Name and Surname: [redacted]

Album no: .....

PRACTICAL EXAM (50 points)

Time to solve the exam: 45 min.

The exam dataset relates to patients with diagnosed lung cancer and the time to patient's death caused by the illness is to be investigated. Below you will find some information about the dataset named lung_cancer.

Attention! Please make the following assumptions:

- Significance level: 0.05
- Please use the default reference level for discrete variables
- Use Efron correction for tied events
- Only the Age and KPS variables are the continuous variables

Variable name	Variable definition	Additional info
Id	Unique patient id	
Therapy	A type of therapy that a patient was assigned for as a lung cancer treatment	*standard- standard therapy* *test- test chemotherapy*
Cell	Type of tumor diagnosed	*adeno* *large* *small* *squamous*
Kps	The Karnofsky scale that is an estimation of patient's general confition and quality of life	*Values differ from 100 to 0;* *100 – perfect condition* *0 - death*
Age	Patient's age at lung cancer diagnosis	*Continuous variable*
Prior	Information if the patient went through a cancer treatment prior to lung cancer diagnosis	*no* *yes*
Time	The time variable is calculated from lung cancer diagnosis until death or censoring (time is measured by days)	*Continuous variable*
Censor	Censoring variable	*1 – patient's death* *0 - censoring*

c) Please interpret the parameter for the explanatory variable ***Therapy***:

i) (1p.) Value of estimated parameter β : .....

ii) (3p.) Interpretation of the parameter β: .....
.....
.....
.....

Exercise 2. (14 p)

Please estimate semiparametric model with only 2 explanatory variables: ***Therapy*** and ***Kps***. Use Efron correction for tied events. Please perform all the tasks even if the explanatory variables will occur to be statistically insignificant

a) Using likelihood ratio test please check if taking out variable **Age** was reasonable (please find attached the table with chi-square distribution):

i) (4p.) State the test hypothesis for the likelihood ratio test:

H0: .....
.....
.....

H1: .....
.....
.....

ii) (1p.) Number of degrees of freedom: .....

iii) (1p.) Value of the test statistics (please do all the calculations below): .....
.....
.....
.....
.....

iv) (1p.) Critical value for the test statistics: .....

v) (3p.) Outcome interpertation: .....

Outcome interpretation: .....

.....

.....

.....

.....

.....

b) Please verify the proportional hazard assumption for the explanatory variable **Kps** using assess option in SAS:

i. (1p.) State what the PH assumption is going to test:

The hazard ratio does not depend on .....

iii. (9p.) Please verify the assumption of the PH for ***Kps*** variable using assess option.

State the test hypothesis:

H0: .....

.....

.....

H1: .....

Please write the p-value for the test: .....

Outcome interpretation: .....

.....

.....

.....

.....

Chi-Square distribution:

n / α	0,99	0,98	0,97	0,96	0,95	0,9	0,1	0,05	0,04	0,03	0,02	0,01
1	0,0002	0,0006	0,0014	0,0025	0,0039	0,0158	2,7055	3,8415	4,2179	4,7093	5,4119	6,6349
2	0,0201	0,0404	0,0609	0,0816	0,1026	0,2107	4,6052	5,9915	6,4378	7,0131	7,8240	9,2103
3	0,1148	0,1848	0,2451	0,3002	0,3518	0,5844	6,2514	7,8147	8,3112	8,9473	9,8374	11,3449
4	0,2971	0,4294	0,5351	0,6271	0,7107	1,0636	7,7794	9,4877	10,0255	10,7119	11,6678	13,2767
5	0,5543	0,7519	0,9031	1,0313	1,1455	1,6103	9,2364	11,0705	11,6443	12,3746	13,3882	15,0863
6	0,8721	1,1344	1,3296	1,4924	1,6354	2,2041	10,6446	12,5916	13,1978	13,9676	15,0332	16,8119
7	1,2390	1,5643	1,8016	1,9971	2,1673	2,8331	12,0170	14,0671	14,7030	15,5091	16,6224	18,4753
8	1,6465	2,0325	2,3101	2,5366	2,7326	3,4895	13,3616	15,5073	16,1708	17,0105	18,1682	20,0902
9	2,0879	2,5324	2,8485	3,1047	3,3251	4,1682	14,6837	16,9190	17,6083	18,4796	19,6790	21,6660
10	2,5582	3,0591	3,4121	3,6965	3,9403	4,8652	15,9872	18,3070	19,0207	19,9219	21,1608	23,2093
11	3,0535	3,6087	3,9972	4,3087	4,5748	5,5778	17,2750	19,6751	20,4120	21,3416	22,6179	24,7250
12	3,5706	4,1783	4,6009	4,9385	5,2260	6,3038	18,5493	21,0261	21,7851	22,7418	24,0540	26,2170
13	4,1069	4,7654	5,2210	5,5838	5,8919	7,0415	19,8119	22,3620	23,1423	24,1249	25,4715	27,6882
14	4,6604	5,3682	5,8556	6,2426	6,5706	7,7895	21,0641	23,6848	24,4855	25,4931	26,8728	29,1412
15	5,2293	5,9849	6,5032	6,9137	7,2609	8,5468	22,3071	24,9958	25,8162	26,8479	28,2595	30,5779
16	5,8122	6,6142	7,1625	7,5958	7,9616	9,3122	23,5418	26,2962	27,1356	28,1907	29,6332	31,9999
17	6,4078	7,2550	7,8324	8,2878	8,6718	10,0852	24,7690	27,5871	28,4450	29,5227	30,9950	33,4087
18	7,0149	7,9062	8,5120	8,9889	9,3905	10,8649	25,9894	28,8693	29,7451	30,8447	32,3462	34,8053
19	7,6327	8,5670	9,2004	9,6983	10,1170	11,6509	27,2036	30,1435	31,0367	32,1577	33,6874	36,1908
20	8,2604	9,2367	9,8971	10,4154	10,8508	12,4426	28,4120	31,4104	32,3206	33,4624	35,0196	37,5662
21	8,8972	9,9146	10,6013	11,1395	11,5913	13,2396	29,6151	32,6706	33,5972	34,7593	36,3434	38,932
22	9,5425	10,6000	11,3125	11,8703	12,3380	14,0415	30,8133	33,9244	34,8673	36,0492	37,6595	40,289
23	10,1957	11,2926	12,0303	12,6072	13,0905	14,8480	32,0069	35,1725	36,1311	37,3323	38,9683	41,63
24	10,8564	11,9918	12,7543	13,3498	13,8484	15,6587	33,1962	36,4150	37,3891	38,6093	40,2704	42,9
25	11,5240	12,6973	13,4840	14,0978	14,6114	16,4734	34,3816	37,6525	38,6416	39,8804	41,5661	44,3
26	12,1981	13,4086	14,2190	14,8509	15,3792	17,2919	35,5632	38,8851	39,8891	41,1460	42,8558	45,6
27	12,8785	14,1254	14,9592	15,6087	16,1514	18,1139	36,7412	40,1133	41,1318	42,4066	44,1400	46,9
28	13,5647	14,8475	15,7042	16,3711	16,9279	18,9392	37,9159	41,3371	42,3699	43,6622	45,4188	48.
29	14,2565	15,5745	16,4538	17,1377	17,7084	19,7677	39,0875	42,5570	43,6038	44,9132	46,6927	49
30	14,9535	16,3062	17,2076	17,9083	18,4927	20,5992	40,2560	43,7730	44,8338	46,1599	47,9618	50
40	22,1643	23,8376	24,9437	25,7989	26,5093	29,0505	51,8051	55,7585	56,9459	58,4278	60,4361	63
50	29,7067	31,6639	32,9509	33,9426	34,7643	37,6886	63,1671	67,5048	68,8039	70,4230	72,6133	7
60	37,4849	39,6994	41,1504	42,2656	43,1880	46,4589	74,3970	79,0819	80,4820	82,2251	84,5799	8
70	45,4417	47,8934	49,4953	50,7243	51,7393	55,3289	85,5270	90,5312	92,0241	93,8813	96,3875	10
80	53,5401	56,2128	57,9553	59,2902	60,3915	64,2778	96,5782	101,8795	103,4588	105,4221	108,0693	1
90	61,7541	64,6347	66,5093	67,9437	69,1260	73,2911	107,5650	113,1453	114,8057	116,8688	119,6485	1
100	70,0649	73,1422	75,1419	76,6705	77,9295	82,3581	118,4980	124,3421	126,0794	128,2367	131,1417	

.....

.....

b) Compare the model 1 (the model from the Exercise 1) with the model 2 (the model from Exercise 2) using Akaike Information Criterion:

i) (2p) Please fill in the value of the criterion below :

Model 1: .....

Model 2: .....

ii) (2p) Please indicate the best model and justify the choice:.....

.....

.....

Exercise 3. (20 p)

Estimate the model as in Exercise 1. Please perform all the tasks even if the explanatory variables will occur to be statistically insignificant

a) Verify the assumption of <u>linearity</u> with respect to explanatory variable **Age**:

i. *(1p.)* State linearity of what is being tested:

..... is linear with respect to explanatory variable ***Age***

ii. *(9p.)* Please verify the assumption from the previous point using assess option in SAS.

Please explain why is it a non-parametic test: .....

.....

.....

.....

State the test hypothesis:

H0: .....

.....

.....

H1: .....

Please write the p-value for the test: .....

Exercice 1. (16 p)

Please estimate semiparametric model with only 3 explanatory variables: ***Therapy***, ***Age*** and ***Kps***. Use Efron correction for tied events. Please perform all the tasks even if the explanatory variables will occur to be statistically insignificant.

a) Determine if explanatory variable *Therapy* is statistically significant:

i) (4p.) State the test hypothesis for the type 3 test:

H0: .....

.....

.....

H1: .....

.....

.....

ii) (1p.) Value of test statistic: .....

iii) (2p.) Outcome interpretation: .....

.....

.....

.....

.....

b) Please interpret the parameter for the explanatory variable ***Age***:

i) (1p.) Value of estimated parameter β : .....

ii) (1p.) What is the functional form of relationship between parameter β ar

ratio: .....

iii) (3p.) Interpretation of the parameter β : .....

.....

.....

.....