobei zusätzlich angenommen wird, dass die abhängige Variable y eine asymmetrische Laplace-Verteilung, d.h. die Dichtefunktion

$$f(y) = \frac{\tau(1-\tau)}{\sigma} \exp\left(-\rho_\tau\left(\frac{y-\mu}{\sigma}\right)\right) \quad \text{mit } \mu \in \mathbb{R}, \sigma > 0 \text{ und } \tau \in (0, 1)$$

besitzt. Weisen Sie nach, dass der QR-Schätzer $\hat{\boldsymbol{\beta}}_\tau$ unter dieser zusätzlichen Voraussetzung auch der ML-Schätzer für den Vektor mit den Regressionskoeffizienten $\boldsymbol{\beta}_\tau$ ist.

Hinweis: Verwenden Sie dabei, dass links von μ die Wahrscheinlichkeitsmass der asymmetrischen Laplace-Verteilung genau τ und rechts von μ dementsprechend genau $1 - \tau$ beträgt. D.h. dass das τ -Quantil einer asymmetrischen Laplace-Verteilung bei μ liegt.

Aufgabe 4

Weisen Sie explizit nach, dass die Geometrische-Verteilung $\text{Geo}(p)$ mit der Wahrscheinlichkeitsfunktion

$$f(y; p) = \begin{cases} p(1-p)^{y-1} & \text{für } y \in \mathbb{N} \\ 0 & \text{sonst} \end{cases}$$

mit $p \in (0, 1)$ zur Exponential-Dispersions-Familie gehört und berechnen Sie den Erwartungswert und die Varianz der Verteilung.

Aufgabe 5

Ein Kreditinstitut sucht ein geeignetes Modell für die Einschätzung der Kreditwürdigkeit von potenziellen Kreditnehmern. Dabei werden die Merkmale „Alter“ (x_1), „Monatsgehalt“ (x_2), „Laufzeit“ (x_3) und „Höhe des Kredites“ (x_4) erhoben und die abhängige Variable y gibt an, ob der potenzielle Kreditnehmer kreditwürdig ist ($y = 1$) oder nicht ($y = 0$). Mit Hilfe eines Datensatzes bestehend aus $n = 300$ Beobachtungen wurden ein klassisches lineares Modell, ein Logit- und ein Probit-Modell angepasst. Man erhielt folgende Ergebnisse:

	Lineares Modell		Logit-Modell		Probit-Modell	
	Estimate	p value	Estimate	p value	Estimate	p value
$\hat{\beta}_0$	-1,105	$1,81 \cdot 10^{-4}$	-8,437	$3,34 \cdot 10^{-7}$	-4,627	$9,67 \cdot 10^{-7}$
$\hat{\beta}_1$	$1,328 \cdot 10^{-2}$	$2,15 \cdot 10^{-3}$	$6,936 \cdot 10^{-2}$	$2,40 \cdot 10^{-3}$	$3,820 \cdot 10^{-2}$	$4,27 \cdot 10^{-3}$
$\hat{\beta}_2$	$2,031 \cdot 10^{-4}$	$1,19 \cdot 10^{-4}$	$1,068 \cdot 10^{-3}$	$1,46 \cdot 10^{-4}$	$5,683 \cdot 10^{-4}$	$5,24 \cdot 10^{-4}$
$\hat{\beta}_3$	$2,790 \cdot 10^{-2}$	$3,49 \cdot 10^{-5}$	$1,456 \cdot 10^{-1}$	$7,09 \cdot 10^{-5}$	$8,455 \cdot 10^{-2}$	$6,68 \cdot 10^{-5}$
$\hat{\beta}_4$	$2,303 \cdot 10^{-6}$	$2,23 \cdot 10^{-2}$	$1,159 \cdot 10^{-5}$	$2,46 \cdot 10^{-2}$	$6,859 \cdot 10^{-6}$	$2,25 \cdot 10^{-2}$

- Beurteilen Sie für einen Neukunden, der 52 Jahre alt ist, ein Monatsgehalt von 4413 € erhält und einen Kredit in Höhe von 103.361 € mit einer Laufzeit von 17 Jahren beantragt hat die Kreditwürdigkeit mit Hilfe der drei angegebenen Modelle. Interpretieren Sie die Ergebnisse und beurteilen Sie, welches der Modelle am wenigsten geeignet ist.
- Welchen von den folgenden 10 potenziellen Kreditnehmern sollte das Kreditinstitut einen Kredit gewähren, wenn es für seine Einschätzung das Logit- und Probit-Modell verwendet und einer Person ein Kredit gewährt werden soll, wenn deren Ausfallwahrscheinlichkeit kleiner als 50% ist? Geben Sie den Anteil an richtig klassifizierten Kreditnehmern an.

i	x_1	x_2	x_3	x_4	y_i
1	52	4413	17	103361	1
2	45	3649	17	61044	1
3	37	3548	14	85822	1
4	52	4040	20	84167	1
5	38	5134	11	116721	1
6	40	2609	13	65051	0
7	44	4571	13	67189	0
8	23	2333	6	57439	0
9	32	4041	7	74440	0
10	31	4013	10	62341	0

- Wie würden sich die Anteile der richtig klassifizierten potenziellen Kreditnehmer jeweils ändern, wenn die Kosten für einen nicht zurückgezahlten Kredit im Durchschnitt 6 mal so hoch sind, wie die Kosten eines irrtümlicherweise nicht vergebenen Kredites? Wie verändert sich der Anteil der vergebenen Kredite?