Part 6
Lecture 2 Survival Analysis - Non Parametric

Who I am…

Pascal Tyrrell, PhD

Associate Professor

Department of Medical Imaging, Faculty of Medicine

Institute of Medical Science, Faculty of Medicine

Department of Statistical Sciences, Faculty of Arts and Science

STUDY 2 LIFETEST PROCEDURE

Forty rats are randomly allocated to two groups of 20 animals. In one group the rats were exposed to a carcinogen and in the other group they were exposed to a placebo.

The outcome variable was the time in days from randomization to death.

```
PROC   FORMAT ;  VALUE RX  1   =   "DRUG X"   0 ="PLACEBO" ;    RUN ;
DATA  EXPOSED ;  INPUT  DAYS  STATUS  TREATMENT  SEX  $  @@ ;
FORMAT  TREATMENT  RX.  ;       DATALINES  ;
179  1  1  F    378  0  1  M      256  1  1  F    355  1  1  M     262  1  1  M
319  1  1  M    256  1  1  F      256  1  1  M    255  1  1  M     171  1  1  F
224  0  1  F    325  1  1  M      225  1  1  F    325  1  1  M     287  1  1  M
217  1  1  F    319  1  1  M      255  1  1  F    264  1  1  M     256  1  1  F
237  0  0  F    291  1  0  M      156  1  0  F    323  1  0  M     270  1  0  M
253  1  0  M    257  1  0  M      206  1  0  F    242  1  0  M     206  1  0  F
157  1  0  F    237  1  0  M      249  1  0  M    211  1  0  F     180  1  0  F
229  1  0  F    226  1  0  F      234  1  0  F    268  0  0  M     209  1  0  F
RUN ;

ODS   GRAPHICS  ON  ;
TITLE1  " FIRST OF 3 ANALYSES  " ;
PROC  LIFETEST  DATA  =  EXPOSED
plots=(survival(atrisk=0 to 1000 by 100 test)
                    loglogs
                    logsurv);
TIME  DAYS  *  STATUS ( 0 )  ;
STRATA   TREATMENT  ;    RUN ;
ODS   GRAPHICS   OFF  ;
```


Summary of the Number of Censored and Uncensored Values					
Stratum	TREATMENT	Total	Failed	Censored	Percent Censored
1	DRUG X	20	18	2	10.00
2	PLACEBO	20	18	2	10.00
Total		40	36	4	10.00

Summary Statistics for Time Variable DAYS

Quartile Estimates				
		95% Confidence Interval		
Percent	Point Estimate	Transform	[Lower	Upper)
75	319.000	LOGLOG	256.000	355.000
50	256.000	LOGLOG	255.000	319.000
25	255.000	LOGLOG	171.000	256.000

Mean	Standard Error
271.131	11.877

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 LIFETEST Procedure

Stratum 1: TREATMENT = DRUG X

Product-Limit Survival Estimates						
DAYS		Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000		1.0000	0	0	0	20
171.000		0.9500	0.0500	0.0487	1	19
179.000		0.9000	0.1000	0.0671	2	18
217.000		0.8500	0.1500	0.0798	3	17
224.000	*	.	.	.	3	16
225.000		0.7969	0.2031	0.0908	4	15
255.000		.	.	.	5	14
255.000		0.6906	0.3094	0.1053	6	13
256.000		.	.	.	7	12
256.000		.	.	.	8	11
256.000		.	.	.	9	10
256.000		0.4781	0.5219	0.1146	10	9
262.000		0.4250	0.5750	0.1135	11	8
264.000		0.3719	0.6281	0.1111	12	7
287.000		0.3187	0.6813	0.1071	13	6
319.000		.	.	.	14	5
319.000		0.2125	0.7875	0.0942	15	4
325.000		.	.	.	16	3
325.000		0.1062	0.8938	0.0710	17	2
355.000		0.0531	0.9469	0.0517	18	1
378.000	*	.	.	.	18	0

Note: The marked survival times are censored observations.

Summary Statistics for Time Variable DAYS

Quartile Estimates				
Percent	Point Estimate	95% Confidence Interval		
		Transform	[Lower	Upper)
75	257.000	LOGLOG	237.000	323.000
50	235.500	LOGLOG	206.000	253.000
25	207.500	LOGLOG	156.000	229.000

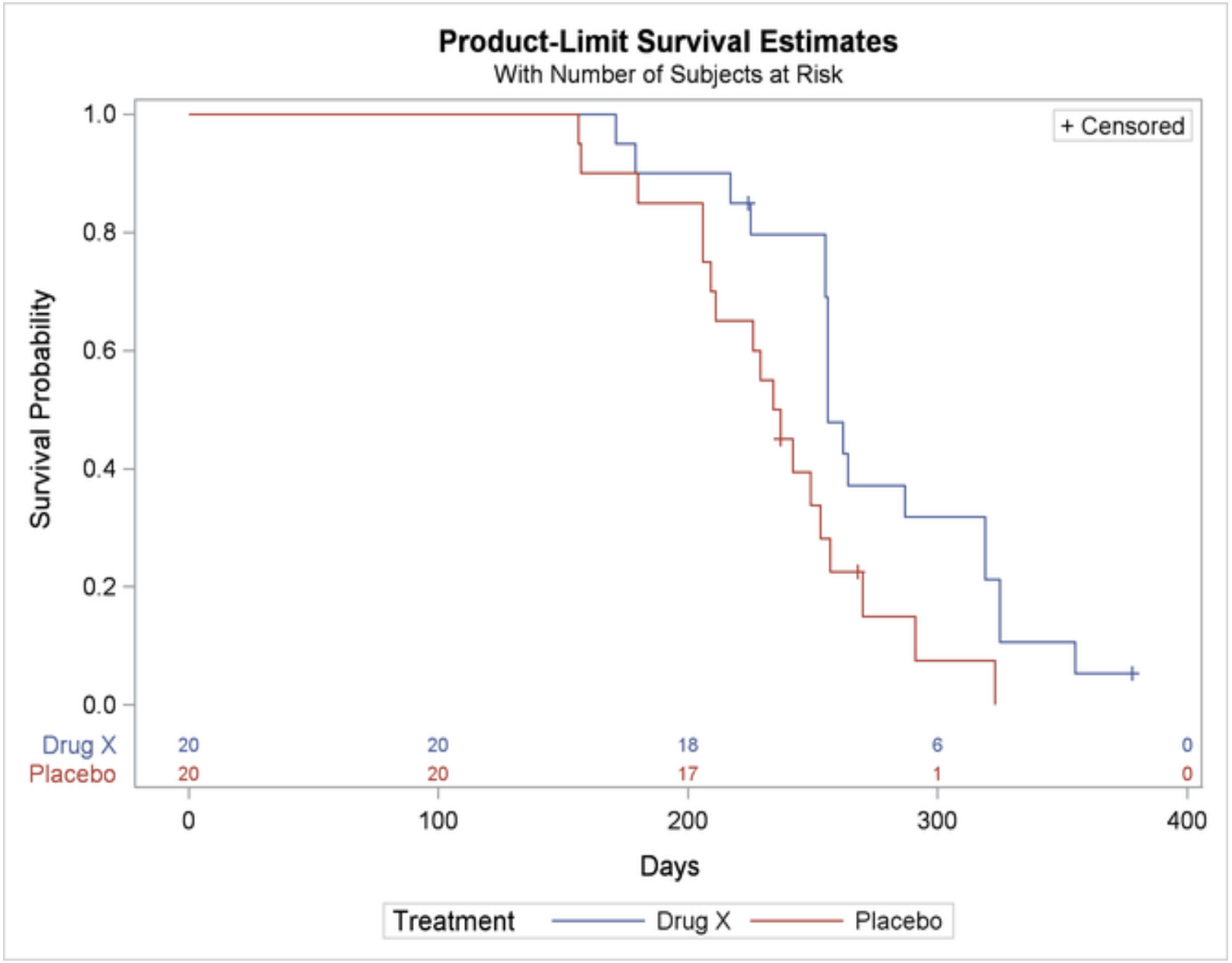
Mean	Standard Error
235.156	10.211

The LIFETEST Procedure

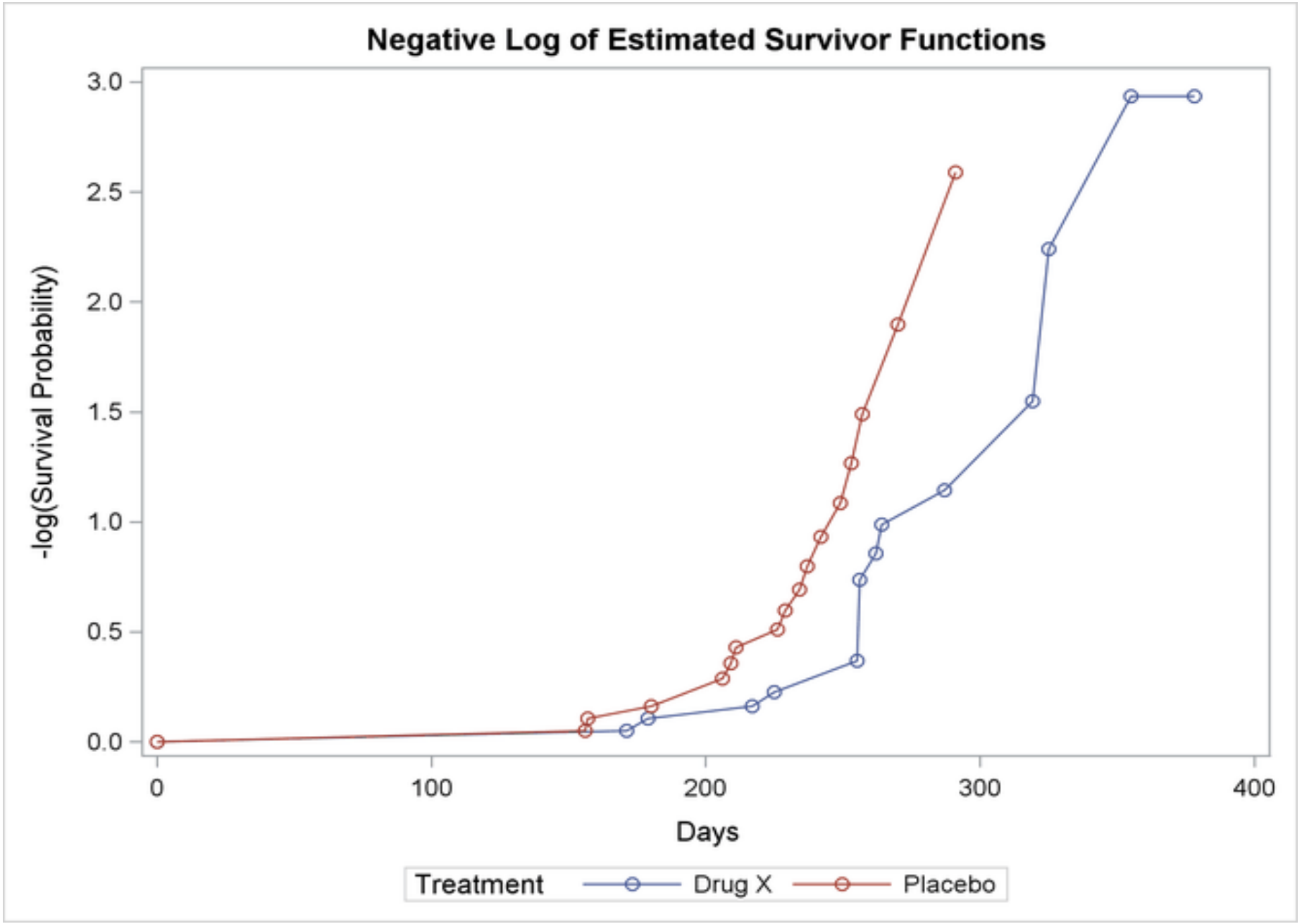
Stratum 2: TREATMENT = PLACEBO

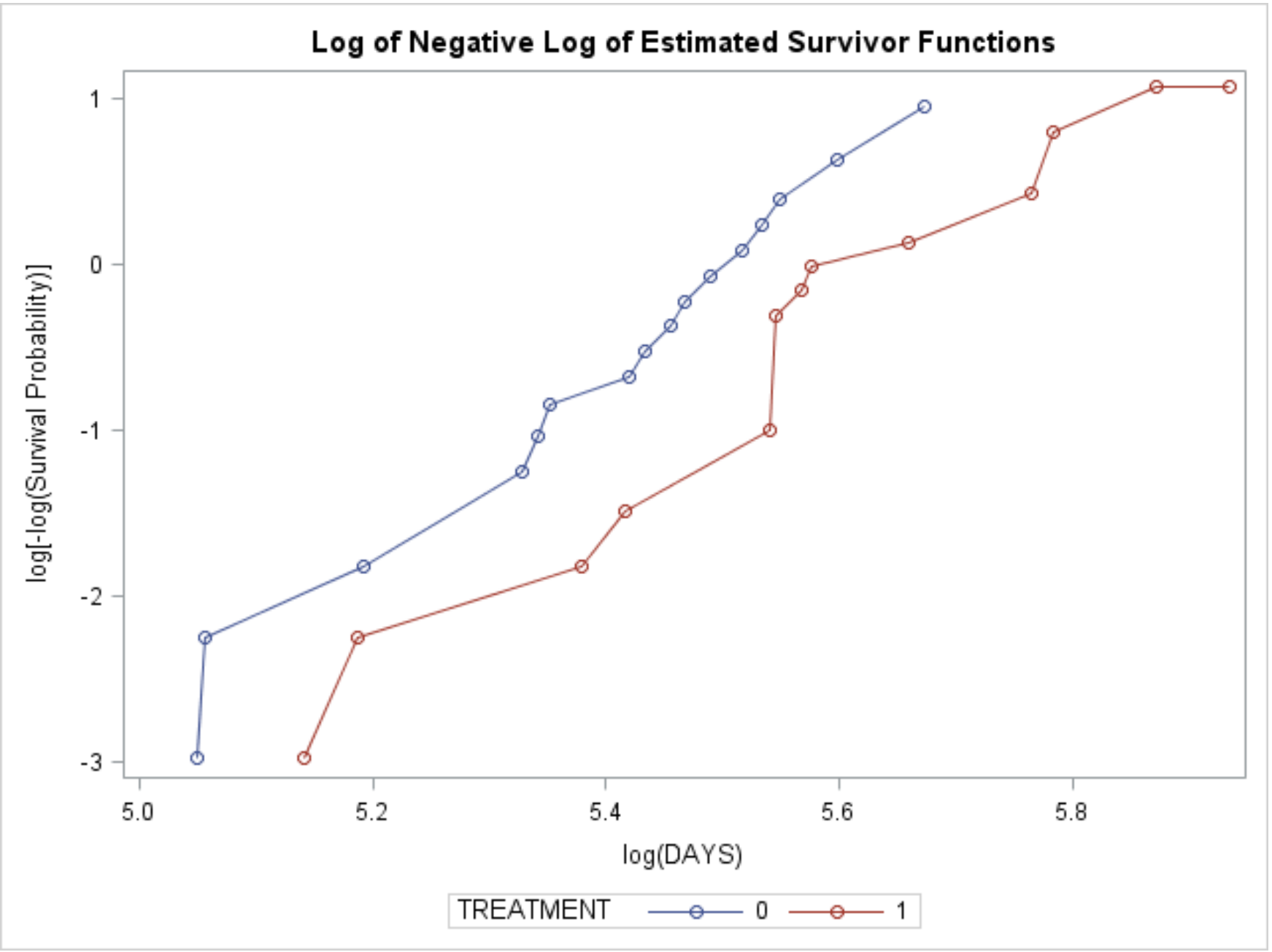
Product-Limit Survival Estimates						
DAYS		Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000		1.0000	0	0	0	20
156.000		0.9500	0.0500	0.0487	1	19
157.000		0.9000	0.1000	0.0671	2	18
180.000		0.8500	0.1500	0.0798	3	17
206.000		.	.	.	4	16
206.000		0.7500	0.2500	0.0968	5	15
209.000		0.7000	0.3000	0.1025	6	14
211.000		0.6500	0.3500	0.1067	7	13
226.000		0.6000	0.4000	0.1095	8	12
229.000		0.5500	0.4500	0.1112	9	11
234.000		0.5000	0.5000	0.1118	10	10
237.000		0.4500	0.5500	0.1112	11	9
237.000	*	.	.	.	11	8
242.000		0.3938	0.6063	0.1106	12	7
249.000		0.3375	0.6625	0.1082	13	6
253.000		0.2813	0.7188	0.1038	14	5
257.000		0.2250	0.7750	0.0971	15	4
268.000	*	.	.	.	15	3
270.000		0.1500	0.8500	0.0891	16	2
291.000		0.0750	0.9250	0.0693	17	1
323.000		0	1.0000	.	18	0

Note: The marked survival times are censored observations.



Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	5.6485	1	0.0175
Wilcoxon	5.0312	1	0.0249
-2Log(LR)	0.1983	1	0.6561


Negative Log of Estimated Survivor Functions
-log(Survival Probability)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
100
200
300
400
Days
Treatment
Drug X
Placebo


Log of Negative Log of Estimated Survivor Functions
log[-log(Survival Probability)]
1
0
-1
-2
-3
5.0
5.2
5.4
5.6
5.8
log(DAYS)
TREATMENT
0
1

RESULTS OF THE TWO-SAMPLE TESTS

Rank tests for homogeneity result in a significant difference between treatments (p =0.018 for LOG-RANK & p =0.025 for WILCOXON test).

DRUG treated rats live significantly longer than those on the PLACEBO.

Because the survival curves for the two treatments differ primarily at longer survival times, and the Wilcoxon test places more weight on short survival times, it becomes less significant than the log-rank test, that is, a larger p value.

TEST	CHI-SQUARE	DF	PR> CHI-SQUARE
LOG-RANK	5.6485	1	0.0175
WILCOXON	5.0312	1	0.0249

FIRST ANALYSIS COMPARING DRUG AND PLACEBO

GROUP	MEDIAN	MEAN	STD ERROR	DF	CHI - SQUARE	P - VALUE
DRUG	256.0	271.13	11.877	1	5.649	0.0175 *
PLACEBO	235.5	235.16	10.211	1	5.031	0.0249 #

* LOG RANK TEST # WILCOXON TEST

Suppose male and female rats have different survival rates.

Therefore we test the treatment effect adjusted for the SEX effect. The variable SEX in the STRATA statement is a stratifying variable and the main variable TREATMENT is the GROUP= option.

The test statistics for the TREATMENT variable are computed by pooling over the strata defined by the values of SEX, thus controlling for the SEX effect. The NOTABLE option is added to avoid estimating a survival curve for each sex.

If variable SEX is associated with group variable TREATMENT then including it in the STRATA statement as a predictor will lower the residual variation and **lower** the p value for the TREATMENT variable.

```
TITLE1  "  SECOND ANALYSIS INCLUDES SEX VARIABLE  "  ;
PROC  LIFETEST  DATA = EXPOSED  NOTABLE ;
TIME  DAYS * STATUS( 0 )  ;
STRATA  SEX  /  GROUP = TREATMENT  ;  RUN  ;
```

SECOND ANALYSIS

STRATIFIED TEST OF EQUALITY OVER GROUP

TEST	CHI SQUARE	DF	PR>CHI-SQUARE
LOG-RANK	7.2466	1	0.0071
WILCOXON	5.9179	1	0.0150

LESSON LEARNED

You may want to show that an EXPOSURE variable such as DRUG or DIET is related to an OUTCOME variables such as DEATH or change in BLOOD PRESSURE.

You may include in your analysis a variable such as AGE or SEX… for 2 reasons!

- ❑ If it is a predictor of the OUTCOME variable then including it in your model will reduce the background random variation thereby reducing the standard error of the OUTCOME variable and reduce the p value.

- ❑ If the predictor variable is also related to the EXPOSURE variable the size of the impact of the EXPOSURE variable may become larger or smaller. The predictor variable is then called a CONFOUNDER.

End of Lecture 2

Next up in Part 7 Lecture 1: Count Data!