**SAS Instructions**
**Parametric regression with LIFEREG**

Suppose that you have the following regression model for some continuous, positive random variable $T$ and a vector of explanatory variables (covariates) $\boldsymbol{x}$:

$$\ln T = \beta_0 + \boldsymbol{\beta}'\mathbf{x} + \sigma W,$$

where $W$ is a continuous random variable on the real line with a distribution that does not involve any unknown parameters. Choices of distributions of $T$ that can be fitted with PROC LIFEREG: exponential, Weibull, log-normal, log-logistic, gamma, generalized gamma. PROC LIFEREG calls $\hat{\alpha}_0$ "Intercept", $\sigma$ "scale" and the other $\hat{\alpha}$ 's by the name of the corresponding explanatory variable.

**General syntax of PROC LIFEREG**

```
PROC LIFEREG DATA=dataset_name COVOUT NOPRINT OUTEST=dataset_name;
      MODEL time_variable = covariates / options;
      CLASS list_variables;
      WEIGHT variable;
      OUTPUT dataset_name options;
      BY list_variables;
RUN;
```

OUTEST: creates dataset with MLEs, maximized likelihood, covariance matrix.

Options in MODEL statement:

- DISTRIBUTION= (any of WEIBULL, EXPONENTIAL, LNORMAL, LLOGISTIC, GAMMA, NORMAL, LOGISTIC)
- NOLOG
- INTERCEPT= (initialize value of intercept)
- NOINT (value of intercept held fixed … uses value 0 if INTERCEPT= not used)
- INITIAL= (initialize values of regression coefficients)
- SCALE= (initialize value of scale)
- NOSCALE (value of scale held fixed … uses value 1 if SCALE= not used)

CLASS: all variables listed are treated as categorical/classification rather than continuous variables.

WEIGHT, OUTPUT, BY: see SAS help files for more details.

**Example**

Let us fit a log-logistic regression model to the dataset from Assignment 1, Question 11, assuming no truncation …

```
DATA a1q11;
  INFILE "a1q11.dat" FIRSTOBS=2;
  INPUT id$ ti delta drug$;
  IF drug="D" THEN treat=1;
    ELSE treat=0;
RUN;

PROC LIFEREG DATA=a1q11;
  MODEL ti*delta(0)=treat / DIST=LLOGISTIC;
RUN;

DATA placebo;
  SET a1q11;
  IF drug="P";
RUN;

PROC LIFEREG DATA=placebo;
  MODEL ti*delta(0)= / DIST=LLOGISTIC;
RUN;

DATA treatmen;
  SET a1q11;
  IF drug="D";
RUN;

PROC LIFEREG DATA=treatmen;
  MODEL ti*delta(0)= / DIST=LLOGISTIC;
RUN;
```

Here are excerpts of the output:

```
Name of Distribution            LLOGISTC
Log Likelihood              -10.69934896

                      Analysis of Parameter Estimates
                                Standard
Variable   DF   Estimate     Error Chi-Square Pr > ChiSq Label

Intercept  1    3.31191    0.12384   715.1644     <.0001 Intercept
treat      1    0.14775    0.16264     0.8252     0.3637
Scale      1    0.21263    0.03796                       Logistic scale
```

```
Name of Distribution            LLOGISTC
Log Likelihood              -4.203140694

                      Analysis of Parameter Estimates

                                Standard
Variable   DF   Estimate    Error Chi-Square Pr > ChiSq Label

Intercept  1    3.30021    0.12087   745.4540     <.0001 Intercept
Scale      1    0.19564    0.05392                      Logistic scale


Name of Distribution             LLOGISTC
Log Likelihood                -6.42326946

                       Analysis of Parameter Estimates

                                Standard
Variable   DF   Estimate    Error Chi-Square Pr > ChiSq Label

Intercept  1    3.46381    0.11343   932.5384     <.0001 Intercept
Scale      1    0.22476    0.05254                     Logistic scale
```

If we wish to test whether it is reasonable to assume that the scale parameter does not depend on the value of the covariate, we simply fit one model with two different scale parameters (the 2nd and 3rd PROC LIFEREGs) and one model with a common scale parameter (the 1st PROC LIFEREG) and we test if the model reduction is appropriate using a likelihood ratio test. Here, the likelihood ratio statistic has value 2*(-6.42 + -4.20 - -10.70) = 0.16. On one degree of freedom, this gives us a p-value of 0.69.