Dept of Mathematics and Statistics
King Fahd University of Petroleum & Minerals

AS475: Survival Models for Actuaries
Dr. Mohammad H. Omar
Major 2 Exam Term 132 FORM A
Wednesday April 2 2014
7.00pm-8.30pm

Name_____ ID#:_____ Serial #:___

Instructions.

1. Please turn off your cell phones and place them under your chair. Any student caught with mobile phones on during the exam will be considered under the **cheating rules** of the University.
2. If you need to leave the room, please do so quietly so not to disturb others taking the test. No two person can leave the room at the same time. No extra time will be provided for the time missed outside the classroom.
3. Only materials provided by the instructor can be present on the table during the exam.
4. Do not spend too much time on any one question. If a question seems too difficult, leave it and go on.
5. Use the blank portions of each page for your work. Extra blank pages can be provided if necessary. If you use an extra page, indicate clearly what problem you are working on.
6. Only answers supported by work will be considered. Unsupported guesses will not be graded.
7. While every attempt is made to avoid defective questions, sometimes they do occur. In the rare event that you believe a question is defective, the instructor cannot give you any guidance beyond these instructions.
8. Mobile calculators, I-pad, or communicable devices are disallowed. Use regular scientific calculators or financial calculators only. Write important steps to arrive at the solution of the following problems.

The test is 90 minutes, GOOD LUCK, and you may begin now!

Question	Total Marks	Marks Obtained	Comments
1	1+4+1+3+3+1=13		
2	2+2+2=6		
3	2+3=5		
4	3+3=6		
5	2+2+2=6		
6	1+3=4 bonus		
Total	36		

Extra blank page

1. (1+4+1+3+3+1=13 points) Consider the survival data for 137 patients from the Veteran's Administration Lung Cancer Trial cited by *Kalbfleisch and Prentice* (1980). The variables in this dataset are listed as follows:

Variable #		**Variable name**	**Coding**
1		Treatment	$Standard = 1$, $test = 2$
2	indicator	Cell type 1	$Large = 1$, $other = 0$
3	variables	Cell type 2	$Adeno = 1$, $other = 0$
4	for	Cell type 3	$Small = 1$, $other = 0$
5	cell type	Cell type 4	$Squamous = 1$, $other = 0$
6		Survival time	(Days) integer counts
7		Performance status	$0 = worst, \ldots, 100 = best$
8		Disease duration	(Months) integer counts
9		Age	(Years) integer counts
10		Prior therapy	$None = 0$, $some = 10$
11		Status	$0 = censored$, $1 = died$

A survival model was fitted (on survival time variable 6) for these data, yielding the following edited computer results:

Variable name	Coef.	Std. Err.	$p > \|z\|$	Haz. Ratio .	[95% Conf	interval]
1 Treatment	0.290	0.207	0.162	1.336	0.890	2.006
3 Adeno cell	0.789	0.303	0.009	2.200	1.216	_____
4 Small cell	0.457	0.266	0.086	1.579	0.937	2.661
5 Squamous cell	−0.400	0.283	0.157	0.671	0.385	1.167
7 Perf. status	−0.033	0.006	0.000	0.968	0.958	0.978
8 Disease dur	0.000	0.009	0.992	1.000	0.982	1.018
9 Age	−0.009	0.009	0.358	0.991	0.974	1.010
10 Prior therapy	0.007	0.023	0.755	1.007	0.962	1.054

Log likelihood = −475.180

a) Write the **Cox model** used to obtain the above computer results.

b) Perform an **omnibus test** of the significance of this survival model. (Use 5% significance level).

c) Complete the *blank* information in the computer printout above.

d) Report the **hazard ratio** that compares persons with *adeno cell* type to persons with *large cell* type and **write** the formula for this hazard ratio

e) Based on the computer results and at 5% significance level, is there an effect of **treatment** on survival time?

f) Write an expression of the **estimated survival curve** for a person who was given the test treatment and who had a squamous cell type, where the variables to be adjusted are performance status, disease duration, age, and prior therapy.

2. (2+2+2=6 points) Consider the following computer output for the same data set in question 1.

Variable name	Coef.	Std. Err.	$p > \|z\|$	Haz. Ratio .	[95% Conf	interval]	*P(PH)*
1 Treatment	0.290	0.207	0.162	1.336	0.890	2.006	0.628
3 Adeno cell	0.789	0.303	0.009	2.200	1.216	_ _ _ _	0.033
4 Small cell	0.457	0.266	0.086	1.579	0.937	2.661	0.081
5 Squamous cell	0.400	0.283	0.157	0.671	0.385	1.167	0.078
7 Perf. status	0.033	0.006	0.000	0.968	0.958	0.978	0.000
8 Disease dur	0.000	0.009	0.992	1.000	0.982	1.018	0.919
9 Age	0.009	0.009	0.358	0.991	0.974	1.010	0.198
10 Prior therapy	0.007	0.023	0.755	1.007	0.962	1.054	0.145

a) **Which variable(s)** appear to violate the proportional hazards assumption of the Cox PH model.
b) **How do** you check for **specific violations** of the proportional hazards assumption?
c) Given your answer in (a) above, **provide suggestion** for improving the model.

3. (2+3=5 points) A computer analysis on leukimia patients ($n = 42$) shows the following results:

Variable name	Coef.	Std. Err.	$p > \|z\|$	Haz. Ratio .	[95% Conf	interval]
$Log\ WBC$	1.170	0.499	0.019	3.222	1.213	8.562
Rx	0.267	0.566	0.637	1.306	0.431	3.959
$Sex \times Log\ WBC$	0.469	0.720	0.515	1.598	0.390	6.549
$Sex \times Rx$	1.592	0.923	0.084	4.915	0.805	30.003

No. of subjects = 42 Log likelihood = -55.835 Stratified by Sex

(a) Write the Cox model represented by the computer printout above.

(b) If you know that a reduced model with only $Log\ WBC$ and Rx as predictors results in a log likehood or -57.560, at a 5% significance level, what can you conclude about the **full** model?

4. (3+3=6 points). Consider the following group of people participating in BUPA's medical insurance policy.

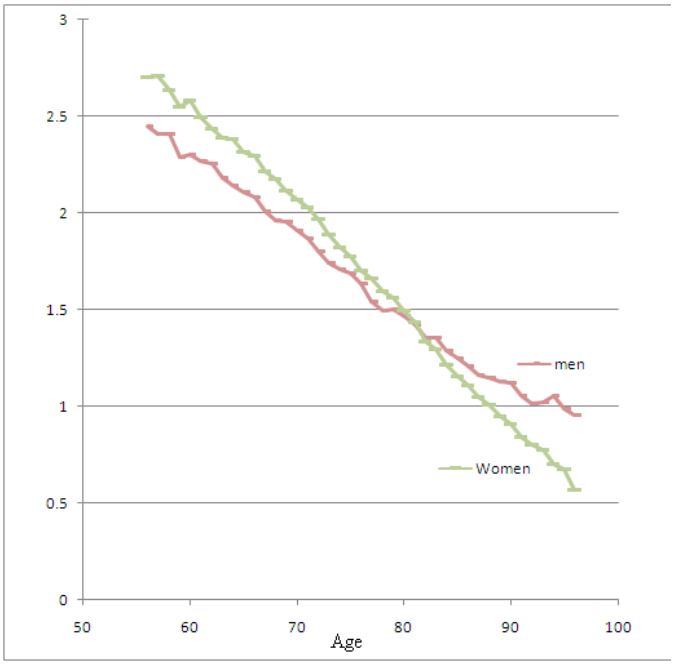
ID	$TIME$	$STATUS$	$SMOKE$
Barry	2	1	1
Gary	3	1	0
Harry	5	0	0
Larry	8	1	1

where $TIME$ = Survival time (in years), $STATUS = 1$ for event, 0 for censorship, and $SMOKE$= 1 for a smoker, 0 for a nonsmoker. Consider also the Cox model $h(t, X) = h_0(t) \exp[\beta_1 SMOKE]$.

a) **Write** the expression for the **likelihood function** to estimate the Cox PH model with $SMOKE$ as a single predictor variable.

b) Given your answer in (a) above, **write the procedure how** you would obtain the **maximum likelihood estimates (MLE)** for this model.

5. (2+2+2=6 points) A survival analysis study in Valencia region of Spain provides survival probabilities for men and women of ages 56 to 96. The following graphical display was conducted to examine the proportional hazards assumption of the cox model with respect to gender.



a) **Label** the Y axis of this graphical display

b) **Decide** from the graph whether or not the proportional hazards assumptions is met.

c) Given your decision in (b), **describe briefly** one strategy you would use to provide a model for the hazard rates.

6. (3+1=4 bonus points) You are given:

(i) A Cox proportional hazards model was used to study losses on two groups of policies.
(ii) A single covariate z was used with $z = 0$ for a policy in Group 1 and $z = 1$ for a policy in Group 2.
(iii) A sample of three policies was taken from each group. The losses were:

Group 1:	275	325	520
Group 2:	215	250	300

(iv) The baseline survival function is

$$S_0(x) = \left(\frac{200}{x}\right)^{\alpha}, \quad x > 200, \quad \alpha > 0.$$

Calculate the maximum likelihood estimate of the coefficient β.

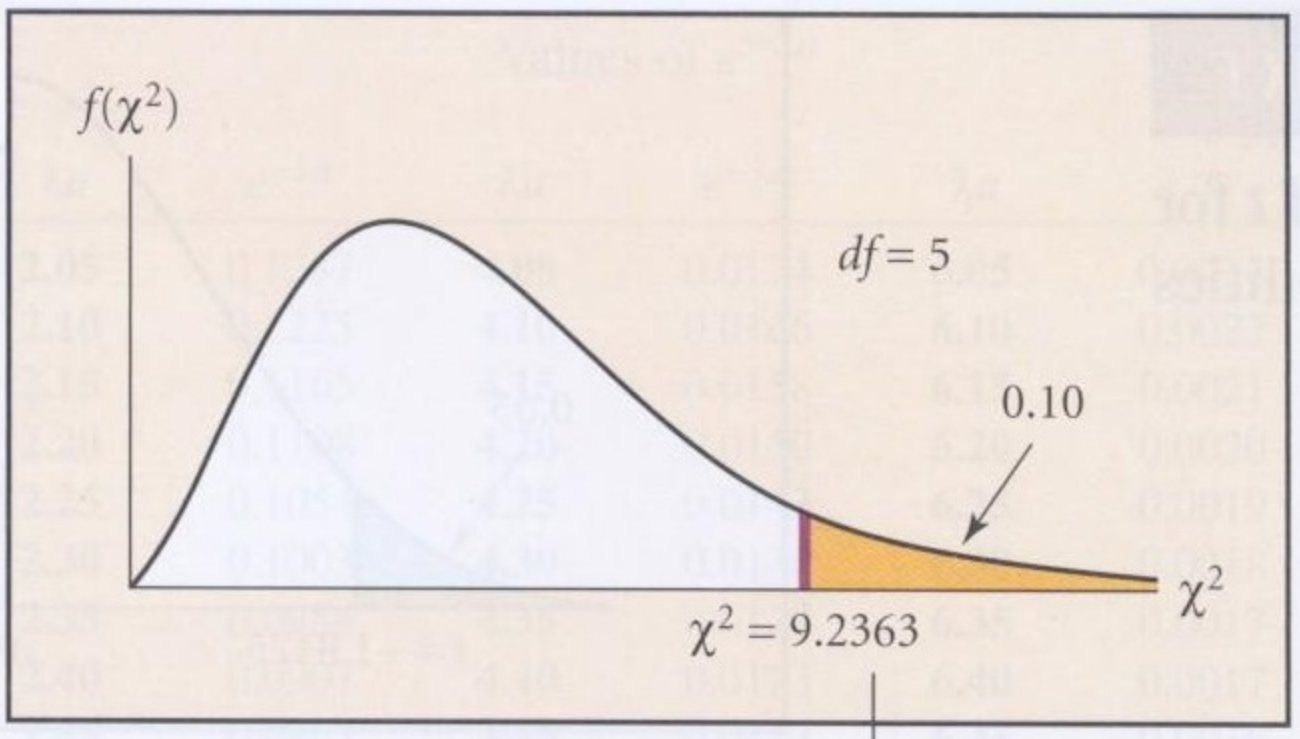
a) -0.92
b) -0.40
c) 0.40
d) 0.92
e) 2.51
Work Shown (3 points)

Answer is _____.

END OF TEST PAPER

APPENDIX G

Values of χ^2 for Selected Probabilities



PROBABILITIES (OR AREAS UNDER CHI-SQUARE DISTRIBUTION CURVE ABOVE GIVEN CHI-SQUARE VALUES)

	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
df					*Values of Chi-Squared*					
1	0.0000	0.0002	0.0010	0.0039	0.0158	2.7055	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.6052	5.9915	7.3778	9.2104	10.5965
3	0.0717	0.1148	0.2158	0.3518	0.5844	6.2514	7.8147	9.3484	11.3449	12.8381
4	0.2070	0.2971	0.4844	0.7107	1.0636	7.7794	9.4877	11.1433	13.2767	14.8602
5	0.4118	0.5543	0.8312	1.1455	1.6103	9.2363	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	2.2041	10.6446	12.5916	14.4494	16.8119	18.5475
7	0.9893	1.2390	1.6899	2.1673	2.8331	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	3.4895	13.3616	15.5073	17.5345	20.0902	21.9549
9	1.7349	2.0879	2.7004	3.3251	4.1682	14.6837	16.9190	19.0228	21.6660	23.5893
10	2.1558	2.5582	3.2470	3.9403	4.8652	15.9872	18.3070	20.4832	23.2093	25.1881
11	2.6032	3.0535	3.8157	4.5748	5.5778	17.2750	19.6752	21.9200	24.7250	26.7569
12	3.0738	3.5706	4.4038	5.2260	6.3038	18.5493	21.0261	23.3367	26.2170	28.2997
13	3.5650	4.1069	5.0087	5.8919	7.0415	19.8119	22.3620	24.7356	27.6882	29.8193
14	4.0747	4.6604	5.6287	6.5706	7.7895	21.0641	23.6848	26.1189	29.1412	31.3194
15	4.6009	5.2294	6.2621	7.2609	8.5468	22.3071	24.9958	27.4884	30.5780	32.8015
16	5.1422	5.8122	6.9077	7.9616	9.3122	23.5418	26.2962	28.8453	31.9999	34.2671
17	5.6973	6.4077	7.5642	8.6718	10.0852	24.7690	27.5871	30.1910	33.4087	35.7184
18	6.2648	7.0149	8.2307	9.3904	10.8649	25.9894	28.8693	31.5264	34.8052	37.1564
19	6.8439	7.6327	8.9065	10.1170	11.6509	27.2036	30.1435	32.8523	36.1908	38.5821
20	7.4338	8.2604	9.5908	10.8508	12.4426	28.4120	31.4104	34.1696	37.5663	39.9969
21	8.0336	8.8972	10.2829	11.5913	13.2396	29.6151	32.6706	35.4789	38.9322	41.4009
22	8.6427	9.5425	10.9823	12.3380	14.0415	30.8133	33.9245	36.7807	40.2894	42.7957
23	9.2604	10.1957	11.6885	13.0905	14.8480	32.0069	35.1725	38.0756	41.6383	44.1814
24	9.8862	10.8563	12.4011	13.8484	15.6587	33.1962	36.4150	39.3641	42.9798	45.5584
25	10.5196	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3140	46.9280
26	11.1602	12.1982	13.8439	15.3792	17.2919	35.5632	38.8851	41.9231	45.6416	48.2898
27	11.8077	12.8785	14.5734	16.1514	18.1139	36.7412	40.1133	43.1945	46.9628	49.6450
28	12.4613	13.5647	15.3079	16.9279	18.9392	37.9159	41.3372	44.4608	48.2782	50.9936
29	13.1211	14.2564	16.0471	17.7084	19.7677	39.0875	42.5569	45.7223	49.5878	52.3355
30	13.7867	14.9535	16.7908	18.4927	20.5992	40.2560	43.7730	46.9792	50.8922	53.6719