Math 473 Exam 3 Spring 2017 Name_____

	Value	Std. Error	z	p
(Intercept)	3.74398	8.54299	0.438	0.661
age	-0.36903	0.46004	-0.802	0.422
weight	0.00364	0.00242	1.502	0.133
length	-0.01203	0.03319	-0.362	0.717
Log(scale)	-0.22451	0.21043	-1.067	0.286

Loglik(model)= -101.8 Loglik(intercept only)= -103
Chisq= 2.39 on 3 degrees of freedom, p= 0.49 n = 50

hard way: $x^2(NIF) = [-2(-103)] - [-2(-101.8)] = 2.40$

1) Data from Collett (2003, p. 366) is on black ducks. Y is survival time in days, predictors are *age* (0 hatch year bird, 1 bird aged $\geq$ 1 year), *weight* of duck in grams and *length* of wing in mm. Weibull regression was used.

a) Test $\beta = 0$. $H_0\ \underline{\beta} = \underline{0}$ $H_A\ \underline{\beta} \neq \underline{0}$

$x^2(NIF) = 2.39$

pval = 0.49

fail to reject H_0, there is not a WPH survival relationship between Y and the predictors age, weight, length

b) Test $\beta_2 = 0$. $H_0\ \beta_2 = 0$ $H_A\ \beta_2 \neq 0$

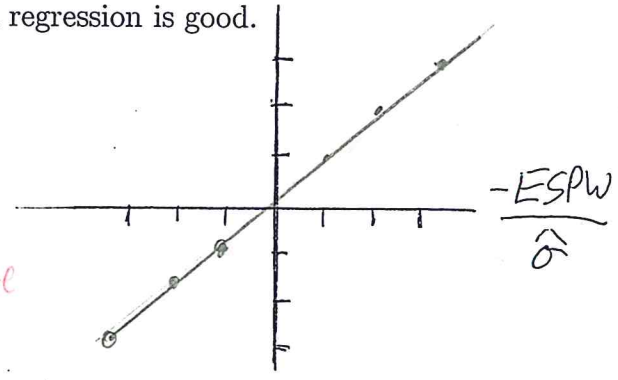
$Z_{02} = 1.502$

pval = 0.133

fail to reject H_0, weight is not needed in the WPH survival model given the other predictors age and length are in the model.

c) Sketch the Weibull EE plot if the Weibull regression is good.

identity line



pts off if they say identity line but plot is poor

27

```
Full Model:     coef exp(coef)  se(coef)     z      p
age         9.05e-03  1.01e+00  1.16e-02  0.78 0.4352
ph.ecog     7.07e-01  2.03e+00  2.23e-01  3.17 0.0015
ph.karno    2.07e-02  1.02e+00  1.13e-02  1.84 0.0663
pat.karno -1.33e-02  9.87e-01  8.05e-03 -1.65 0.0985
meal.cal  -5.27e-06  1.00e+00  2.63e-04 -0.02 0.9840
wt.loss   -1.52e-02  9.85e-01  7.89e-03 -1.93 0.0540

Likelihood ratio test=21.6  on 6 df, p=0.00145

Reduced Model:coef exp(coef) se(coef)     z      p
ph.ecog     0.71138   2.03679  0.22491  3.16 0.0016
ph.karno    0.01895   1.01913  0.01111  1.71 0.0880
pat.karno -0.01375   0.98635  0.00793 -1.73 0.0831
wt.loss   -0.01547   0.98465  0.00781 -1.98 0.0476

Likelihood ratio test=20.9  on 4 df, p=0.000326
```

2) The *R* lung data examines survival of lung cancer patients. A stratified proportional hazards model has been fit with stratification on *sex*. Predictors are *age*, *ph.ecog*: the physician's estimate of the ECOG performance score (0-4), *ph.karn*: the estimate of the Karnofsy score, *pat.karno*: the patient's assessment of their Karnofsy score, *meal.cal*: calories consumed at meals excluding beverages and snacks, and *wt.loss*: weight loss in the last 6 months.

a) Test $\boldsymbol{\beta} = \mathbf{0}$ for the full model. $H_0\ \beta = 0 \quad H_A\ \beta \neq 0$

$X^2(NIF) = 21.6$

$pval = 0.00145$

reject H_0 there is a SPH survival relationship between y and the predictors (age, ..., wt.loss).

b) Test whether the reduced model is good.

H_0 the reduced model is good H_A use the full model

$X^2(RIF) = X^2(NIF) - X^2(NIR) = 21.6 - 20.9 = 0.7$

$pval = P(\chi^2_2 > 0.7) > 0.25$

df	0.25
2	2.77

fail to reject H_0, the reduced model is good

18

$r = \#\text{ uncensored cases} = \sum_{i=1}^{14} \delta_i = 3$

3) Suppose that the 14 failure times (in minutes) given below follow an exponential (λ) distribution.

1, 8, 10, 59+, 72+, 76+, 113+, 117+, 124+, 145+, 149+, 153+, 182+, 320+

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

$$= \frac{r}{\sum T_i} = \frac{3}{1529} = 0.001962$$

6.7.8.21ab

b) Find a 95% CI for λ. $\hat{\lambda} \pm 1.96 \frac{\hat{\lambda}}{\sqrt{r}} =$

$$0.001962 \pm 1.96 \frac{0.001962}{\sqrt{3}} =$$

$\hat{\lambda} > 0$, so poor estimator

$$0.001962 \pm 0.002220 = [-0.0002580, 0.004182]$$

c) $\hat{S}(t) = exp(-t\hat{\lambda})$. Find $\hat{S}(100)$ so $t = 100$ minutes.

$$\exp[-100(.001962)] = \exp(-0.1962)$$

$$= \boxed{0.8218}$$

27

```
terms in PH model                                        AIC
trt  celltype  karno  diagtime  age   prior  964.79
trt  celltype  karno            age   prior  962.79
trt  celltype  karno            age          960.92
trt  celltype  karno                         959.83
     celltype  karno                         959.53
               karno                         970.87
coxph(formula = Surv(time, status) ~ celltype + karno, data = veteran)

                     coef exp(coef) se(coef)     z       p
celltypesmallcell 0.71533   2.04487  0.25269  2.83  0.0046
celltypeadeno     1.15773   3.18271  0.29294  3.95 7.7e-05
celltypelarge     0.32564   1.38492  0.27668  1.18  0.2392
karno            -0.03106   0.96942  0.00518 -6.00 2.0e-09

Likelihood ratio test=59.4  on 4 df, p=3.93e-12
n= 137, number of events= 128
```

4) The above output is for backward elimination using the R veteran data set from Kalbfleisch and Prentice (1980), *The Statistical Analysis of Failure Time Data.* Patients were randomly selected to receive two treatment regimens for lung cancer. The predictors were *trt*: 1=standard, 2=test; *celltype*: 1=squamous, 2=smallcell, 3=adeno, 4=large; *karno*: Karnofsky performance score (100=good); *diagtime*: months from diagnosis to randomization; *age*: in years; and *prior*: prior therapy 0=no, 1=yes. Note that *celltype* has 4 levels so was coded with 3 indicator variables.

a) What predictors are in the model I_{min}?

celltype, karno

b) The output above has variables x_1, x_2, x_3 and x_4 where x_i are indicators for *celltype* for $i = 1, 2, 3$ and $x_4 = karno$. Find the $ESP = \hat{\boldsymbol{\beta}}^T \boldsymbol{x}$ if $x_1 = 0$, $x_2 = 0$, $x_3 = 1$ and $x_4 = 80$.

$0.71533(0) + 1.15773(0) + 0.32564(1) - 0.03106(80)$

$= -2.1592$

18

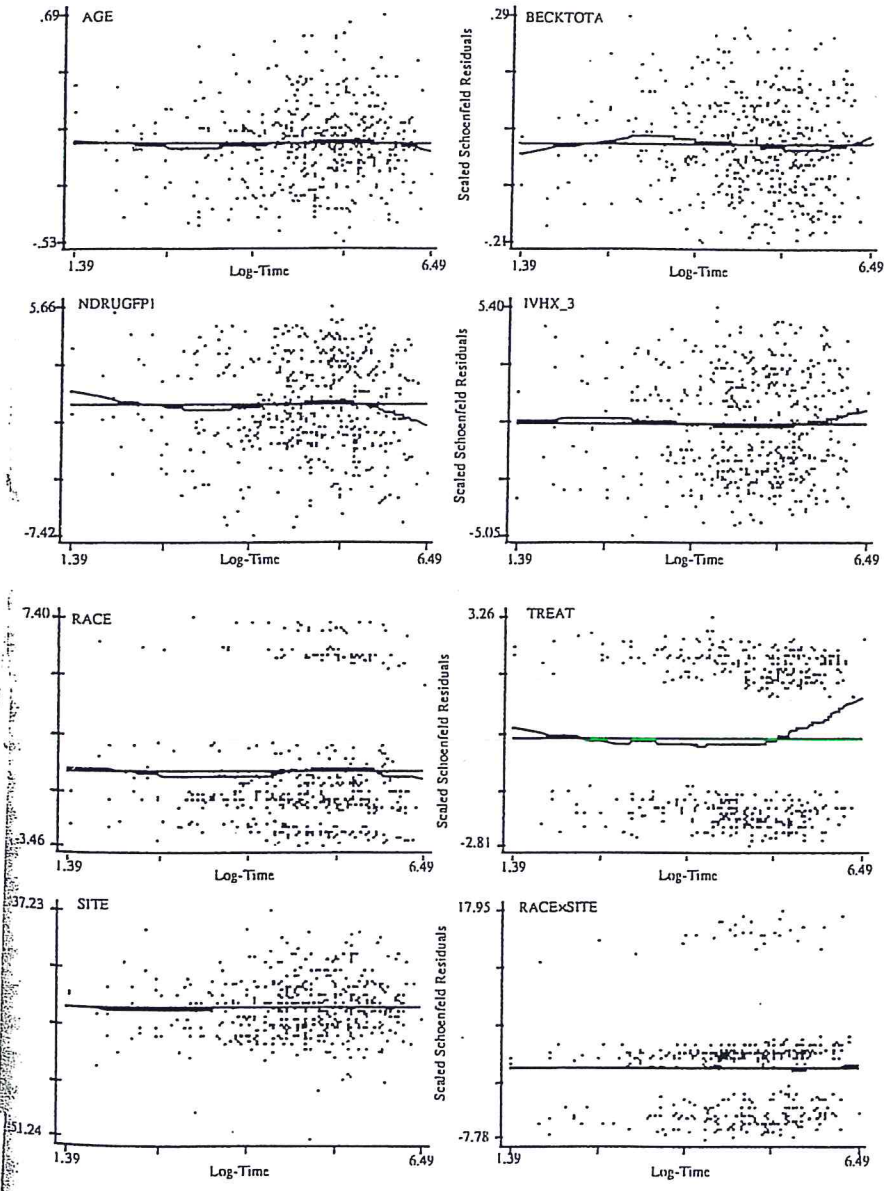


Figure 6.4 Graphs of the scaled Schoenfeld residuals and their lowess smooth obtained from the model in Table 5.11 for covariates (top left to bottom right): AGE, BECKTOTA, NDRUGFP1, IVHX_3, RACE, TREAT, SITE and RACExSITE. The zero line is drawn for reference.

Q8d21

25) The above plots suggest that the PH assumption may be reasonable, except perhaps for one variable. Which variable may suggest that the PH assumption is violated?

treat

10