Math 473 HW9 Spring 2021. Due Friday, April 9.

Quiz 8 covers HW7 and HW8.

1) 3.10: This problem considers the ovarian data from Collett (2003, p. 189, 344-346).

a) Obtain the *R* code for 3.10a from (http://parker.ad.siu.edu/Olive/survhw.txt). Copy and paste the plot into *Word*.

b) Now obtain the R code for 3.10b and put the plot into *Word*.

c) Can the Exponential regression model be used or should the more complicated Weibull regression model be used?

d) Examine the cumulative hazard plot Figure 7.1 below. Which plot has tighter clustering about the identity line, the EE plot from a) or Figure 7.1?

2) 3.11: Copy and paste the two source commands from the top of (http://parker.ad.siu.edu/Olive/survhw.txt) to get programs `phdata` and `wregsim2` into *R*.

Make the left window small by moving the cursor to the lower right corner of the window, then hold the right mouse button down and drag the window to the left.

The program `wregsim2` generates Weibull proportional hazards regression data with baseline hazard function $h_0(t) = \gamma t^{\gamma-1}$.

a) Type the command *wregsim2(n=10, gam=1)* 5 times (or use the "up arrow" after typing the command once). This gives 5 simulated Weibull regression data sets with $\gamma = 1$. Increase n by 10 until the plotted points cluster tightly about the identity line in at least 4 out of 5 times. How big is n?

b) Type the command *wregsim2(n =10, gam=5)* 5 times. Increase n by 10 until the plotted points cluster tightly about the identity line in at least 4 out of 5 times. How big is n?

c) Type the command *wregsim2(n=10, gam=0.5)* 5 times. Increase n by 10 until the plotted points cluster tightly about the identity line in at least 4 out of 5 times. How big is n?

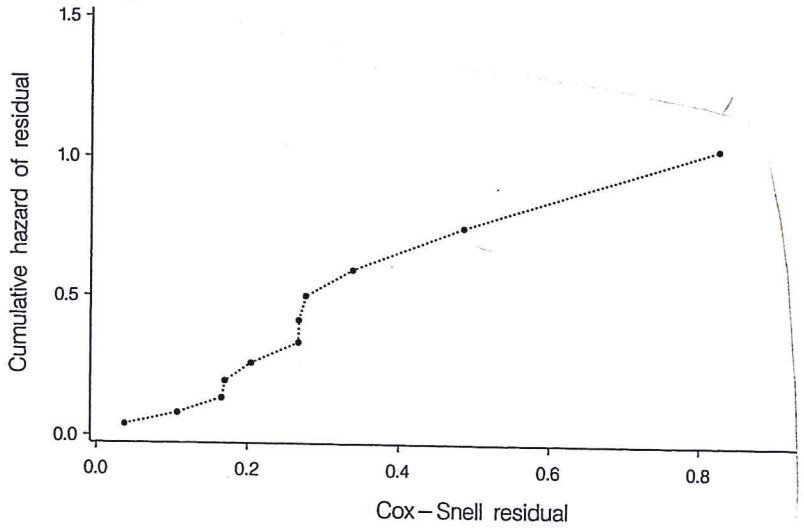

Figure 7.1 *Cumulative hazard plot of the Cox-Snell residuals.*

3) 3.12: If necessary copy and paste the two source commands as done for problem 3.11 to get programs `phdata` and `wregsim3` into R.

Make the left window small by moving the cursor to the lower right corner of the window, then hold the right mouse button down and drag the window to the left.

The program `wregsim3` generates Weibull proportional hazards regression data with baseline hazard function $h_0(t) = \gamma t^{\gamma-1}$. This is also an AFT model with $\alpha = 0$, $\boldsymbol{\beta}' = -(1/\gamma, ..., 1/\gamma)$ and $\sigma = 1/\gamma$. The program generate 100 Weibull AFT data sets and for each run i computes $\hat{\alpha}_i$, $\hat{\boldsymbol{\beta}}_i$ and $\hat{\sigma}_i$. Then the averages are reported. Want mnint ≈ 0, mncoef $\approx -(1/\gamma, ..., 1/\gamma)$ and mnscale $\approx 1/\gamma$.

a) Make a table (by hand) with headers

```
n   gamma   mnint  mncoef                    mnscale
```

Fill in the table for $n = 20, \gamma = 1; n = 100, \gamma = 1; n = 200, \gamma = 1; n = 20, \gamma = 5; n = 100, \gamma = 5; n = 200, \gamma = 5; n = 20, \gamma = 0.5; n = 100, \gamma = 0.5; n = 200, \gamma = 0.5$ by using the commands *wregsim3(n=20, gam=1)*, ..., *wregsim3(n=200, gam=0.5)*.

b) Are the estimators close to parameters $\alpha, \boldsymbol{\beta}$ and σ for $n = 20$? How about for $n = 100$?

4) 3.13: If necessary copy and paste the two source commands as done for problem 3.2 to get programs `wphsim` and `swhat` into R. Type the command *wphsim(n=999)* to make a slice survival plot based on the WPH survival function. Are the KM curve and Weibull estimated survival function close for the plot in the bottom right corner? Include the plot in *Word*. Recent versions of R may make 3 curves of cirles. The center curve is the KM curve while the 2 outer curves are pointwise CI bands. (When 3 curves of cirles are made, if the plusses are near or within the circles, then the plots suggest that the WPH model is good.)