You can consult your textbook and notes during the exam, and you are permitted to use R to complete any calculations that are necessary. You will need to make it clear on your exam paper how you performed the calculation that lead to your answer.

It is not permitted to talk to any other people about the exam or consult websites. We will be taking measures to discover cheating, and any students who we determine have copied their answers will receive a failing grade in the course and be reported to the university.

1. We have a study which observes two times (T_1, T_2): the T_1 is the time of onset of symptoms, and then T_2 is time of death following the symptoms. We would like to use a recurrent events model for the two successive outcomes. In particular, we will suppose that we have a Gap Model where the hazard rate for T_1 is

$$h_1(t) = \frac{4t}{2t+1},$$

and then we have a hazard rate for the time after T_1 occurs $D = T_2 - T_1$ which is $h_2(d) = 2$.

   (a) Calculate the survival function $S_1(t)$ for the first occurrence.

   (b) Calculate the marginal survival function $S(t_2)$ for the time T_2.

2. The service times recorded for 45 employees were

3	6	7+	7+	10	10	10	10+	13
14	15	15+	19	19	20+	20	22	22
27	27	28	29+	30	31+	34	35	36
37	38	39	39	39	39	40	42	43
43	47	47+	50+	58+	59+	70+	74+	78+

   Observations with "+" were censored at that time.

   (a) Calculate the estimate of the hazard probability at time $t = 47$.

   (b) Estimate the survival function at time $t = 13$.

   (c) Estimate the 20% percentile of the distribution of the service times. (i.e. find a such that $\mathbb{P}\{T \leq a\} = 0.2$.)

3. The data set `Neph.txt` contains observation from a small kidney disease study. The first column in the data file is an indicator of whether or not the subjects received a nephrectomy treatment. The second column is a factor dividing the subjects into three age ranges. The last two columns measure the time until an event and the status. We wish to test whether the nephrectomy treatment extended the times until failure.

   (a) Estimate the median survival time and give a 95% confidence interval.

   (b) Perform a likelihood ratio test for an effect from the nephrectomy treatment when we control for the differences between the ages using a Cox PH model. Report a P value for the test.

   (c) Repeat this test except use a model that is stratified on the age groups.

   (d) Describe how the two models in (b) and (c) are different, and what factors you would consider when deciding which one is the correct test to use.