

## Projekt 7: Analyse von Lebenszeiten anhand einer Krebsstudie

Die Daten im Datensatz `HEADNECK.DAT` stammen aus einer Studie, in der bei Krebsformen des Kopfes und des Halses zwei Therapien verglichen wurden. Bei der einen Therapie wurde nur Bestrahlung angewendet, bei der anderen Therapie wurde zusätzlich zur Bestrahlung noch eine Chemotherapie durchgeführt. Dabei wurden die Patienten zufällig einer der beiden Therapien zugeordnet. Der Datensatz hat folgende Gestalt:

Nur Bestrahlung				Bestrahlung und Chemotherapie			
y	d	y	d	y	d	y	d
7	1	185	0	37	1	519	1
34	1	218	1	84	1	528	0
42	1	225	1	92	1	547	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Die Variable  $y$  gibt die Tage nach der Therapie bis zum Rückfall bzw. bis zum Ende der Studie an. Die Variable  $d$  zeigt an, ob der Rückfall beobachtet wurde ( $d = 1$ ) oder ob vorher die Studie beendet wurde ( $d = 0$ ).

### Aufgaben

1. Wie groß sind die geschätzten Wahrscheinlichkeiten für beide Therapien, dass der Rückfall erst nach 50, nach 100, nach 200 bzw. nach 800 Tagen nach der Therapie kommt? Was ist der geschätzte Median für die Zeit bis zum Rückfall bei beiden Therapien?  
Schätzen Sie dazu die Rückfallzeit-Funktionen (survival function) für beide Therapien mit und ohne Berücksichtigung der Zensierung, d.h. mit und ohne Berücksichtigung der Variable  $d$ . Welchen Effekt hat die Berücksichtigung der Zensierung?
2. Wie sehen für die beiden Therapien die Rückfallrisiken nach 50, 100, 200 und 800 Tagen aus?
3. Kann man schließen, dass eine der beiden Therapien die Rückfallzeit verlängert?

### Literatur

- Sachs, L. und Hedderich, J. (2009). *Angewandte Statistik: Methodensammlung mit R*. Springer, Berlin.
- Therneau, T.M. und Grambsch, P.M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer, New York.

Klein, J.P. und Moeschberger, M.L. (2003) *Survival Analysis. Techniques for Censored and Truncated Data*, Springer, New York.

Kleinbaum, D.G. und Klein, M. (2005). *Survival Analysis. A Self-Learning Text*. Springer, New York.

### **Literatur zu den Daten**

Efron, B. (1988). Three examples of computer intensive statistical inference. *Sankhya, Series A*, **50**, 338-362.

aus

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. und Ostrowski, E. (1994). *Small Data Sets*. Chapman & Hall, London, Seite 320.

### **Abgabetermin**

Abgabe bis spätestens **Montag, den 26. Juli 2010**, ins Fach von Frau Müller.