APPENDIX 04 - GLIM Macros for survival analysis

In this appendix, the author, purpose, formal arguments, macro arguments, scalar arguments, subsidiary macros, uses, output, further information and references for each of GLIM macros WEIBULL and RESPLOT are taken from GLIM Macro library Release 1.0, Macro library description (Filename PACKAGE\GLIM\MACLIB.DOC), January 1985. GLIM Programs 2-6 provide examples on the use of these macros and are given in APPENDIX 03.

Macro WEIBULL in subfile WEIBULL

Author

B. J. Francis, Centre for Applied Statistics,
Cartmel College, University of Lancaster, U.K.

Purpose

The macro WEIBULL allows the fitting of a specified regression model to survival data using either the exponential or Weibull distributions. The survival data may be partially right-censored, or can be entirely uncensored.

Estimation:
The likelihood for the exponential or Weibull distributions can be expressed as a Poisson likelihood, with a log-linear model for the Poisson mean corresponding to a log-linear model for the hazard function.
If the shape parameter is fixed or known, then this procedure corresponds to using the censor variate as y-variate, specifying the log of the vector of survival times as an offset, and fitting a Poisson model with log link. Thus, for the exponential distribution, a single Poisson fit in GLIM is all that is needed. The Weibull distribution requires an iterative procedure - fixing the shape parameter, fitting the model and re-estimating the shape parameter - until convergence is achieved.

Full computational details are given in Aitkin and Clayton (1980).
The hazard function in t has the following form:

```
Exponential h(t)=exp(B'X)
Weibull     h(t)=a*(t**(a-1))*exp(B'X)
```

where a is the shape (index) parameter
B is the vector of parameters for a specified model
X is the vector of explanatory variables.
(and * and ** have their usual meanings of operators)

Maximum likelihood estimates of a and B are given by the macro. The value of -2 ln(maximised likelihood) or the 'deviance' is also given.
Standard errors of B are displayed with the estimate of B. For the exponential distribution, these standard errors are correct, but for the Weibull, the standard errors of B are slightly underestimated, as they do not allow for the fact that a is an estimated parameter, and not fixed. No standard error for a is given.
Macro WEIBULL may be called as many times as is necessary, changing the model formula specified in macro MODEL (and other arguments if necessary) as appropriate.

APPENDIX 04 - GLIM Macros for survival analysis

Formal arguments

%1 The variate containing the survival times, some of which may be right-censored.

%2 An indicator or censor variate. The elements of the variate should be 1 if the corresponding survival time is uncensored, and 0 if the corresponding survival time is censored.

%3 (optional) A scalar. Determines whether the Weibull or exponential distribution is fitted. If the argument is not set or set to a non-zero scalar, then the Weibull distribution is fitted, using an iterative fitting procedure. The starting value of the shape parameter is taken from %A. If %A is zero or negative, then a starting value of 1 is used. If the third argument is set to a scalar equal to zero, then the exponential distribution is fitted. The value of the shape parameter is taken to be 1, and only one iteration of the iterative procedure is carried out.

Macro arguments

MODEL (obligatory) The model formula to be fitted needs to be stored in a macro called MODEL.
CONV (obligatory) The convergence criterion for the change in the estimate of the shape parameter. The default is .001.
CYCLE (optional) Should be set to contain either the GLIM directive $CYCLE$, or $RECYCLE$. Determines whether the underlying GLIM IRLS procedure recycles from the previous estimates or not. Setting the contents of macro CYCLE to $RECYCLE$ should speed up the fitting of complex models, but persistent recycling in a macro which uses an iterative fitting procedure may occasionally cause divergence of the deviance. The default is $CYCLE$.
DISP (optional) Determines the $DISPLAY options to be used after convergence has occurred. DISP should contain valid display option letters. The default contents of DISP is E.

Scalar arguments

%A The starting value for the estimate of the shape parameter. Used only when the Weibull distribution is being fitted. If %A is zero or negative, then the default starting value of 1 is used.
%W Sets the maximum number of iterations carried out by the macro. If set to a scalar less than or equal to 0, the default setting of 15 is used.

Uses

Within the macro, scalars %A, %D, %F, and %W and vectors OFV_ and LGT_ are used.

Subsidiary macros

Macro WEIBULL calls or uses macros MODEL, MOD1, MOD2, WAR1, MESS, DISP, CONV and CYCLE. By default, DISP is set to E, CONV is set to .001, CYCLE is set to $CYCLE$ and MODEL is undefined.

Output

At each iteration, the current deviance, estimate of the scale parameter and the number of degrees of freedom are displayed. On convergence, or after %W (15) iterations the parameter estimates together with their standard errors are displayed. Note that the standard errors may be slightly underestimated.
On normal exit, the vector %FV contains the scaled residuals from the fit. These are not independent but will have approximately standard exponential distributions.
The following scalars contain useful information:

- %A Estimate of shape parameter
- %D 'Deviance' or -2 ln(maximised likelihood)
- %F No. of degrees of freedom
- %W Current setting of maximum number of iterations

Further information

1. The macro takes NO account of any $OFFSET variate or $SCALE parameter settings. If prior weights are set, the macro will display a warning message and unset them.
2. As the macro may carry out many $FITs, the $OUTPUT is switched off for part of the execution. If the macro fails, for instance, because of an incorrectly specified model, or if break-in is used, the output may still be switched off when control returns to the user at the terminal.
3. On normal exit from the macro:
   The $ERROR setting is Poisson, with $LOG link.
   The GLIM $YVARiate is set to the censor variate specified in formal argument %2.
   The model formula is set to the model formula specified in macro MODEL.
   The $OFFSET is unset.
   The $WEIGHT is unset.
   Any settings of the $ERROR, $LINK, $WEIGHT, $OFFSET and $CYCLE made before the call of macro WEIBULL are lost. The display is inhibited.

References

"Aitkin, M. and Clayton, D. (1980), 'The fitting of Exponential, Weibull and Extreme Value distributions to complex censored survival data using GLIM', Applied Statist, Vol. 29, pp 156-163"

Macro RESPLOT in subfile WEIBULL

Author

M. A. Aitkin, Centre for Applied Statistics,
Cartmel College, University of Lancaster, U.K.

Purpose

After a fit of a Weibull model using macro WEIBULL, the macro RESPLOT displays two quantile-quantile residual plots. The macro uses the vector %FV, which on exit from macro WEIBULL contains the scaled residuals, which will have approximately standard exponential distributions.

Formal arguments
%1 An indicator or censor variate. The elements of the variate should be 1 if the corresponding element of %FV is uncensored, and 0 if the corresponding element of %FV is censored.

Macro arguments
-None-

Uses
Within the macro, vectors WK1_, WK2_, WK3_ and WK4_ are used.

Subsidiary macros
-None-

Output
The macro displays two quantile-quantile residual plots. The first plot is uncorrected for heterogeneity of the variance of the scaled residuals. The second is variance stabilised. If the probability model holds, both plots should give straight lines with slopes of unity.

Further information
1. Macro RESFLOT does not destroy the contents of either %1 or %FV.

References
"Aitkin, M. and Clayton, D. (1980), 'The fitting of Exponential, Weibull and Extreme Value distributions to complex censored survival data using GLIM', Appl. Statist, Vol. 29, pp 156-163"

Listing (From $input %plc 80 weib$)

```
!
!
!-------------------------------------------------------------------------------
!  Author: B. J. Francis, Centre for Applied Statistics,
!                         University of Lancaster, U. K.
!  Version 1.0 GLIM 3.77 January 1985
!  Main macros:
!       WEIBULL  Fits the exponential or Weibull transformation to survival
!                data. Survival times may be right-censored.
!       RESPLOT  produces quantile-quantile plots of the scaled residuals
!       For macro WEIBULL:
!              Formal arguments:
!                 %1       (obligatory) Variate containing the survival
!                          times, some of which may be right censored.
!                 %2       (obligatory) Indicator or censor variate. Element
!                          take the value 1 if uncensored and 0 if censored.
!                 %3       (optional) A scalar. If not set, or set equal to
!                          a non-zero scalar, then the Weibull distribution
!                          is fitted. If set equal to a zero scalar, then the
```

```
!                        exponential distribution is fitted.
!               Macro arguments:
!                  MODEL    The model formula requested (no default)
!                  CONV     (optional) The convergence criterion(default .001)
!                  DISP     (optional) The $DISP options used after
!                             convergence (default E)
!                  CYCLE    (optional) Takes the values $CYC$ or $RECY$
!                             determines whether $recycling is used for the
!                             iterative fitting of the Weibull distribution.
!                             (Default  $CYCLE$)
!               Scalar arguments:
!                  %A       Starting value for the shape parameter (default 1)
!                  %W       Maximum number of iterations (default 15)
!       For macro RESPLOTS:
!               Formal arguments:
!                  %1       (obligatory) Indicator or censor variate. Element
!                           take the value 1 if uncensored and 0 if censored.
!       Output:
!             WEIBULL:Displays for each iteration
!                     the deviance and the number of degrees of freedom
!                     On convergence, or after %W (15) iterations, displays
!                     by default the estimates of the parameters and s.errors
!             RESPLOT:Displays two Q-Q residual plots, the second plot
!                     being variance stabilised. Can only be called after a
!                     call to WEIBULL
!       Example of use:
!                  $mac model a+b+c $endm
!                  $mac disp et $endm        !  optional
!                  $use weibull t c$
!                  ...macro prompts user for grid values...
!-------------------------------------------------------------------------------
!
$MAC MESS $PRI '-- Standard errors of estimates'!
  given below are underestimated'!
$$E
!
$MAC MOD1 $PRI 'Exponential fit' $$ENDM
!
$MAC MOD2 $PRI 'Weibull fit' $$ENDM
!
$MAC WAR1 $PRI '--Weights not available. No weights used in fit. $WEI $$ENDM
!
$MAC CONV .001 $ENDM
!
$MAC DISP E $ENDM
!
$MAC CYCLE $CYCLE $ENDM
!
$MAC WB1!
$CAL %Z6=%Z6-1: %Z6=%GT(%Z6,0)*%Z6!                update iteration counter
$OUT $CAL OFV_=%A*%1!                              update offset variate
$USE CYCLE$FIT #MODEL$OUT %Z4!                     fit model suppressing usual output
$CAL %D=%Z2 - 2*%CU(%Z2*%LOG(%A*%FV)-%FV)!         calculate deviance
: %F=%DF-%NE(%A,1)!                                calculate correct DF
$PRI *5 %D %A *1 %F,6$!                            and print them out
```

```
$CAL %Z9=%EQ(%Z3,0) $EX %Z9!                    only one iteration for exponential
$CAL %Z8=%CU(%1*(%FV-%2))
: %Z8=0.5*(%A-%Z1/%Z8)!                         new increment for shape parameter
: %Z8=%IF(%LT(%Z8,-5),-5,%Z8)!                  should not be less than -5
: %Z9=%LE(-#CONV,%Z8)*%GE(#CONV,%Z8)!
$EXI %Z9!                                       test for convergence
$CAL %Z9=%EQ(%A,1)$SW %Z9 MOD2!                 if end of last iteration, print MOD2
$CAL %A=%A-%Z8 $$ENDM!                          update shape parameter
!
$MAC WB2 $CAL %Z3=(%1/=0)$$END
!
$MAC WEIBULL!
$DEL OFV_ LGT_!
$PR *M 20 '-- Model is   MODEL $
$CAL LGT_=%LOG(%1+0.5*%EQ(%1,0))!               lgt_ is transformed survival time
 : %Z1=%CU(%2)!
 : %Z2=2*%CU(LGT_*%2)!
 : %Z4=%COC!
 : %Z6=%IF(%GT(%W,0),%W,15)!                    max no. of iterations
$YVA %2 $ERR P$CYC!
$SWI %PWF WAR1$OFF OFV_!                        switch weights off if set
$CAL %Z3=1$ARG WB2 %3$SWI %A3 WB2$!             Check to see if third argument set
$CAL %A=%IF(%NE(%Z3,0)*%GT(%A,0),%A,1)!         initial value of snape parameter
$CA %Z9=%EQ(%A,1)+1$SW %Z9 MOD2 MOD1$!          if %a=1 then print mod1 else mod2
$pr :   Deviance   shape     df'!               Print headings
 :                parameter'!
$ARG WB1 LGT_ %2$WHI %Z6 WB1!                   use wb1 until convergence
$SWI %Z3 MESS!
$DIS #DISP$CYC$OFF!
$DEL LGT_ OFV_$!
$$END
!
!
$M RESPLOTS!
$DEL WK1_ WK2_ WK3_ WK4_!
$SOR WK3_ %1 %FV : WK4_ %FV!
$CAL WK1_=(%CU(1)-%EXP(-WK4_)*%CU((1-WK3_)*%EXP(WK4_)))/(%NU+1)!
 : WK2_= -%LOG(1-WK1_)!
$PRI 'Residual plot'$
$PLO WK2_ WK4_ '+'$!
$CAL WK1_=%ANG(1-WK1_) : WK4_=%ANG(%EXP(-WK4_))!
$PRI 'Variance stabilised residual plot'
$PLO WK1_ WK4_ '+'$!
$DEL WK1_ WK2_ WK3_ WK4_$E
!
$M DEL $DEL RESP WB1 WB2 WEIB CONV DISP CYCLE MODEL MOD1 MOD2 MESS WAR1 $ENDM
!
$RETURN
```