

Aufgabe 13:

Ein Datensatz zu 474 Angestellten einer US Bank (R package AER) enthält die Variable `job` mit den Ausprägungen `custodial`, `admin` und `manage`. Das multinomiale Logit-Modell mit den Einflussgrößen `education` (metrisch in Jahren) und `minority` (kategorial) ergibt:

Call:

```
multinom(formula = job ~ education + minority, data = BankWages)
```

Coefficients:

	(Intercept)	education	minorityyes
admin	-2.232145	0.4550169	-1.174635
manage	-30.268977	2.2002726	-3.293551

Std. Errors:

	(Intercept)	education	minorityyes
admin	0.9622174	0.08961043	0.4347329
manage	4.1483824	0.27296417	0.8724398

Dieser output wurde mit der Funktion `multinom` aus dem Paket `nnet` in R erzeugt.

- (a*) Interpretieren Sie den Koeffizient 0.4550169 (vgl. letzte Übung).
- (b) Erklären Sie, wie im multinomialen Logit-Modell die Signifikanz einzelner Parameter getestet werden kann.
- (c) Testen Sie, ob der Zusammenhang zwischen der Dauer der Ausbildung und einem Job im Management zum Signifikanzniveau von 5% signifikant ist.

Im Folgenden soll die ordinale Struktur der Variable `job` berücksichtigt werden (`custodial < admin < manage`). Für alle männlichen Angestellten wird in R mit der Funktion `vglm` aus dem Paket `VGAM` ein kumulatives Logit-Modell gefittet:

Call:

```
vglm(formula = job ~ education + minority, family = cumulative(parallel = T),  
      data = BankWages, subset = gender == "male")
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept):1	7.95135	1.07295	7.411	1.26e-13 ***
(Intercept):2	14.17211	1.46984	9.642	< 2e-16 ***
education	-0.87000	0.09288	-9.367	< 2e-16 ***
minorityyes	1.05644	0.41066	2.573	0.0101 *

- (d) Geben Sie das zugrundeliegende Modell an.
- (e) Interpretieren Sie den geschätzten Koeffizienten für `minority` und einen der Intercepts.
- (f) Mit der Funktion `polr` aus dem Paket `MASS` lassen sich auch kumulative Modelle schätzen. Vergleichen Sie die beiden Modelle. Wie lassen sich die Unterschiede erklären?

Call:

```
polr(formula = job ~ education + minority, data = BankWages,  
      subset = gender == "male")
```

Coefficients:

	Value	Std. Error	t value
education	0.870	0.09307	9.348
minorityyes	-1.056	0.41199	-2.564

Intercepts:

	Value	Std. Error	t value
custodial admin	7.9514	1.0769	7.3833
admin manage	14.1721	1.4744	9.6124

Aufgabe 14:

Die Münchner Gründerstudie untersucht die Überlebenschancen von in den Jahren 1985/6 neu gegründete Unternehmen. Unter Anderem wurde das Alter des/r Gründers/in (`age`) in Jahren erhoben. Daneben gibt es die Information über den Art des Kundenkreises (`kart` = 1: breit gestreut, 0: kleine Zahl großer Kunden) und ob das Startkapital mehr als 25.000 DM betrug (ja: `stkap` = 1). Die Überlebenszeit (`b7`) (bis zur Insolvenz) wird in 6-Monats-Intervallen gemessen (1: kleiner/gleich 6 Monate, 2: länger als 6 Monate kleiner/gleich 12 Monate, ..., 7: länger als 36 Monate). Mit der Funktion `vglm` aus dem Paket `VGAM` wurde ein sequentielles logit Modell geschätzt:

Call:

```
vglm(formula = b7 ~ stkap + kart + age, family = sratio(link = "logit",  
               parallel = T), data = founder)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept):1	-1.921405	0.314757	-6.104	1.03e-09	***
(Intercept):2	-1.364200	0.303741	-4.491	7.08e-06	***
(Intercept):3	-1.569356	0.311843	-5.033	4.84e-07	***
(Intercept):4	-1.597379	0.315962	-5.056	4.29e-07	***
(Intercept):5	-1.814809	0.326582	-5.557	2.74e-08	***
(Intercept):6	-1.780629	0.328528	-5.420	5.96e-08	***
stkap1	-1.121824	0.127188	-8.820	< 2e-16	***
kart1	0.511786	0.147822	3.462	0.000536	***
age	-0.028313	0.006651	-4.257	2.07e-05	***

- Geben Sie das zugrundeliegende Modell an.
- Interpretieren Sie den geschätzten Koeffizienten für das Startkapital.

Im Folgenden soll für ein Unternehmen, das von einer/m 25-Jährigen mit weniger als 25.000 DM Startkapital und breit gestreutem Kundenkreis gegründet wurde, folgende Werte berechnet werden:

- Bestimmen Sie für dieses Unternehmen den Hazard für das 1-te Zeitintervall.
- Wie hoch ist die Wahrscheinlichkeiten, dass dieses Unternehmen nicht im ersten Jahr Insolvenz anmelden muss.
- Skizzieren Sie die Survival-Kurven getrennt nach breit gestreutem und kleinem Kundenkreis (bei ansonsten gleichen Eigenschaften der Unternehmen) für alle Zeitpunkte.