STAT 4008 Assignment 3 **Due: March 29, 2021**

1. You are given the following right-censored sample

$$12, 15+, 17, 17, 18, 19+, 20, 20, 20+, 21+, 24, 27$$

Test if the hazard function $h(t) = 0.2, 0 < t < 27$.

2. You are given the following data:

Group 1	:	2,2,3+,3,4,4+,5,5,6+
Group 2	:	2, 3,3,3+,4,4,4+,5,5+5+,6,7

+:censored

Test if these two groups of data have the same distribution.

3. The following table gives survival data from 30 patients with AML. Two possible prognostic factors are considered:

$$x_1 = \begin{cases} 1 & \text{if patient} \geq 50 \text{ years old} \\ 0 & \text{otherwise} \end{cases}$$

$$x_2 = \begin{cases} 1 & \text{if cellularity of mallow clot section is } 100\% \\ 0 & \text{otherwise} \end{cases}$$

Table. Survival Times and Data of Two Possible Prognostic Factors and 30 AML Patients

Survival Time	x_1	x_2	Survival Time	x_1	x_2
18	0	0	8	1	0
9	0	1	2	1	1
28+	0	0	26+	1	0
31	0	1	10	1	1
39+	0	1	4	1	0
19+	0	1	3	1	0
45+	0	1	4	1	0
6	0	1	18	1	1
8	0	1	8	1	1
15	0	1	3	1	1
23	0	0	14	1	1
28+	0	0	3	1	0
7	0	1	13	1	1
12	1	0	13	1	1
9	1	0	35+	1	0

Test if the prognostic factors are significance.

4. A data set is given in "ass3data.xls" on the blackboard system. It contains "Times" in Column 1, "Group" in Column 2 and "Treatment" in Column 3. Note that there is no censoring in the data set.

   (a) test if there is difference among groups

   (b) test if there is difference among treatments

5. A data set is given in "ass3q5.csv" on the blackboard. It contains Time, Status (0: lived; 1: death) and Smoking Status. Use the seed 123457, generate a subsample of size 100 and then test if there are any difference among the Smoking Status based on this subsample.