Masterkurs Wirtschaftsmathematik und Statistik

Dauer der Prüfung: 90 Minuten **28. Juni 2023**

ZUNAME:			
VORNAME:		MATR.NR.:	

ERLAUBT: **Formelsammlung des Instituts, nicht-programmierbare Taschenrechner**

VERBOTEN: **alle sonstigen Unterlagen, programmierbare Taschenrechner, Handys, etc.**

Für eine positive Beurteilung müssen aus den Beispielen 1,2,3 sowie 4,5 jeweils mindestens 10 Punkte erreicht werden!

1. a) (6 Punkte) Betrachtet wird eine logistische Regression basierend auf der linearen Funktion $z = -2 - 3x_1 + 2x_2 + x_3$.

   Es ist eine Stichprobe mit den folgenden 5 Datensätzen gegeben:

i	x_1	x_2	x_3	y
1	1	1	1	0
2	-1	-1	-1	0
3	-1	-2	1	0
4	2	2.5	3	0
5	-2.5	-1.5	-3	0

   Bestimmen Sie die hit-ratio und die Devianz für das Modell. Testen Sie mit Hilfe der Devianz die Güte des Modells ($\alpha = 0.05$). Was ist die Aussage Ihres Tests? Was können Sie näherungsweise über das Signifikanzniveau dieser Aussage aussagen?
   [$h = 4/5$, $Dev = 3.096$, $K = [3.841, \infty)$, $\alpha^* \approx 0.08$]

   b) (2 Punkte) Gegeben ist nun die Beobachtung: $x_1 = 2$, $x_2 = 1$, $x_3 = 4$. Wie lauten die Odds für das Eintreten von y? Wie ändern sich diese Odds näherungsweise, wenn x_2 auf 2 erhöht wird? [Odds: 0.135, Erhöhung um den Faktor e^2 auf 1]

   c) (2 Punkte), unabhängig von a.,b.
   Gegeben ist eine logistische Regression mit der linearen Funktion $z = 3 + b_1 x_1 + 2.5x_2 - 2x_3$. Welche Werte für b_1 aus den natürlichen Zahlen liefern die größte hit-ratio für die beiden Datensätze $x_1 = -3$, $x_2 = 2$, $x_3 = -2$, $y = 1$ sowie $x_1 = 2$, $x_2 = -1$, $x_3 = 2$, $y = 0$?
   [$b_1 = 1$]

Quantile der χ^2-Verteilungen:

f	γ =0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995	0,999
1	0,0	0,0002	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879	10,828
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597	13,816
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838	16,266
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860	18,467

2. Gegeben sind 9 Objekte mit jeweils einem reellen Messwert z_j.

Objekt j	1	2	3	4	5	6	7	8	9
z_j	-12	-9	-6	-2	-1	-1	3	7	13

   a) (5 Punkte) Führen Sie das hierarchisch-agglomerative Clusterungsverfahren mit *Centroid Clustering* bis zum Ende durch.

   b) (3 Punkte), unabhängig von a.)
   Gegeben ist die folgende Clusterung der obigen Objekte:
   $C_1 = \{1, 2, 3\}$, $C_2 = \{4, 5, 6, 7\}$, $C_3 = \{8, 9\}$.
   Bestimmen Sie für diese Clusterung den ESS (error sum of squares) Wert. Bestimmen Sie auch die Hausdorff-Distanz der Cluster C_1 und C_2. [ESS= 50.75, 10]

3. Gegeben ist die folgende Funktion:

$$f(x, y, z) = \frac{x^2}{y} + e^{xy} + \frac{\ln z}{z}$$

   a) (7 Punkte) Bestimmen Sie die Hessematrix von f im Punkt $P = (0, 1, 1)$.

   b) (3 Punkte) Bestimmen Sie die Definitheit der Matrix aus 3a).

   [

$$H_f(0,1,1) \begin{pmatrix} 3 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

   ist indefinit]

4. Gegeben ist das Optimierungsproblem

$$\begin{array}{rlcl} f(x,y,z) = & x + y + z & & \\ \text{unter} & x^2 + y^2 & = & 4 \\ & x + z & = & 1 \end{array}$$

   mit der Zielfunktion $f(x, y, z)$ und zwei Nebenbedingungen.

   a) (6 Punkte) Bestimmen Sie alle kritischen Punkte des Problems mit der Lagrange-Methode!

   b) (6 Punkte) Untersuchen Sie jenen kritischen Punkt, der den größten Zielfunktionswert hat, und stellen Sie mit der geränderten Hessematrix fest, ob es sich um einen Hoch-, Tief- oder Sattelpunkt handelt!

   [kritische Punkte $P_1 = (0, 2, 1)$ und $P_2\,(0, -2, 1)$. P_1 ist ein Maximum.]