

Logistische Regression

Ziel der Übungseinheit

Die dritte Übungseinheit beschäftigt sich mit der logistischen Regression. Grundsätzlich wird in einem Regressionsmodell der Zusammenhang zwischen einer Zielgröße und einer oder mehreren Einflussgrößen durch eine Funktion beschrieben. Die Regressionskoeffizienten werden aus den beobachteten Daten geschätzt.

Man verwendet die lineare Regression, wenn die Zielgröße stetig ist. (Die Bezeichnung „linear“ bedeutet nicht, dass der Zusammenhang zwischen den Einflussgrößen und der Zielgröße linear sein muss. D.h. die Funktion, die den Zusammenhang beschreibt, muss nicht unbedingt eine Gerade, sondern kann auch z.B. eine quadratische Funktion sein). Die logistische Regression wird verwendet, wenn die Zielgröße binär ist. Die Einflussgrößen können in beiden Fällen sowohl quantitativ als auch qualitativ sein.

Studienbeschreibung

Wie in der zweiten Praktikumseinheit liegt der dritten Übungseinheit der anonymisierte Teil-Datensatz „AML99.sav“ der AMLCG-99-Studie zu Grunde. Hierbei handelt es sich um eine multizentrische, randomisierte, klinische Studie zur Therapie der akuten myeloischen Leukämie bei Erwachsenen. Diese Studie wurde von der deutschen *Acute Myeloid Leukemia Cooperative Group (AMLCG)* durchgeführt. Koordiniert wurde sie in Münster, München und Braunschweig. Die statistische Auswertung fand/findet im Institut für Biometrie und Klinische Forschung (IBKF) der WWU Münster statt. Ziel der Studie ist der randomisierte Vergleich zweier Induktions- und zweier Erhaltungstherapien.

Die Behandlung der AML erfolgt durch intensive Chemotherapie. Diese wird in zwei Stufen durchgeführt. Die erste Stufe bildet die sogenannte Induktionstherapie, deren Ziel das Erreichen einer kompletten Remission (complete remission, CR) ist. Erreichen der CR bedeutet, dass im Knochenmark weniger als 5% Blasten, im Blut Neutrophile über 1.500/cmm und Thrombozyten über 100.000/cmm vorliegen.¹

Die Behandlung besteht aus mehrtägigen Chemotherapiekursen. Die in dieser Studie eingesetzten Substanzen waren **TAD** (Thioguanin, **ARA-C** (Cytarabin) und **Daunorubicin**), bzw. **HAM** (High dose **ARA-C** und **Mitoxantron**). Im Rahmen der Studie sollte die Standard-Doppelinduktionstherapie TAD-HAM mit der innovativen Doppelinduktionstherapie HAM-HAM verglichen werden.

Die Ergebnisse wurden u. a. zusammen mit den Daten der früheren Studie AMLCG92 publiziert². Gegenstand dieser Publikation war es, die Altersabhängigkeit der Therapieergebnisse aufzuzeigen.

¹ Büchner et al: Double induction containing either two courses or one course of high-dose cytarabine plus mitoxantrone and postremission therapy by either autologous stem-cell transplantation or by prolonged maintenance for acute myeloid leukemia. J Clin Oncol. 2006 Jun 1;24(16):2480-9.

² Büchner T, Berdel WE, Haferlach C, Haferlach T, Schnittger S, Müller-Tidow C, Braess J, Spiekermann K, Kienast J, Staib P, Grüneisen A, Kern W, Reichle A, Maschmeyer G, Aul C, Lengfelder E, Sauerland MC, Heinecke A, Wörmann B, Hiddemann W: Age-related risk profile and chemotherapy dose response in acute myeloid leukemia: a study by the German Acute Myeloid Leukemia Cooperative Group. J Clin Oncol. 2009; 27: 61–69.

Datenbeschreibung

Variablenname	Bedeutung und Kodierung
PID	Laufende Nummer (zufällig)
THERAPIE	Therapiegruppe <i>1 = TAD-HAM</i> <i>2 = HAM-HAM</i>
GESCHLECHT	Geschlecht <i>1 = Männlich</i> <i>2 = Weiblich</i>
ALTER	Alter bei Diagnosestellung (in Jahren)
KGW	Körpergewicht in kg
KGR	Körpergröße in cm
ZUSTAND	Allgemeinzustand bei Aufnahme <i>0 = Normal</i> <i>1 = Leicht beeinträchtigt</i> <i>2 = Stark beeinträchtigt</i>
ZUSTAND_BIN	Allgemeinzustand bei Aufnahme mit zusammengefassten Kategorien <i>0 = Normal oder leicht beeinträchtigt</i> <i>1 = Stark beeinträchtigt</i>
FIEBER	Körpertemperatur in °C
log_LEUKO_0	Logarithmierte Leukozytenzahl bei Aufnahme
HB_0	Hb in g/dl bei Aufnahme
HB_2	Hb in g/dl nach Induktionskurs 2
LDH_0	Lactatdehydrogenase (LDH) in U/l bei Aufnahme
log_LDH_0	Logarithmierter LDH-Wert bei Aufnahme
THROMBO_0	Thrombozyten pro mm ³ bei Aufnahme
THROMBO_2	Thrombozyten pro mm ³ nach Induktionskurs 2
BLASTEN_0	Blasten im Knochenmark in % bei Aufnahme
BLASTEN_2	Blasten im Knochenmark in % nach Induktionskurs 2
KARYOTYP	<i>Karyotyp</i> <i>1 = Günstig</i> <i>2 = Intermediär</i> <i>3 = Ungünstig</i>
CR	Komplette Remission (CR) nach Induktion <i>0 = Nein</i> <i>1 = Ja</i>

Aufgabe 1

Welchen Merkmalstyp / welches Skalenniveau haben die folgenden Variablen?

- a) Alter (ALTER): _____
- b) Geschlecht (GESCHLECHT): _____
- c) Allgemeinzustand (ZUSTAND): _____

Aufgabe 2

Untersuchen Sie mit einer geeigneten deskriptiven Darstellung den Einfluss folgender Variablen auf das Erreichen der kompletten Remission (CR):

- Geschlecht
- Karyotyp
- Alter

Welche dieser Variablen scheinen einen Einfluss auf die CR-Rate / auf die Chance einer CR zu haben?

Aufgabe 3

a) Mit welcher Methode können Sie den gemeinsamen Einfluss von mehreren Variablen auf eine binäre Zielgröße (z.B. das Erreichen der kompletten Remission) überprüfen?

- ☐ Cox-Regression
- ☐ Logistische Regression
- ☐ Lineare Regression

- b) Führen Sie die ausgewählte Methode durch und untersuchen Sie den gemeinsamen Einfluss von Geschlecht (GESCHLECHT), der Behandlungsgruppe (THERAPIE), dem Alter (ALTER), dem Hb-Wert bei Aufnahme (HB_0), dem körperlichen Zustand (ZUSTAND) und dem Karyotyp (KARYOTYP) auf das Erreichen der kompletten Remission (CR).

Achten Sie dabei darauf, dass für folgende kategoriale Variablen die Referenzkategorie wie folgt gewählt wird:

Variablenname	Referenzkategorie
GESCHLECHT	<i>Männlich</i>
THERAPIE	<i>TAD-HAM</i>
ZUSTAND	<i>Stark beeinträchtigt</i>
KARYOTYP	<i>Ungünstig</i>

Vervollständigen Sie die folgende Tabelle mit Ihren Ergebnissen:

Variable	Odds Ratio	95%-Konfidenzintervall
Intercept	1,386	-
Zustand – normal vs. stark beeinträchtigt		
Zustand – leicht beeinträchtigt vs. stark beeinträchtigt		
Karyotyp – günstig vs. ungünstig		
Karyotyp – intermediär vs. ungünstig		
Behandlungsgruppe – HAM-HAM vs. TAD-HAM		
Geschlecht – weiblich vs. männlich		
Hb (x g/dl vs. (x-1) g/dl)		
Alter (x Jahre vs. (x-1) Jahre)		

c) Führen Sie eine Rückwärts-Selektion mittels Likelihood-Ratio-Tests (LR-Tests) durch.

Welche Variablen erweisen sich als wichtige Einflussgrößen für das Erreichen der kompletten Remission? Wie groß und wie sicher ist ihr Einfluss? Tragen Sie Ihre Ergebnisse in die folgende Tabelle ein.

Variable	Odds Ratio	95%-Konfidenzintervall	p-Wert des Likelihood Ratio-Tests

Wie interpretieren Sie diese Ergebnisse?

Aufgabe 4

Wie unterscheidet sich die Chance auf komplette Remission eines 60-jährigen von der Chance eines 50-jährigen, wenn sich diese beiden Patienten nur hinsichtlich des Alters unterscheiden?

Bonusfrage 1

Für Patienten im Alter von über 60 Jahren, bei denen prinzipiell eine intensive Chemotherapie möglich ist, wurden in Zusammenarbeit mit den Studiengruppen German AML Cooperative Group (AMLCG) und Studienallianz Leukämie (SAL) Prognosescores entwickelt, die die Wahrscheinlichkeit des Ansprechens einer intensiven Chemotherapie und das Risiko einer therapie-assoziierten Mortalität voraussagen^{3,4}. Diese Scores wurden durch eine retrospektive Analyse der in der AMLCG99-Studie intensiv behandelten älteren Patienten ermittelt und anhand einer unabhängigen Kohorte in der AML96-Studie der SAL intensiv behandelter Patienten überprüft.

Ermitteln Sie mit Hilfe dieses Scores die Wahrscheinlichkeit für das Erreichen einer kompletten Remission für einen 67-jährigen Patienten mit einer ‚de novo AML‘ und normalem Karyotyp, der bei Erstdiagnose Fieber hatte, und zu dem folgenden Laborwerte dokumentiert wurden: Hb-Wert 9,8 g/dl, Thrombozyten 120.000 pro µl, LDH 580 U/L und Fibrinogen 170 mg/dl.

Webseite: www.aml-score.org

Benutzername: *IBKF3*

Passwort: *ibkf3*

Wahrscheinlichkeit für Erreichen der kompletten Remission: _____

Mit welchem Ihnen bekannten statistischen Verfahren kann ein Prognosescore entwickelt werden?

³ Krug U, Röllig C, Koschmieder A, Heinecke A, Sauerland MC, Schaich M, Thiede C, Kramer M, Braess J, Spiekermann K, Haferlach T, Haferlach C, Koschmieder S, Rohde C, Serve H, Wörmann B, Hiddemann W, Ehninger G, Berdel WE, Büchner T, Müller-Tidow C, German Acute Myeloid Leukaemia Cooperative Group, Study Alliance Leukemia Investigators: Complete remission and early death after intensive chemotherapy in patients aged 60 years or older with acute myeloid leukaemia: a web-based application for prediction of outcomes. Lancet 2010; 376: 2000–8.

⁴ Behandlung älterer Patienten mit akuter myeloischer Leukämie
The Treatment of Elderly Patients With Acute Myeloid Leukemia
Dtsch Arztebl Int 2011; 108(51-52): 863-70; DOI: 10.3238/arztebl.2011.0863.

Bonusfrage 2

Bestimmen Sie, um welchen Faktor sich die Chance auf komplette Remission der folgenden beiden Patienten unterscheidet:

Patient 1 ist 50 Jahre alt, Patient 2 hat ein Alter von 60 Jahren.

Beide Patienten sind körperlich leicht beeinträchtigt.

Patient 1 weist den günstigen Karyotyp auf, Patient 2 jedoch den ungünstigen.
