Read Bull and Spiegelhalter TUTORIAL IN BIOSTATISTICS: SURVIVAL ANALYSIS IN OBSERVATIONAL STUDIES (1997).

B.   The following questions refer to the same data set used in Homework 5, contained in file **lungca**. Survival time up to 90 days, the dependent variable, is the next column (X_9) in the data set. Use the status at the end of 90 days (Y) as the censoring variable, with censoring being a value of 0, and variables $X_1, X_2, \cdots, X_8$ as the independent variables to answer the following questions.

1.   Use **PROC LIFETEST** to obtain plots of the survival distribution for each treatment group. Does treatment appear to affect survival? (If you are familiar with the logrank test, interpret it.) Program as follows:

           PROC LIFETEST   PLOTS = (S,LLS);
           TIME DTHSUR*STATUS(0);
           STRATA TREAT;

       where DTHSUR is your label for X_9, STATUS is Y, and TREAT is X_1.

2.   Perform Cox's proportional hazards regression

       a.   What is your conclusion regarding the central question of the study? Do the two treatment groups differ with respect to 90-day survival after adjustment for confounders?

       b.   What is the estimated risk of dying in the test group compared to the standard group? Include a 95% confidence interval. Show your work. Make a statement regarding the effect of treatment on survival with these results.

       c.   Is there confounding in examining this association? Provide evidence.

       d.   What other variables are associated with 90-day survival? Provide evidence.

       e.   How do these results compare with those of the logistic regression analyses? Justify your responses.

       f.   Is the proportional hazards assumption satisfied for the treatment groups? Is it satisfied for the other covariates? Explain your answers.

3.   These individuals were in reality followed much longer than 90 days, about 3 years. The real follow-up data are given in the last two fields on the data set: X_{10} is the survival time and X_{11} is the status at the end of the study. Repeat the proportional hazards analysis on the real survival time and evaluate whether treatment had an effect. What are your conclusions? Are the assumptions of the proportional hazards model met?