Math 473 HW 8 Spring 2023. Due Friday, April 7.

1) 3.2: Suppose that the lifetimes of a certain brand of lightbulb follow an exponential (λ) distribution. 20 light bulbs are tested for 1000 hours. The failure times are below.

71, 88, 254, 339, 372, 403, 498, 499, 593, 774, 935,

$1000+, 1000+, 1000+, 1000+, 1000+, 1000+, 1000+, 1000+, 1000+$

a) Find $\hat{\lambda}$.

b) Find a 95% CI for λ.

2) 3.15: This problem considers the ovarian data from Collett (2003, p. 344-346).

a) Obtain the SAS program for 3.15 from (http://parker.ad.siu.edu/Olive/survhw.txt). Print the output. (It should be very similar to that on exam 3 review p. 2.)

b) Find the ESP if *age* = 40 and *treat 1* = 1. (Comment: treatment takes on 2 levels so only one indicator is needed. SAS output includes a 2nd indicator *treat 2* but its coefficient is $\hat{\beta}_3 = 0$ and hence can be ignored. In general if the category takes on J levels, SAS will give nonzero output for the first J − 1 levels and a line of 0s for the Jth level. This means level J was omitted and the line of 0s should be ignored.) Hint: Should get a number near −3.7.

c) Give a 95% CI for β_1 corresponding to age from output and the CI using the formula.

d) Give a 95% CI for β_2 corresponding to treat 1 from output and the CI using the formula.

e) If the model statement in the SAS program is changed to

model survtime*status(0)=;

then the null model is fit and the SAS output says Log Likelihood −29.76723997.

Test $\boldsymbol{\beta} = \mathbf{0}$ with the LR test.

(Hint: The full model log likelihood log(L) = −20.56313339. Want −2 log(L) for both the full and null models for the LR test.)

f) Suppose the reduced model does not include *treat*. Then SAS output says Log Likelihood −21.7830. Test whether the reduced model is good.

(Hint: The log likelihood for the full model is log(L) = −20.56313339. Want −2 log(L) for the full and reduced models for the change in LR test.)

3) 3.8: This problem considers the same data as in 2) but uses *R*.

a) Obtain the *R* code for 3.8 from (http://parker.ad.siu.edu/Olive/survhw.txt). Click on the left screen then hit *Enter*. Copy and paste both the output (it should be very similar to that on p. 3 of the exam 3 review) into *Word*. Also copy and paste the plot into *Word*.

b) The plot is a log censored response plot. The top line is the identity line and the bottom line the least squares line. Is the slope of the least squares line near 1?

4) 3.9: Use the source commands near the top of (http://parker.ad.siu.edu/Olive/survhw.txt) to get `survpack` into R. The programs `phdata`, `weyp` and `wregsim` will be used.

The program `wregsim` generates Weibull proportional hazards regression data with baseline hazard function $h_0(t) = kt^{k-1}$.

a) Type the command *wregsim(gam=1)* 5 times (or use the "up arrow" after typing the command once). This gives 5 simulated Weibull regression data sets with $k = 1$. Hence the Weibull regression is also an exponential regression. Include the last plot in *Word*.

b) Type the command *wregsim(gam=5)* 5 times. To judge linearity, ignore the cases on the bottom of the plot with low density (points with log(time) less than -2). (These tend to be censored cases because time $Y = W^{1/k}$ where $W \sim EXP(\lambda = \exp(SP))$ where $E(W) = 1/\lambda$. $Z \sim EXP(.1)$ has mean 10 and if $Z_i < Y_i$ then Z_i is usually very small.) Do the plots seem linear ignoring the cases on the bottom of the plot? Do not include the plot.

c) Type the command *wregsim(gam=0.5)* 5 times. (Now censored cases tend to be large because time $Y = W^{1/k} = W^2$ where $W \sim EXP(\lambda)$. $Z \sim EXP(.1)$ has mean 10 and if $Z_i < Y_i$ then $Y_i > 10$, usually.) Do the plots seem linear (ignoring cases on the bottom of the plot)? (The plot is linear if it is roughly box shaped or ellipsoidal, possibly ignoring some of the points with $\log(time) < -9$. Since the error distribution is left skewed, most of the plotted points will fall below the identity line, even if the plot is linear.) Do not include the plot.