**VDPAM 527**

**Dr. Chong Wang**

**Shahan Azeem**

**Home Work # 13**

**12-13-13**

The data set in " http://www.public.iastate.edu/~chwang/data/Cancer.csv " comes from a paper that appeared in the Journal of the American Statistical Association by Wei et al (1989). In this study, tumors were removed from the bladders of 86 patients. Subsequently, the patients were assigned to be treated either with a placebo or with the drug thiopeta. The variables in data include:

Time: This is the time (in months) to first recurrence of a tumor or the time at which censoring took place.

Censored: This equals 1 if censoring took place and 0 otherwise.

Treatment: This equals 1 for placebo and 2 for thiopeta.

Number: This equals 1 if originally one tumor was removed and 2 if two or more tumors were removed.

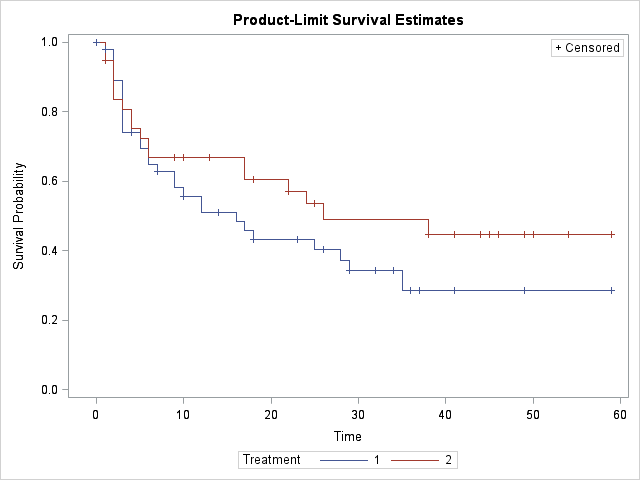
You may find it useful to note the following:

Among 48 subjects receiving placebo, 29 developed new tumors while in the study.

Among 38 subjects receiving thiopeta, 18 developed new tumors while in the study.

[100] 1. Begin by examining the role of treatment without taking into account the number of tumors removed. In what follows let S1(t) denote the survival function for patients receiving placebo and S2(t) the survival function for patients receiving thiopeta.

[25] a. Create and submit a plot of the Kaplan-Meier estimates for S1(t) and S2(t). Clearly label which estimate pertains to which stage.



A + sign in both the curves indicates times when censoring is taking place.

[25] b. Interpret S1(6 ) and S2(6 ).

S1 (6) = 0.6498

This is the estimated probability (64.98%) of avoiding development of new tumor for at least full six months in the placebo group.

S2 (6) = 0.6687

This is the estimated probability (66.87 %) of avoiding development of new tumor for at least full six months in the thiopeta treated group.

[25] c. Provide point and interval estimates for S1 (6 ) and S2(6 ).

𝑆 𝑡i ±𝑧1−𝛼/2∙𝑠𝑒 𝑆 𝑡i

For α=0.05 the critical value of 𝑧1−𝛼/2=1.96

For S1(6):

0.6498 + 1.96 (0.0707)

0.6498 + 0.1386

Point and interval estimate S1(6)= **0.6498 ( 0.5112, 0.7884)**

For S2(6):

0.6687 + 1.96 (0.0783)

0.6687 + 0.15346

Point and interval estimate S2 (6) = **0.6687 (0.5152 , 0.8222 )**

[25] d. At level 0.05 test the null hypothesis that S1(t) and S2(t ) are identical against the alternative hypothesis that they differ.

𝐻0: 𝑆1 (t)=𝑆2 (t) for every time t

𝐻1: 𝑆1 (t) ≠𝑆2 (t) for at least one time t

| **Test of Equality over Strata** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > Chi-Square** |
| Log-Rank | 1.5209 | 1 | 0.2175 |
| **Wilcoxon** | 0.7660 | 1 | 0.3815 |
| **-2Log(LR)** | 3.2114 | 1 | 0.0731 |

**Chi-square= 1.5209**

**p-value= 0.2175**

Since p-value > 0.05 we fail to reject null at significance 0.05. Therefore, we may not conclude that survival functions of placebo and thiopeta treated patients differ (at 0.05 level of significance).

**SAS Output:**

|  |
| --- |
| The SAS System |

| **Obs** | **Time** | **Censored** | **Treatment** | **Number** |
| --- | --- | --- | --- | --- |
| **1** | 0 | 1 | 1 | 1 |
| **2** | 1 | 1 | 1 | 2 |
| **3** | 4 | 1 | 1 | 1 |
| **4** | 7 | 1 | 1 | 1 |
| **5** | 10 | 1 | 1 | 1 |
| **6** | 6 | 0 | 1 | 1 |
| **7** | 14 | 1 | 1 | 1 |
| **8** | 18 | 1 | 1 | 1 |
| **9** | 5 | 0 | 1 | 2 |
| **10** | 12 | 0 | 1 | 1 |
| **11** | 23 | 1 | 1 | 2 |
| **12** | 10 | 0 | 1 | 2 |
| **13** | 3 | 0 | 1 | 1 |
| **14** | 3 | 0 | 1 | 1 |
| **15** | 7 | 0 | 1 | 2 |
| **16** | 3 | 0 | 1 | 1 |
| **17** | 26 | 1 | 1 | 2 |
| **18** | 1 | 0 | 1 | 1 |
| **19** | 2 | 0 | 1 | 2 |
| **20** | 25 | 0 | 1 | 2 |
| **21** | 29 | 1 | 1 | 2 |
| **22** | 29 | 1 | 1 | 2 |
| **23** | 29 | 1 | 1 | 1 |
| **24** | 28 | 0 | 1 | 2 |
| **25** | 2 | 0 | 1 | 2 |
| **26** | 3 | 0 | 1 | 1 |
| **27** | 12 | 0 | 1 | 2 |
| **28** | 32 | 1 | 1 | 2 |
| **29** | 34 | 1 | 1 | 1 |
| **30** | 36 | 1 | 1 | 1 |
| **31** | 29 | 0 | 1 | 1 |
| **32** | 37 | 1 | 1 | 2 |
| **33** | 9 | 0 | 1 | 1 |
| **34** | 16 | 0 | 1 | 1 |
| **35** | 41 | 1 | 1 | 2 |
| **36** | 3 | 0 | 1 | 1 |
| **37** | 6 | 0 | 1 | 2 |
| **38** | 3 | 0 | 1 | 1 |
| **39** | 9 | 0 | 1 | 1 |
| **40** | 18 | 0 | 1 | 1 |
| **41** | 49 | 1 | 1 | 2 |
| **42** | 35 | 0 | 1 | 1 |
| **43** | 17 | 0 | 1 | 2 |
| **44** | 3 | 0 | 1 | 1 |
| **45** | 59 | 1 | 1 | 1 |
| **46** | 2 | 0 | 1 | 2 |
| **47** | 5 | 0 | 1 | 2 |
| **48** | 2 | 0 | 1 | 2 |
| **49** | 1 | 1 | 2 | 2 |
| **50** | 1 | 1 | 2 | 1 |
| **51** | 5 | 0 | 2 | 1 |
| **52** | 9 | 1 | 2 | 2 |
| **53** | 10 | 1 | 2 | 1 |
| **54** | 13 | 1 | 2 | 1 |
| **55** | 3 | 0 | 2 | 2 |
| **56** | 1 | 0 | 2 | 2 |
| **57** | 18 | 1 | 2 | 1 |
| **58** | 17 | 0 | 2 | 2 |
| **59** | 2 | 0 | 2 | 1 |
| **60** | 17 | 0 | 2 | 1 |
| **61** | 22 | 1 | 2 | 1 |
| **62** | 25 | 1 | 2 | 2 |
| **63** | 25 | 1 | 2 | 2 |
| **64** | 25 | 1 | 2 | 1 |
| **65** | 6 | 0 | 2 | 1 |
| **66** | 6 | 0 | 2 | 1 |
| **67** | 2 | 0 | 2 | 1 |
| **68** | 26 | 0 | 2 | 2 |
| **69** | 38 | 1 | 2 | 1 |
| **70** | 22 | 0 | 2 | 1 |
| **71** | 4 | 0 | 2 | 1 |
| **72** | 24 | 0 | 2 | 1 |
| **73** | 41 | 1 | 2 | 2 |
| **74** | 41 | 1 | 2 | 1 |
| **75** | 1 | 0 | 2 | 1 |
| **76** | 44 | 1 | 2 | 1 |
| **77** | 2 | 0 | 2 | 1 |
| **78** | 45 | 1 | 2 | 2 |
| **79** | 2 | 0 | 2 | 2 |
| **80** | 46 | 1 | 2 | 2 |
| **81** | 49 | 1 | 2 | 2 |
| **82** | 50 | 1 | 2 | 1 |
| **83** | 4 | 0 | 2 | 1 |
| **84** | 54 | 1 | 2 | 2 |
| **85** | 38 | 0 | 2 | 1 |
| **86** | 59 | 1 | 2 | 2 |

|  |
| --- |
| The SAS System |

The LIFETEST Procedure

Stratum 1: Treatment = 1

| **Product-Limit Survival Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Time** |  | **Survival** | **Failure** | **Survival Standard Error** | **Number Failed** | **Number Left** |
| **0.0000** |  | 1.0000 | 0 | 0 | 0 | 48 |
| **0.0000** | **\*** | . | . | . | 0 | 47 |
| **1.0000** |  | 0.9787 | 0.0213 | 0.0210 | 1 | 46 |
| **1.0000** | **\*** | . | . | . | 1 | 45 |
| **2.0000** |  | . | . | . | 2 | 44 |
| **2.0000** |  | . | . | . | 3 | 43 |
| **2.0000** |  | . | . | . | 4 | 42 |
| **2.0000** |  | 0.8917 | 0.1083 | 0.0457 | 5 | 41 |
| **3.0000** |  | . | . | . | 6 | 40 |
| **3.0000** |  | . | . | . | 7 | 39 |
| **3.0000** |  | . | . | . | 8 | 38 |
| **3.0000** |  | . | . | . | 9 | 37 |
| **3.0000** |  | . | . | . | 10 | 36 |
| **3.0000** |  | . | . | . | 11 | 35 |
| **3.0000** |  | 0.7395 | 0.2605 | 0.0647 | 12 | 34 |
| **4.0000** | **\*** | . | . | . | 12 | 33 |
| **5.0000** |  | . | . | . | 13 | 32 |
| **5.0000** |  | 0.6947 | 0.3053 | 0.0681 | 14 | 31 |
| **6.0000** |  | . | . | . | 15 | 30 |
| **6.0000** |  | 0.6498 | 0.3502 | 0.0707 | 16 | 29 |
| **7.0000** |  | 0.6274 | 0.3726 | 0.0717 | 17 | 28 |
| **7.0000** | **\*** | . | . | . | 17 | 27 |
| **9.0000** |  | . | . | . | 18 | 26 |
| **9.0000** |  | 0.5810 | 0.4190 | 0.0735 | 19 | 25 |
| **10.0000** |  | 0.5577 | 0.4423 | 0.0742 | 20 | 24 |
| **10.0000** | **\*** | . | . | . | 20 | 23 |
| **12.0000** |  | . | . | . | 21 | 22 |
| **12.0000** |  | 0.5092 | 0.4908 | 0.0752 | 22 | 21 |
| **14.0000** | **\*** | . | . | . | 22 | 20 |
| **16.0000** |  | 0.4838 | 0.5162 | 0.0757 | 23 | 19 |
| **17.0000** |  | 0.4583 | 0.5417 | 0.0758 | 24 | 18 |
| **18.0000** |  | 0.4328 | 0.5672 | 0.0758 | 25 | 17 |
| **18.0000** | **\*** | . | . | . | 25 | 16 |
| **23.0000** | **\*** | . | . | . | 25 | 15 |
| **25.0000** |  | 0.4040 | 0.5960 | 0.0760 | 26 | 14 |
| **26.0000** | **\*** | . | . | . | 26 | 13 |
| **28.0000** |  | 0.3729 | 0.6271 | 0.0763 | 27 | 12 |
| **29.0000** |  | 0.3418 | 0.6582 | 0.0760 | 28 | 11 |
| **29.0000** | **\*** | . | . | . | 28 | 10 |
| **29.0000** | **\*** | . | . | . | 28 | 9 |
| **29.0000** | **\*** | . | . | . | 28 | 8 |
| **32.0000** | **\*** | . | . | . | 28 | 7 |
| **34.0000** | **\*** | . | . | . | 28 | 6 |
| **35.0000** |  | 0.2849 | 0.7151 | 0.0819 | 29 | 5 |
| **36.0000** | **\*** | . | . | . | 29 | 4 |
| **37.0000** | **\*** | . | . | . | 29 | 3 |
| **41.0000** | **\*** | . | . | . | 29 | 2 |
| **49.0000** | **\*** | . | . | . | 29 | 1 |
| **59.0000** | **\*** | 0.2849 | 0.7151 | . | 29 | 0 |

|  |  |
| --- | --- |
| **Note:** | **The marked survival times are censored observations.** |

Summary Statistics for Time Variable Time

| **Quartile Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Percent** | **Point Estimate** | **95% Confidence Interval** | | |
| **Transform** | **[Lower** | **Upper)** |
| **75** | . | LOGLOG | 28.0000 | . |
| **50** | 16.0000 | LOGLOG | 6.0000 | 29.0000 |
| **25** | 3.0000 | LOGLOG | 3.0000 | 7.0000 |

| **Mean** | **Standard Error** |
| --- | --- |
| 18.2899 | 2.1377 |

|  |  |
| --- | --- |
| **Note:** | **The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.** |

|  |
| --- |
| The SAS System |

The LIFETEST Procedure

Stratum 2: Treatment = 2

| **Product-Limit Survival Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Time** |  | **Survival** | **Failure** | **Survival Standard Error** | **Number Failed** | **Number Left** |
| **0.0000** |  | 1.0000 | 0 | 0 | 0 | 38 |
| **1.0000** |  | . | . | . | 1 | 37 |
| **1.0000** |  | 0.9474 | 0.0526 | 0.0362 | 2 | 36 |
| **1.0000** | **\*** | . | . | . | 2 | 35 |
| **1.0000** | **\*** | . | . | . | 2 | 34 |
| **2.0000** |  | . | . | . | 3 | 33 |
| **2.0000** |  | . | . | . | 4 | 32 |
| **2.0000** |  | . | . | . | 5 | 31 |
| **2.0000** |  | 0.8359 | 0.1641 | 0.0613 | 6 | 30 |
| **3.0000** |  | 0.8080 | 0.1920 | 0.0653 | 7 | 29 |
| **4.0000** |  | . | . | . | 8 | 28 |
| **4.0000** |  | 0.7523 | 0.2477 | 0.0717 | 9 | 27 |
| **5.0000** |  | 0.7245 | 0.2755 | 0.0743 | 10 | 26 |
| **6.0000** |  | . | . | . | 11 | 25 |
| **6.0000** |  | 0.6687 | 0.3313 | 0.0783 | 12 | 24 |
| **9.0000** | **\*** | . | . | . | 12 | 23 |
| **10.0000** | **\*** | . | . | . | 12 | 22 |
| **13.0000** | **\*** | . | . | . | 12 | 21 |
| **17.0000** |  | . | . | . | 13 | 20 |
| **17.0000** |  | 0.6050 | 0.3950 | 0.0828 | 14 | 19 |
| **18.0000** | **\*** | . | . | . | 14 | 18 |
| **22.0000** |  | 0.5714 | 0.4286 | 0.0848 | 15 | 17 |
| **22.0000** | **\*** | . | . | . | 15 | 16 |
| **24.0000** |  | 0.5357 | 0.4643 | 0.0867 | 16 | 15 |
| **25.0000** | **\*** | . | . | . | 16 | 14 |
| **25.0000** | **\*** | . | . | . | 16 | 13 |
| **25.0000** | **\*** | . | . | . | 16 | 12 |
| **26.0000** |  | 0.4911 | 0.5089 | 0.0902 | 17 | 11 |
| **38.0000** |  | 0.4464 | 0.5536 | 0.0924 | 18 | 10 |
| **38.0000** | **\*** | . | . | . | 18 | 9 |
| **41.0000** | **\*** | . | . | . | 18 | 8 |
| **41.0000** | **\*** | . | . | . | 18 | 7 |
| **44.0000** | **\*** | . | . | . | 18 | 6 |
| **45.0000** | **\*** | . | . | . | 18 | 5 |
| **46.0000** | **\*** | . | . | . | 18 | 4 |
| **49.0000** | **\*** | . | . | . | 18 | 3 |
| **50.0000** | **\*** | . | . | . | 18 | 2 |
| **54.0000** | **\*** | . | . | . | 18 | 1 |
| **59.0000** | **\*** | 0.4464 | 0.5536 | . | 18 | 0 |

|  |  |
| --- | --- |
| **Note:** | **The marked survival times are censored observations.** |

Summary Statistics for Time Variable Time

| **Quartile Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Percent** | **Point Estimate** | **95% Confidence Interval** | | |
| **Transform** | **[Lower** | **Upper)** |
| **75** | . | LOGLOG | . | . |
| **50** | 26.0000 | LOGLOG | 6.0000 | . |
| **25** | 5.0000 | LOGLOG | 2.0000 | 17.0000 |

| **Mean** | **Standard Error** |
| --- | --- |
| 23.5565 | 2.7247 |

|  |  |
| --- | --- |
| **Note:** | **The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.** |

| **Summary of the Number of Censored and Uncensored Values** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Stratum** | **Treatment** | **Total** | **Failed** | **Censored** | **Percent Censored** |
| **1** | **1** | 48 | 29 | 19 | 39.58 |
| **2** | **2** | 38 | 18 | 20 | 52.63 |
| **Total** |  | 86 | 47 | 39 | 45.35 |

|  |
| --- |
| The SAS System |

The LIFETEST Procedure

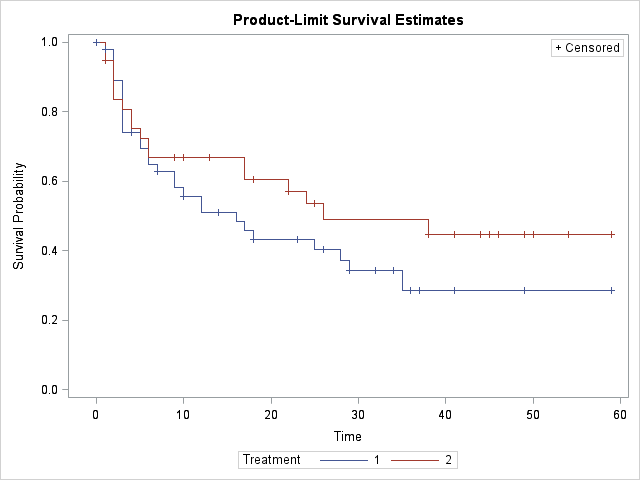
Testing Homogeneity of Survival Curves for Time over Strata

| **Rank Statistics** | | |
| --- | --- | --- |
| **Treatment** | **Log-Rank** | **Wilcoxon** |
| **1** | 4.0878 | 174.00 |
| **2** | -4.0878 | -174.00 |

| **Covariance Matrix for the Log-Rank Statistics** | | |
| --- | --- | --- |
| **Treatment** | **1** | **2** |
| **1** | 10.9867 | -10.9867 |
| **2** | -10.9867 | 10.9867 |

| **Covariance Matrix for the Wilcoxon Statistics** | | |
| --- | --- | --- |
| **Treatment** | **1** | **2** |
| **1** | 39524.5 | -39524.5 |
| **2** | -39524.5 | 39524.5 |

| **Test of Equality over Strata** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > Chi-Square** |
| **Log-Rank** | 1.5209 | 1 | 0.2175 |
| **Wilcoxon** | 0.7660 | 1 | 0.3815 |
| **-2Log(LR)** | 3.2114 | 1 | 0.0731 |



**SAS code:**

filename data url' http://www.public.iastate.edu/~chwang/data/Cancer.csv ';

**data** CR;

infile data delimiter=',' MISSOVER FIRSTOBS=**2** ;

input Time Censored Treatment Number;

**run**;

\* Printing the data set;

**PROC** **PRINT** DATA = CR;

**RUN**;

\* Kaplan-Meier estimation of survival curves and the log rank test;

symbol1 c=blue; symbol2 c=orange;

**proc** **lifetest** data=CR plots=(s);

time Time\*Censored(**1**);

strata Treatment;

**run**;