

Dr. Gert Hullen, Bundesinstitut für
Bevölkerungsforschung beim
Statistischen Bundesamt, Wiesbaden

Survivor-Analyse und Ereignisanalyse

Sterbetafel, Kaplan-Meier-Schätzung und Cox-
Regression mit SPSS - Eine kritische Würdigung

SPSS 2. Academic Convention, Marburg, 3.-4. Juni 2003

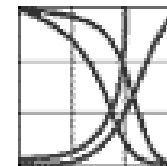
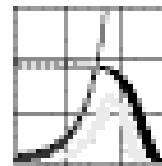
Überlebenszeitanalyse mit SPSS

- Sterbetafelanalyse (life table analysis)
- Survivorfunktion mit Kaplan-Meier-Schätzern
- Ereignisanalyse (event history analysis) mit Cox-Regression
- Prüfung mit Version SPSS 10

Anwendungen

Schätzung der Wartezeit auf den Eintritt eines Ereignisses sowie
Schätzung der Einflüsse auf die Wartezeit in

- Medizin
- Demographie (Versicherungsmathematik, ...)
- business demography (Glühbirnen, Gebrauchtwagen, ...)



Besonderheiten

- zensierte Daten
- parametrische Schätzungen: exponentiell, Weibull, Sichel
- nicht-parametrische Schätzungen: proportionale Hazards, Cox-Regression

Untersuchungsobjekte mit zensierten Daten

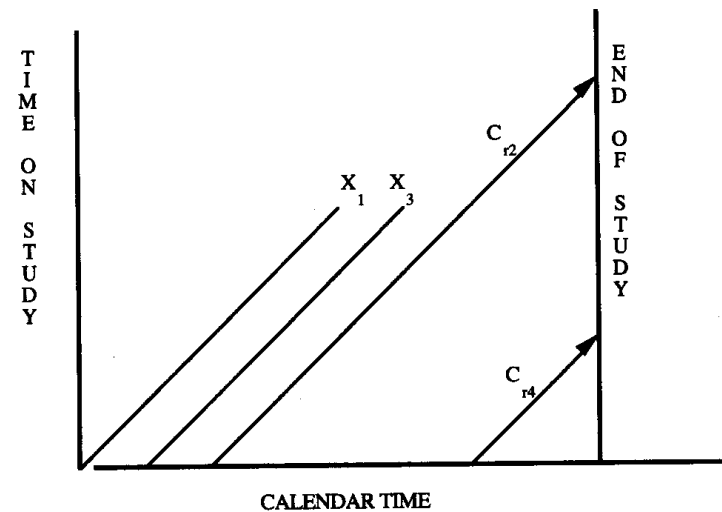


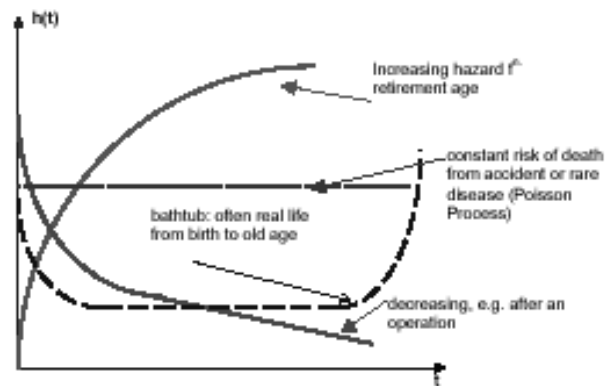
Figure 3.5 *Lexis diagram for generalized Type I censoring in Figure 3.3*

Aus: Klein, John P.; Moeschberger, Melvin L: Survival Analysis. Techniques for Censored and Truncated Data. 2. Aufl. New York: Springer, 68

Hazardfunktionen

Single Sample Models

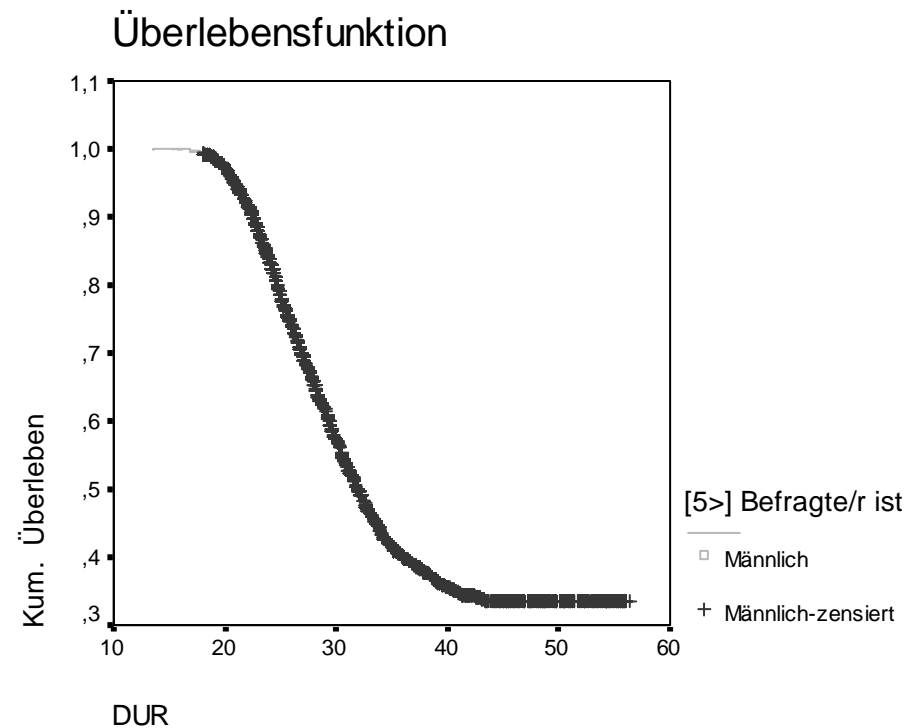
Hazard functions:



23

Kaplan-Meier Schätzung: Eingabe und Plot

```
USE ALL.  
KM  
dur by F51 /STATUS=des(1)  
/PRINT TABLE MEAN /PLOT.
```



Kaplan-Meier Schätzungen: Tabellen-Ausgabe

Survival Analysis for DUR

Factor F51 = Männlich

| Time | Status | Cumulative Survival | Standard Error | Cumulative Events | Number Remaining |
|------|--------|------------------------|-------------------|----------------------|---------------------|
| 14 | 1 | ,9997 | ,0003 | 1 | 3652 |
| 15 | 1 | ,9995 | ,0004 | 2 | 3651 |
| 16 | 1 | ,9992 | ,0005 | 3 | 3650 |
| 16 | 1 | ,9989 | ,0005 | 4 | 3649 |
| 16 | 1 | ,9986 | ,0006 | 5 | 3648 |
| 17 | 1 | ,9984 | ,0007 | 6 | 3647 |
| 17 | 1 | ,9981 | ,0007 | 7 | 3646 |
| 17 | 1 | ,9978 | ,0008 | 8 | 3645 |
| 17 | 1 | | | 9 | 3644 |
| 17 | 1 | ,9973 | ,0009 | 10 | 3643 |
| 17 | 1 | ,9970 | ,0009 | 11 | 3642 |
| 17 | 1 | | | 12 | 3641 |
| 17 | 1 | ,9964 | ,0010 | 13 | 3640 |
| 17 | 1 | ,9962 | ,0010 | 14 | 3639 |
| 18 | 1 | ,9959 | ,0011 | 15 | 3638 |
| 18 | 1 | ,9956 | ,0011 | 16 | 3637 |
| 18 | 1 | | | 17 | 3636 |

Kaplan-Meier-Schätzung

"Die Kaplan-Meier-Methode zur Ereignisdatenanalyse eignet sich nur für kleine Fallzahlen, wenn die Einteilung der Überlebenszeit in Intervalle nicht möglich ist. Bei großen Fallzahlen wird die Textausgabe zu umfangreich, da jeder Fall dokumentiert wird." (Bühl, Zöfel 2002, S. 559)

Ties

Cox-Regression: Syntax

```
[TIME PROGRAM]
[commands to compute time dependent covariates ]
[CLEAR TIME PROGRAM]
COXREG [/VARIABLES=] depvar [ WITH indep_varlist]
  /STATUS=varname [EVENT](valuelist) [LOST(valuelist)]
[/STRATA=varname]
[/CATEGORICAL=varlist]
[/CONTRASTS (indep_cat_var)={DEVIATION (refcat)}]
  {SIMPLE (refcat)}
  {DIFFERENCE}
  {HELMERT}
  {REPEATED}
  {POLYNOMIAL(metric)}
  {SPECIAL (matrix)}
  {INDICATOR (refcat)}
[/ [METHOD= {ENTER **}    [{mthd_varlist}]
  {BSTEP [{COND*}] {ALL    }
  {LR}
  {WALD}
  {FSTEP [{COND*}
  {LR}
  {WALD}
[/MISSING={EXCLUDE **}]
  {INCLUDE}
[/PRINT={[[ALL][SUMMARY][BASELINE][CORR][DEFAULT**]]
  [CI({95*})]
  {n}
```

```
[/CRITERIA={BCON} ({1e-7**})]
  {PCON} {eps }
[/ITERATE ({20**})]
  {n }
[/LCON ({1e-5**})]
  {pct }
[/PIN ({0.05 **})]
  {eps }
[/POUT ({0.1 **}) ]
  {eps }
[/PLOT={NONE**} [SURVIVAL] [HAZARD] [LML] [OMS]]
[/PATTERN=[varname (value) ...] [BY varname]]
[/OUTFILE=[ COEFF (file)] [ TABLE (file)]]
[/SAVE=[tempvar [(newvarname)], ...]
[/EXTERNAL]
* Default if file is absent.
** Default if subcommand is omitted
```

Cox-Regression (1)

```
*
//////////////////////////////////////////////////////////////////// Makro.
DEFINE !COXREGR      (           !POS !TOKENS(1)           /!POS !TOKENS(1)
SUBTITLE             !1.
USE ALL.
FILTER BY           !2.
COXREG              dur /STATUS=des(1) /STRATA=kohor_10
                    /METHOD=ENTER nel verh
                    /PLOT SURVIVAL HAZARD
                    /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) .

!ENDDDEFINE.
```

Auswertung der Fallverarbeitung

| | | N | Prozent |
|------------------------------|-------------------------------------------------------------|-------------|---------------|
| Für Analyse verfügbare Fälle | Ereignis ^a | 2992 | 67,4% |
| | Zensiert | 1446 | 32,6% |
| | Insgesamt | 4438 | 100,0% |
| Nicht verwendete Fälle | Fälle mit fehlenden Werten | 0 | ,0% |
| | Fälle mit nicht positiver Zeit | 0 | ,0% |
| | Zensierte Fälle vor dem frühesten Ereignis in einer Schicht | 0 | ,0% |
| | Insgesamt | 0 | ,0% |
| Insgesamt | | 4438 | 100,0% |

a. Abhängige Variable: DUR

Status der Schicht^a

| Stratum | Ereignis | Zensiert | Prozent Zensierte |
|------------------|-------------|-------------|-------------------|
| 40 | 417 | 81 | 16,3% |
| 50 | 863 | 179 | 17,2% |
| 60 | 1256 | 274 | 17,9% |
| 70 | 429 | 625 | 59,3% |
| 80 | 27 | 287 | 91,4% |
| Insgesamt | 2992 | 1446 | 32,6% |

a. Die Schichtenvariable lautet: KOHOR_10

Cox-Regression (2)

```

* ////////////////////////////////////////////////////////////////////////////////////////////////////////////////// Makro.
DEFINE !COXREGR      (          !POS !TOKENS(1)          /!POS !TOKENS(1)

SUBTITLE              !1.
USE ALL.
FILTER BY             !2.
COXREG                dur /STATUS=des(1) /STRATA=kohor_10
                     /METHOD=ENTER nel verh
                     /PLOT SURVIVAL HAZARD
                     /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) .

!ENDDDEFINE.

```

Variablen in der Gleichung

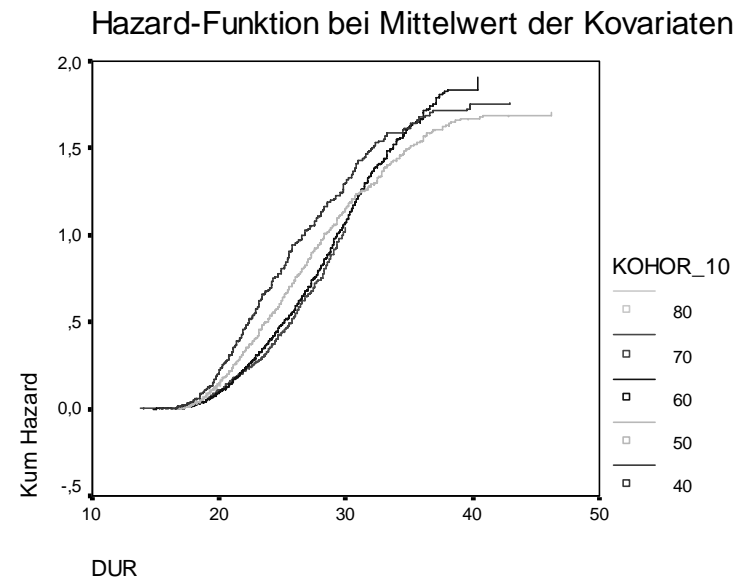
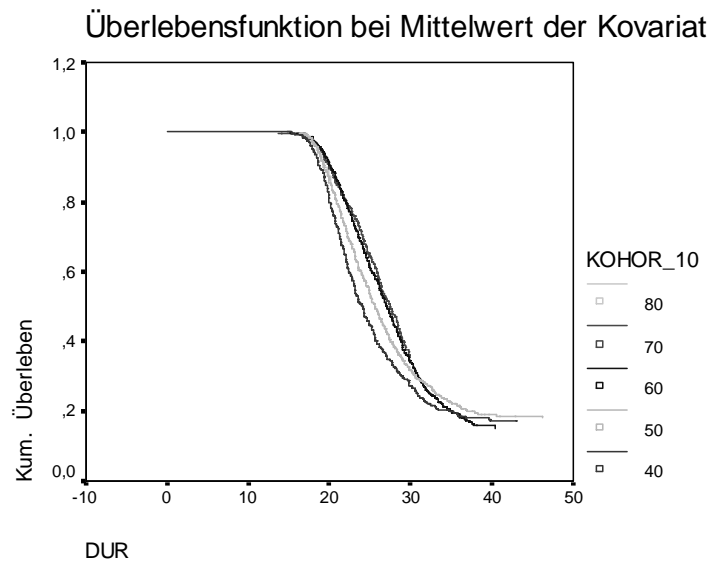
| | B | SE | Wald | df | Signifikanz | Exp(B) |
|------|------|------|--------|----|-------------|--------|
| NEL | ,168 | ,065 | 6,719 | 1 | ,010 | 1,183 |
| VERH | ,414 | ,043 | 90,760 | 1 | ,000 | 1,512 |

Cox-Regression (3)

```
* //////////////////////////////////////// Makro.
DEFINE !COXREGR      (          !POS !TOKENS(1)          /!POS !TOKENS(1)

SUBTITLE              !1.
USE ALL.
FILTER BY              !2.
COXREG                dur /STATUS=des(1) /STRATA=kohor_10
                      /METHOD=ENTER nel verh
                      /PLOT SURVIVAL HAZARD
                      /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) .

!ENDDFINE.
```



Andere Programme

- SAS: Prozeduren LIFETEST (Kaplan-Maier und Sterbetafel), PHREG (Proportional Hazards Regression) mit Optionen für Ties (TIES = {BRESLOW, DISCRETE, EFRON, EXACT})
- Stata
- TDA
- Tabellenkalkulationsprogramme (SISA)

Hazardraten-Modelle (TDA)

##rate model numbers: The following models can be selected with the rate command:

- 1 Cox model
- 2 Exponential model
- 3 Piecewise constant exponential model
- 4 Polynomial rates, I
- 5 Polynomial rates, II
- 6 Gompertz-Makeham models
- 7 Weibull model
- 8 Sickle model
- 9 Log-logistic model, type I
- 10 Log-logistic model, type II
- 12 Log-normal model
- 13 Generalized gamma model
- 14 Inverse Gaussian model
- 16 Piecewise constant exponential model with period-specific effects
- 20 Logistic regression model
- 21 Complementary log-log model

Weiterentwicklungen der Ereignisanalyse

- mehrere parametrische Modelle für Hazardfunktionen (TDA von Rohwer)
- unbeobachtete Heterogenität (Lillard u.a.; Programmpaket aML (Multilevel Multiprocess Modeling) von Lillard/Panis, Rand Corporation; Jan Hoem, Rostock)
- SURV2 und SURV3, Finnish Cancer Registry for relative survival analysis; Paul Dickman, Dept. of Med. Epid., Karolinska Institute, Sweden, paul.dickman@mep.ki.se; Timo Hakulinen, Helsinki, timo.hakulinen@cancer.fi)

Weiterentwicklungen der Graphik

- doppeltes Balkendiagramm für Bevölkerungsstrukturen („Bevölkerungspyramide“)

