Homework #3 [14 points]

Due on Feb 12, 11:59**AM**

**NOTE: Only cut and paste relevant output for each question. Attach the SAS/R codes to the end of your homework.**

1. Use SAS/R to repeat the log-rank test between the survival curves of the two treatment groups in the myelomatosis example in this week’s lecture notes. Use SAS/R to compute the Kaplan-Meier estimators and 95% CI of the survival functions of the two treatment groups (one Kaplan-Meier table for each group). Cut and paste relevant code and output. [4 points]
2. The “hwdata1.csv” provides the time until breast cancer recurrence (in days) for 686 patients with primary node positive breast cancer. The codebook is provided below.

|  |  |  |
| --- | --- | --- |
| Variable | Description | Codes/Values/Range |
| id | Study ID | 1 - 686 |
| diagdate | Date of Diagnosis | ddmmmyy |
| recdate | Date of Recurrence Free Survival | ddmmmyy |
| age | Age at Diagnosis | Years |
| menopause | Menopausal Status | 1 = Yes, 2 = No |
| hormone | Hormone Therapy | 1 = Yes, 2 = No |
| size | Tumor Size | mm |
| grade | Tumor Grade | 1 -3 |
| nodes | Number of Nodes involved | 1 - 51 |
| prog recp | Number of Progesterone Receptors | 1 - 2380 |
| estrg recp | Number of Estrogen Receptors | 1 - 1144 |
| rectime | Time to Recurrence | Days |
| censrec | Recurrence Censoring | 0 = Censored, 1 = Recurrence |

* 1. Use SAS/R to generate a graph of the survival functions in the two hor- mone therapy groups. Interpret the graph. (hint: What do you observe on the difference of the survival functions between the two groups?) [2 points]
  2. Test whether the survival curves for the two hormone therapy groups are the same using the log-rank test at *α* = 0*.*05. Give the null and alternative

hypothesis, test statistic, degrees of freedom, p-value, and conclusion. [3 points]

* 1. Repeat the test in Question 2 using the Wilcoxon test. Does it lead to the same conclusion? Briefly explain why the test statistic of the Wilcoxon test is smaller than that of the log-rank test in this application. [2 points]
  2. Categorize the tumor size into four groups using the sample quartiles (*<* 20, [20*,* 25), [25*,* 35), and *≥* 35). Test whether the survival functions are the same among the four groups using the generalized log-rank test at *α* = 0*.*05. Give the null and alternative hypothesis, test statistic, degrees of freedom, p-value, and conclusion. [3 points]