TMA4275 LIFETIME ANALYSIS

Slides 9 - DRAFT: Parametric inference in lifetime models (cont.)

Bo Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

http://www.math.ntnu.no/∼bo/
bo@math.ntnu.no

NTNU, Spring 2014

Why parametric models?

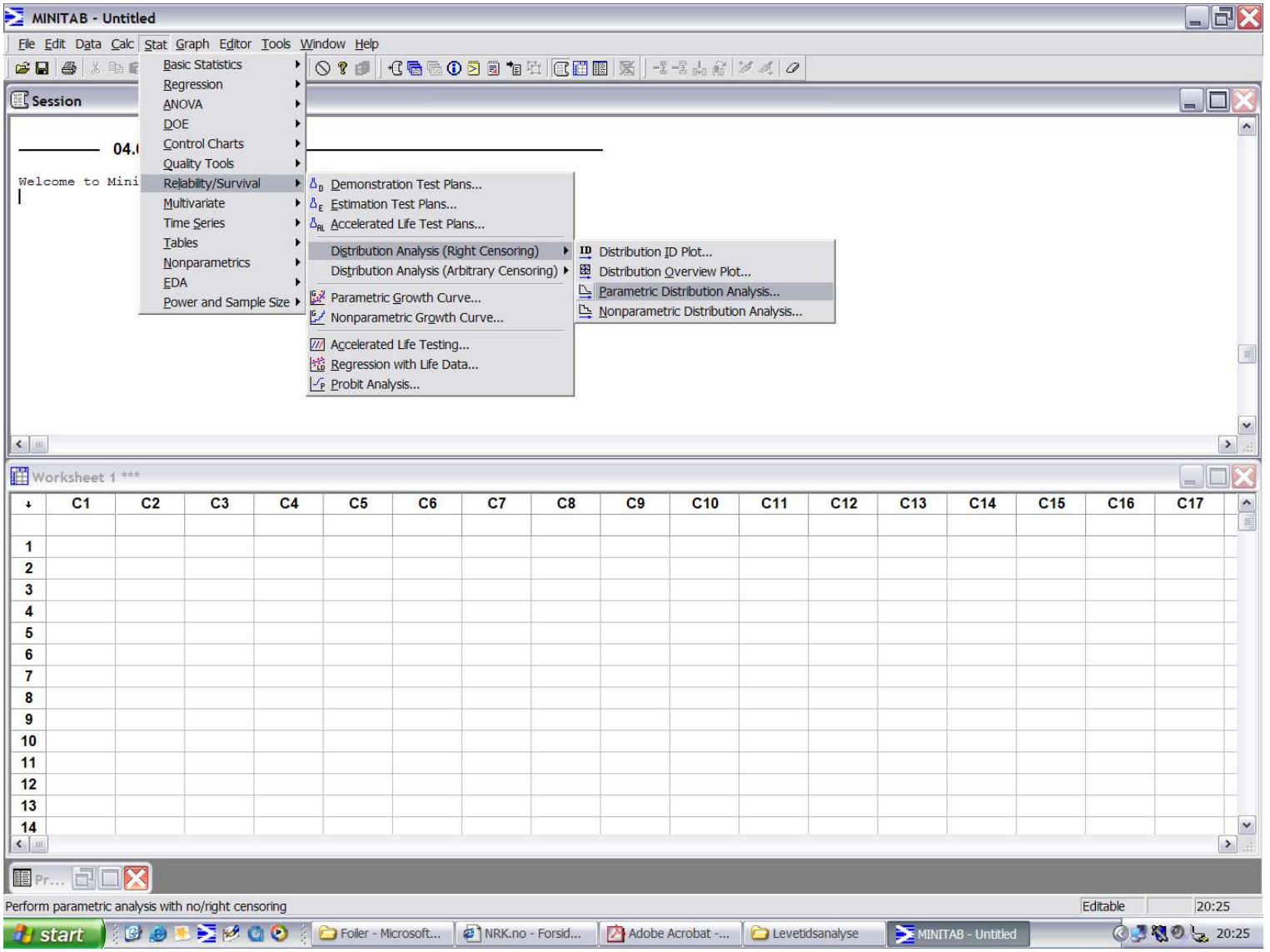
- Complements nonparametric techniques.

- Parametric models can be described concisely with just a few parameters, instead of having to report an entire curve.

- It is possible to use a parametric model to extrapolate (in time) to the lower or upper tail of a distribution.

- Parametric models provide smooth estimates of failure-time distributions. In practice it is often useful to compare various parametric and nonparametric analyses of a data set.

PARAMETRIC LIFETIME ANALYSIS IN MINITAB



DATA OPTIONS

RIGHT CENSORING:

Y_i	δ_i
Observed time	Cens. status 1: Lifetime 0: Censoring

ARBITRARY CENSORING:

Start variable A_i	End variable B_i	
1.7	1.7	Exact lifetime 1.7
2.0	*	Right censoring at time 2.0, i.e. lifetime is > 2.0
*	0.5	Left censoring at time 0.5, i.e. lifetime is < 0.5
1.0	1.5	Interval censoring: Lifetime between 1.0 and 1.5

LIKELIHOOD CONTRIBUTION

Obs. type	Start variable A_i	End variable B_i	Likelihood contribution
Exact lifetime	1.7	1.7	$f(1.7;\theta)$
Right cens.	2.0	*	$1 - F(2.0;\theta)$
Left cens.	*	0.5	$F(0.5;\theta)$
Interval cens.	1.0	1.5	$F(1.5;\theta) - F(1.0;\theta)$

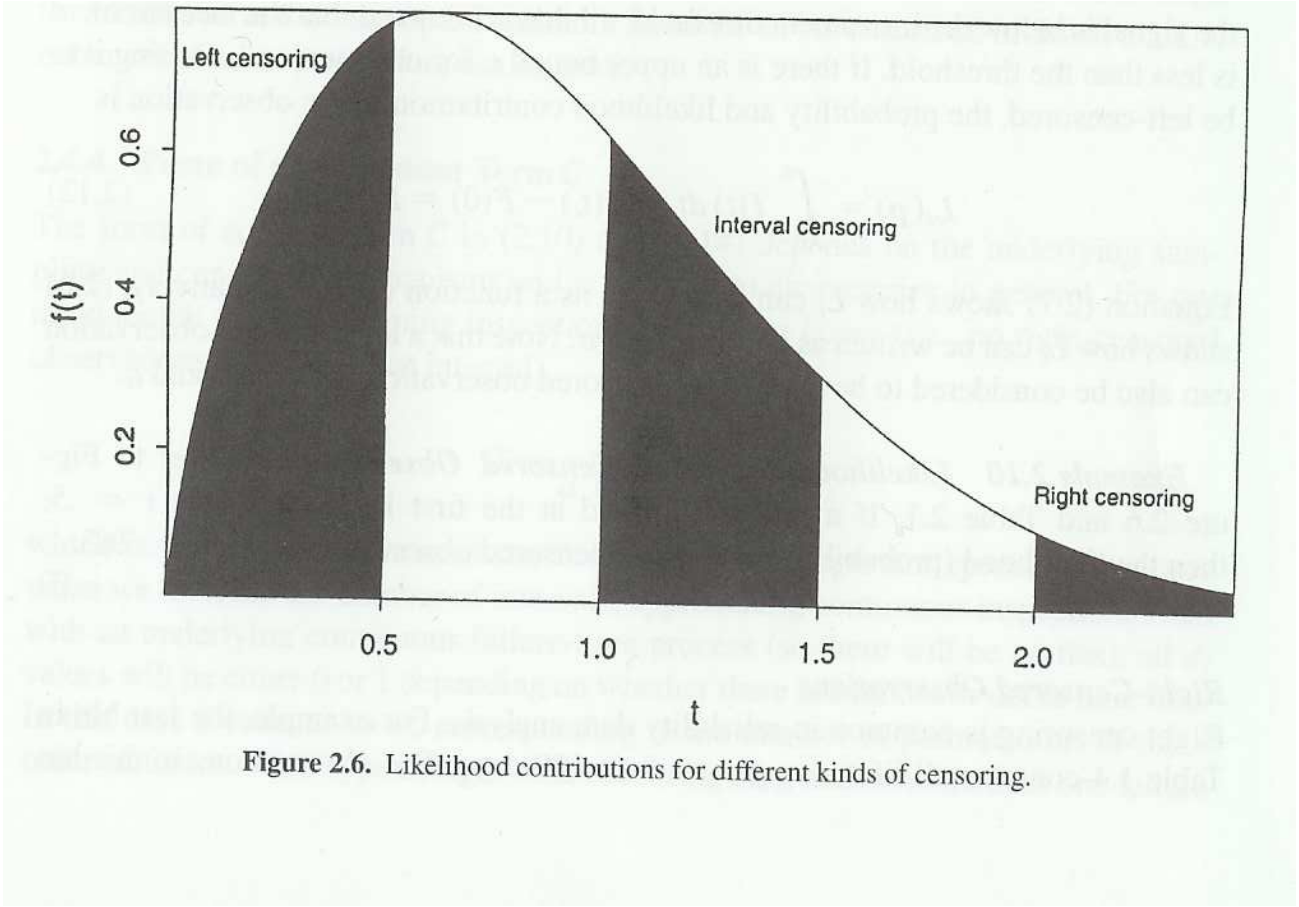
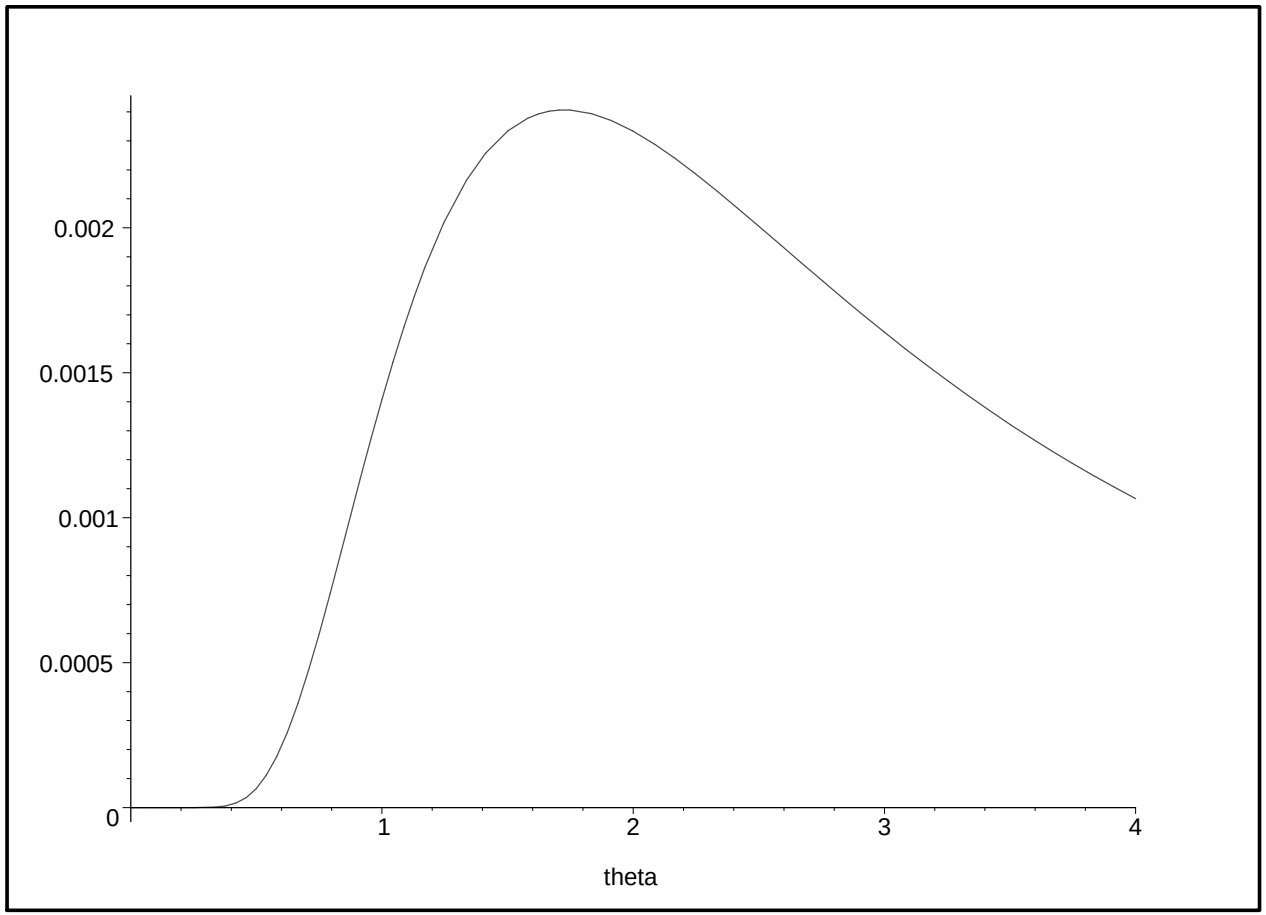


Figure 2.6. Likelihood contributions for different kinds of censoring.

LIKELIHOOD FOR MODEL $f(t;\theta) = (1/\theta)e^{-t/\theta}$

$$L(\theta) = (\frac{1}{\theta}e^{-1.7/\theta}) \cdot (e^{-2.0/\theta}) \cdot (1 - e^{-0.5/\theta}) \cdot (e^{-1.0/\theta} - e^{-1.5/\theta})$$



Maximum likelihood estimate: $\widehat{\theta} = 1.725$

ARBITRARY CENSORED DATA: MINITAB OUTPUT

Distribution Analysis, Start = A and End = B

Variable Start: A End: B

Censoring Information	Count
Uncensored value	1
Right censored value	1
Interval censored value	1
Left censored value	1

Estimation Method: Maximum Likelihood

Distribution: Exponential

Parameter Estimates

Parameter	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Mean	1,72529	0,998421	0,554978	5,36353

Log-Likelihood = -6,029

Goodness-of-Fit
Anderson-Darling (adjusted) = 4,933

Characteristics of Distribution

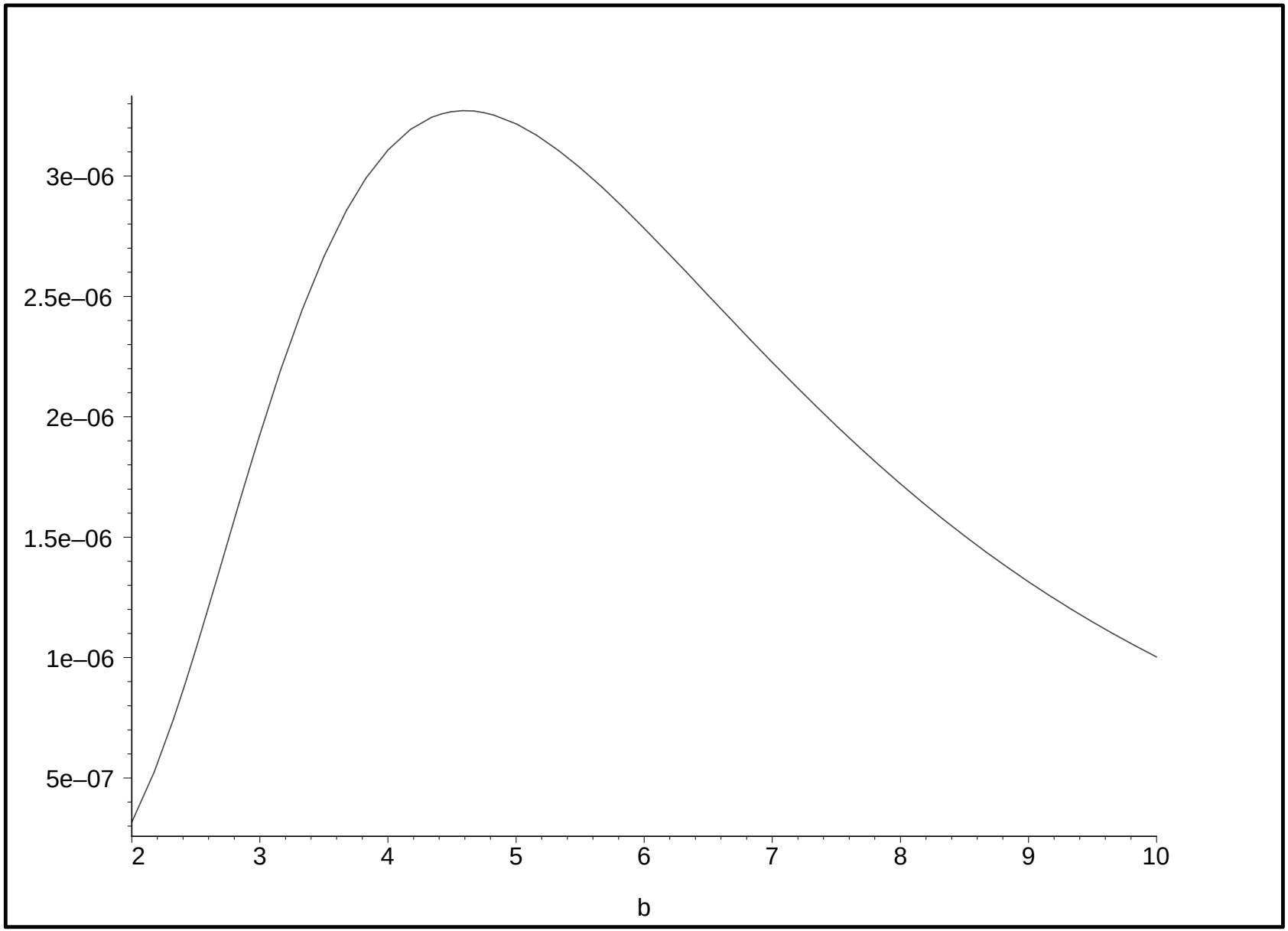
	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Mean(MTTF)	1,72529	0,998421	0,554978	5,36353
Standard Deviation	1,72529	0,998421	0,554978	5,36353
Median	1,19588	0,692053	0,384682	3,71771
First Quartile(Q1)	0,496336	0,287228	0,159657	1,54299
Third Quartile(Q3)	2,39177	1,38411	0,769363	7,43543
Interquartile Range(IQR)	1,89543	1,09688	0,609706	5,89244

EXAMPLE: RIGHT CENSORED DATA

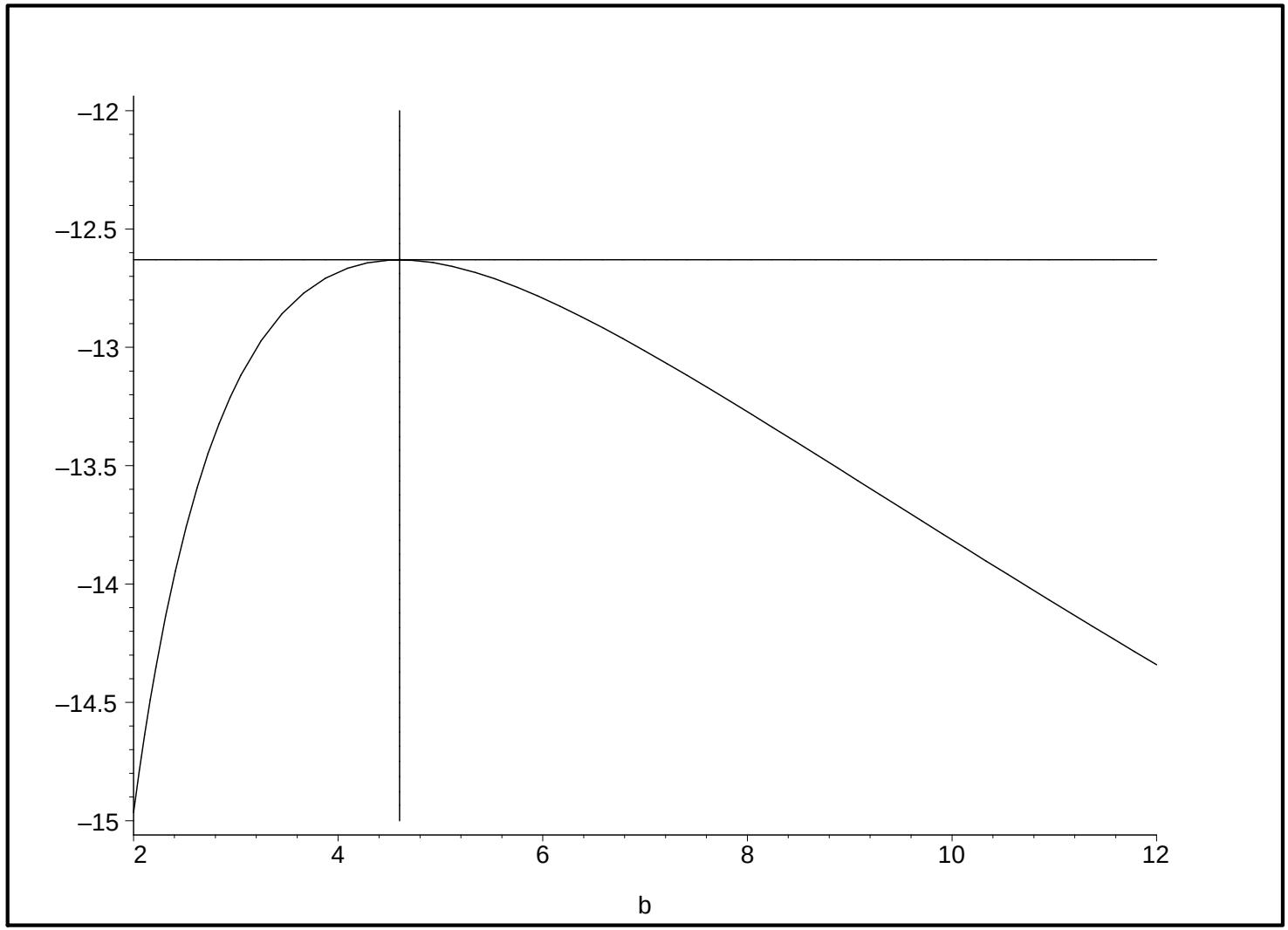
Worksheet 1 ***

↓	C1	C2
	Y	D
1	0,6	0
2	0,8	1
3	2,1	1
4	3,2	1
5	3,3	0
6	4,4	1
7	8,6	1

LIKELIHOOD FUNCTION FOR MODEL $f(t;\theta) = (1/\theta)e^{-t/\theta}$



LOG-LIKELIHOOD FUNCTION



Maximum likelihood estimate: $\widehat{\theta} = 4.6$

Distribution Analysis: Y

Variable: Y

Censoring Information	Count
Uncensored value	5
Right censored value	2

Censoring value: D = 0

Estimation Method: Maximum Likelihood

Distribution: Exponential

Parameter Estimates

Parameter	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Mean	4,6	2,05718	1,91465	11,0516

Log-Likelihood = -12,630

Goodness-of-Fit
Anderson-Darling (adjusted) = 3,767

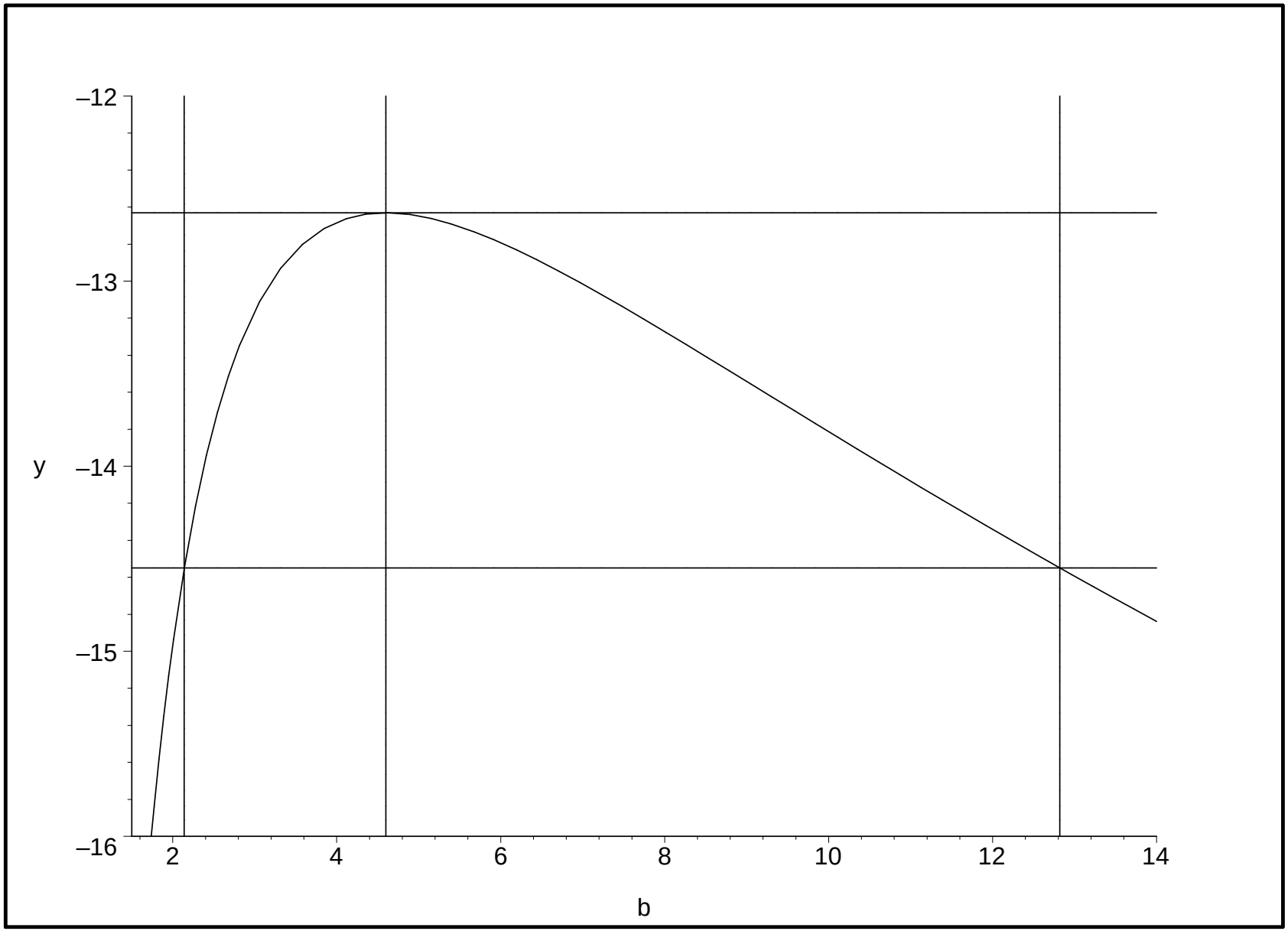
Characteristics of Distribution

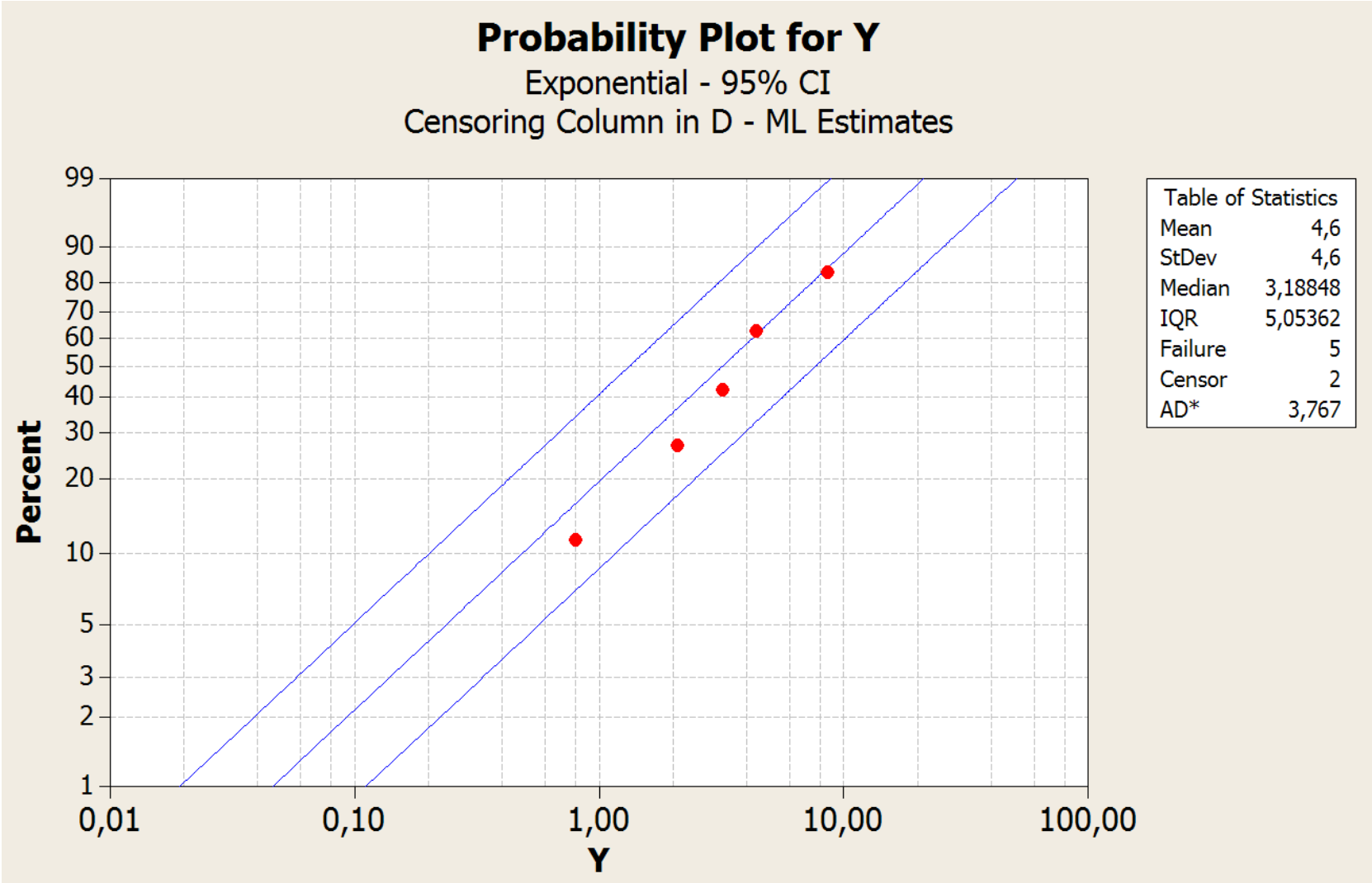
	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Mean(MTTF)	4,6	2,05718	1,91465	11,0516
Standard Deviation	4,6	2,05718	1,91465	11,0516
Median	3,18848	1,42593	1,32713	7,66041
First Quartile(Q1)	1,32334	0,591815	0,550810	3,17936
Third Quartile(Q3)	6,37695	2,85186	2,65427	15,3208
Interquartile Range(IQR)	5,05362	2,26005	2,10346	12,1415

Table of Percentiles

Percent	Percentile	Standard Error	95,0% Normal CI Lower	Upper
1	0,0462315	0,0206754	0,0192429	0,111073
2	0,0929325	0,0415607	0,0386811	0,223273
3	0,140112	0,0626601	0,0583187	0,336624
4	0,187781	0,0839783	0,0781597	0,451150
5	0,235949	0,105520	0,0982086	0,566875
6	0,284627	0,127289	0,118470	0,683825
7	0,333825	0,149291	0,138947	0,802025
8	0,383555	0,171531	0,159646	0,921504
9	0,433829	0,194014	0,180572	1,04229
10	0,484658	0,216746	0,201728	1,16441
20	1,02646	0,459047	0,427241	2,46610
30	1,64070	0,733745	0,682907	3,94184
40	2,34980	1,05086	0,978051	5,64546
50	3,18848	1,42593	1,32713	7,66041
60	4,21494	1,88498	1,75437	10,1265
70	5,53827	2,47679	2,30518	13,3059
80	7,40341	3,31091	3,08151	17,7869
90	10,5919	4,73684	4,40864	25,4473
91	11,0765	4,95358	4,61037	26,6117
92	11,6184	5,19588	4,83588	27,9134
93	12,2326	5,47058	5,09155	29,3892
94	12,9417	5,78770	5,38669	31,0928
95	13,7804	6,16277	5,73577	33,1078
96	14,8068	6,62182	6,16301	35,5739
97	16,1302	7,21363	6,71382	38,7532
98	17,9953	8,04775	7,49015	43,2343
99	21,1838	9,47368	8,81728	50,8947

LOG-LIKELIHOOD FUNCTION




Probability Plot for Y
Exponential - 95% CI
Censoring Column in D - ML Estimates
Percent
99
90
80
70
60
50
40
30
20
10
5
3
2
1
0,01
0,10
1,00
10,00
100,00
Y
Table of Statistics
Mean 4,6
StDev 4,6
Median 3,18848
IQR 5,05362
Failure 5
Censor 2
AD* 3,767

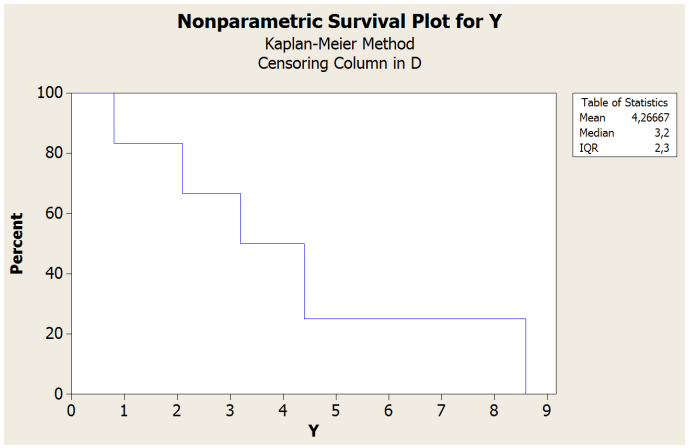
Nonparametric and parametric survival plots

for data from exponential example

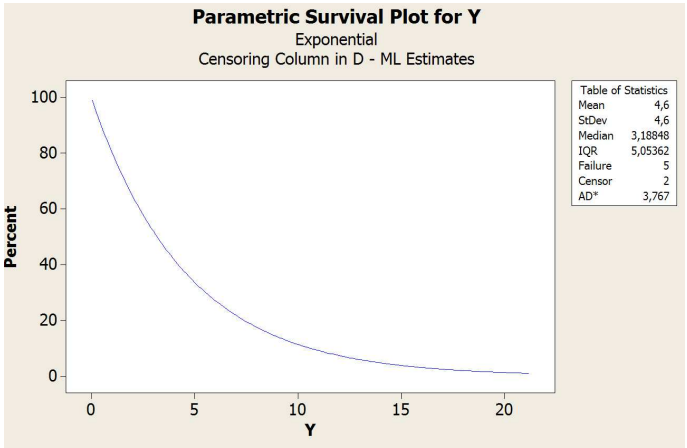
Worksheet 1 ***

↓	C1	C2
	Y	D
1	0,6	0
2	0,8	1
3	2,1	1
4	3,2	1
5	3,3	0
6	4,4	1
7	8,6	1

,



,



RIGHT CENSORED DATA
EXPONENTIAL AND WEIBULL MODELS

Row	C1	C2
1	0,35	1
2	0,50	0
3	0,75	0
4	1,00	1
5	1,30	1
6	1,80	1
7	3,00	0
8	3,15	0
9	4,85	0
10	5,50	1
11	5,50	0
12	6,25	0

Variable: C1
Censoring Information
Count
Uncensored value 5
Right censored value 7
Censoring value: C2 = 0

Estimation Method: Maximum Likelihood Distribution: Exponential

Parameter Estimates

		Standard	95,0% Normal CI	
Parameter	Estimate	Error	Lower	Upper
Shape	1,00000			
Scale	6,790	3,037	2,826	16,313

Log-Likelihood = -14,577

WEIBULL MODEL

Variable: C1

Censoring Information
Count
Uncensored value 5
Right censored value 7
Censoring value: C2 = 0

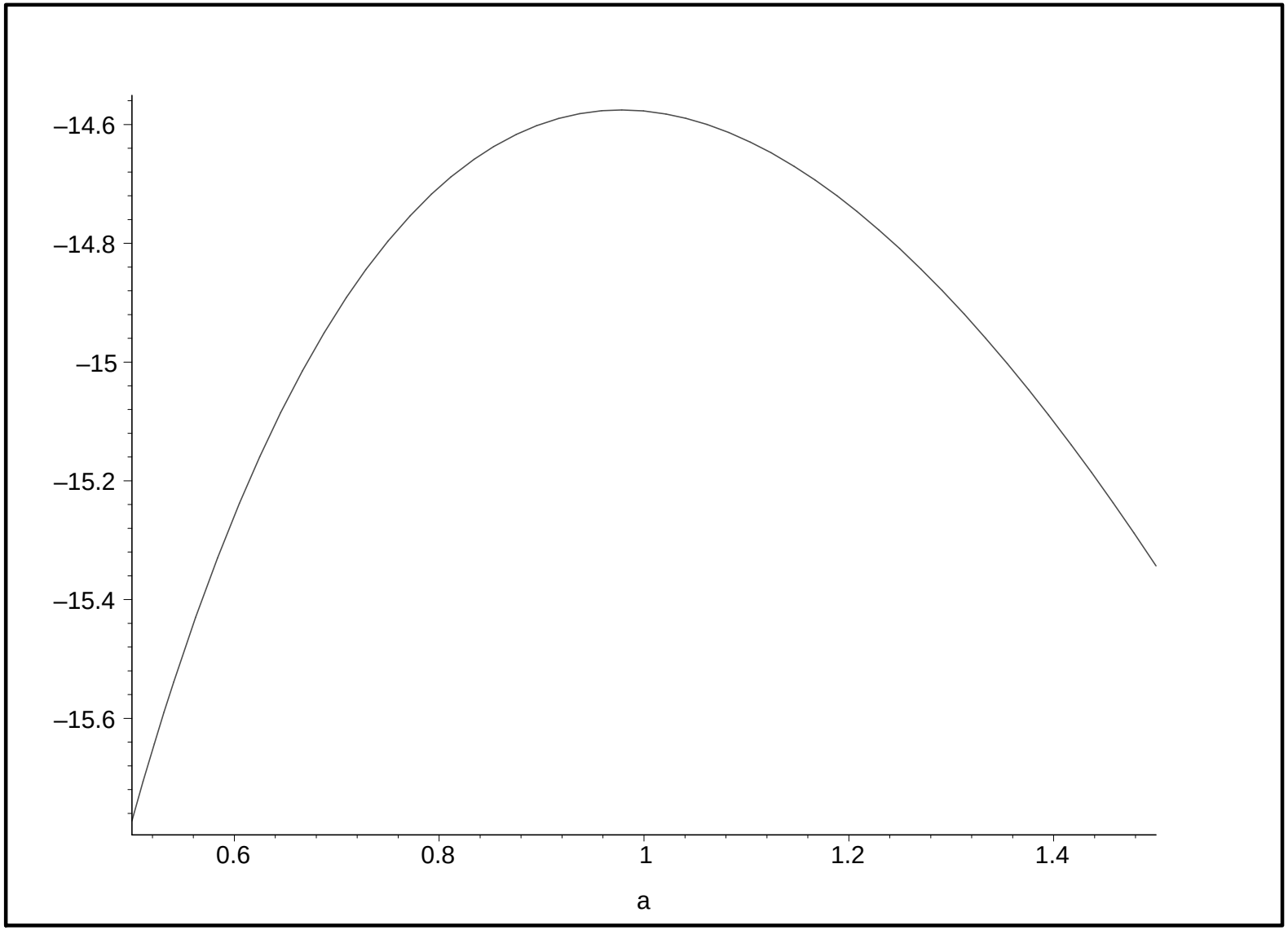
Estimation Method: Maximum Likelihood Distribution: Weibull

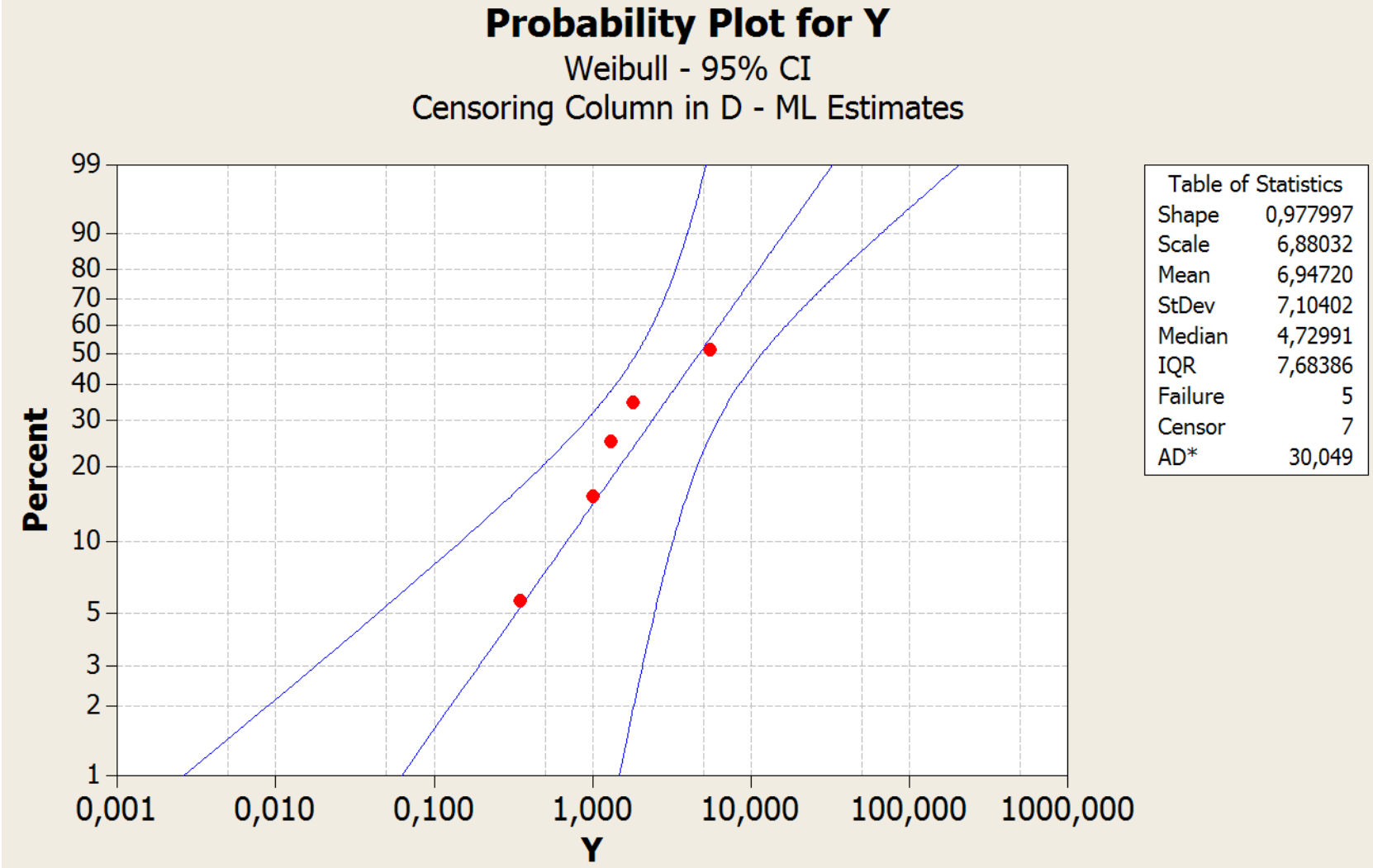
Parameter Estimates

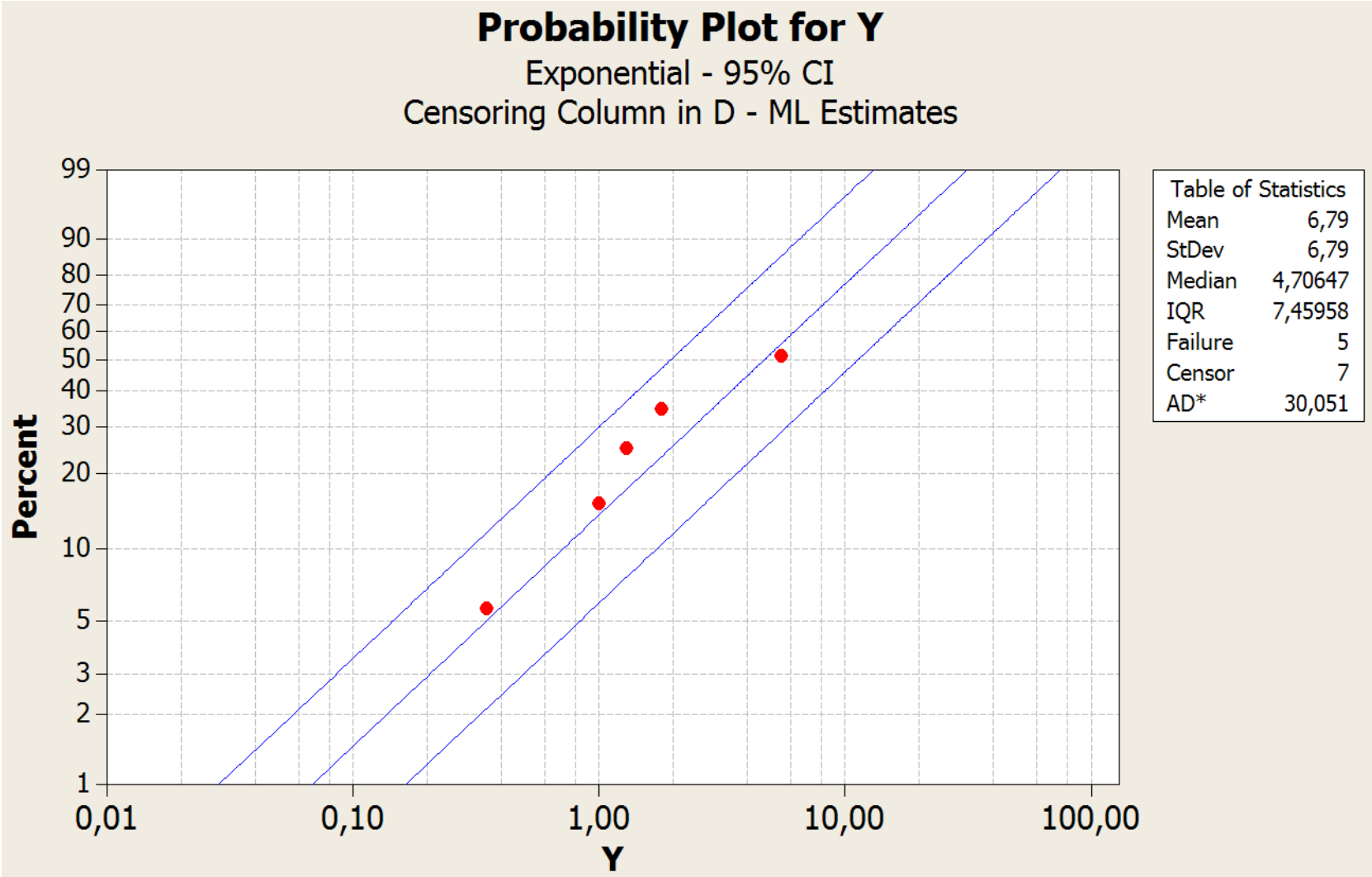
		Standard	95,0% Normal CI	
Parameter	Estimate	Error	Lower	Upper
Shape	0,9780	0,3694	0,4665	2,0504
Scale	6,880	3,517	2,526	18,740

Log-Likelihood = -14,576

PROFILE LOG LIKELIHOOD FOR SHAPE PARAMETER α IN WEIBULL DISTRIBUTION




Probability Plot for Y
Weibull - 95% CI
Censoring Column in D - ML Estimates
Percent
99
90
80
70
60
50
40
30
20
10
5
3
2
1
0,001
0,010
0,100
1,000
10,000
100,000
1000,000
Y
Table of Statistics
Shape 0,977997
Scale 6,88032
Mean 6,94720
StDev 7,10402
Median 4,72991
IQR 7,68386
Failure 5
Censor 7
AD* 30,049


Probability Plot for Y
Exponential - 95% CI
Censoring Column in D - ML Estimates
Percent
99
90
80
70
60
50
40
30
20
10
5
3
2
1
0,01
0,10
1,00
10,00
100,00
Y
Table of Statistics
Mean 6,79
StDev 6,79
Median 4,70647
IQR 7,45958
Failure 5
Censor 7
AD* 30,051

Shock absorber data

Y = kilometers to failure, F = failure mode (0 is censoring)

Row	Y	F			
1	6700	1	19	14300	1
2	6950	0	20	17520	1
3	7820	0	21	17540	0
4	8790	0	22	17890	0
5	9120	2	23	18450	0
6	9660	0	24	18960	0
7	9820	0	25	18980	0
8	11310	0	26	19410	0
9	11690	0	27	20100	2
10	11850	0	28	20100	0
11	11880	0	29	20150	0
12	12140	0	30	20320	0
13	12200	1	31	20900	2
14	12870	0	32	22700	1
15	13150	2	33	23490	0
16	13330	0	34	26510	1
17	13470	0	35	27410	0
18	14040	0	36	27490	1
			37	27890	0
			38	28100	0

Shock Absorber Failure Data

First reported in O'Connor (1985).

- Failure times, in number of kilometers of use, of vehicle shock absorbers.
- Two failure modes, denoted by M1 and M2.
- One might be interested in the distribution of time to failure for mode M1, mode M2, or in the overall failure-time distribution of the part.

  Here we do not differentiate between modes M1 and M2. We will estimate the distribution of time to failure by either mode M1 or M2.

Shock absorber data

Estimation Method: Maximum Likelihood Distribution: Lognormal

Parameter Estimates

Parameter	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Location	10,1448	0,144175	9,86219	10,4273
Scale	0,530068			
		0,112683	0,349447	0,804047

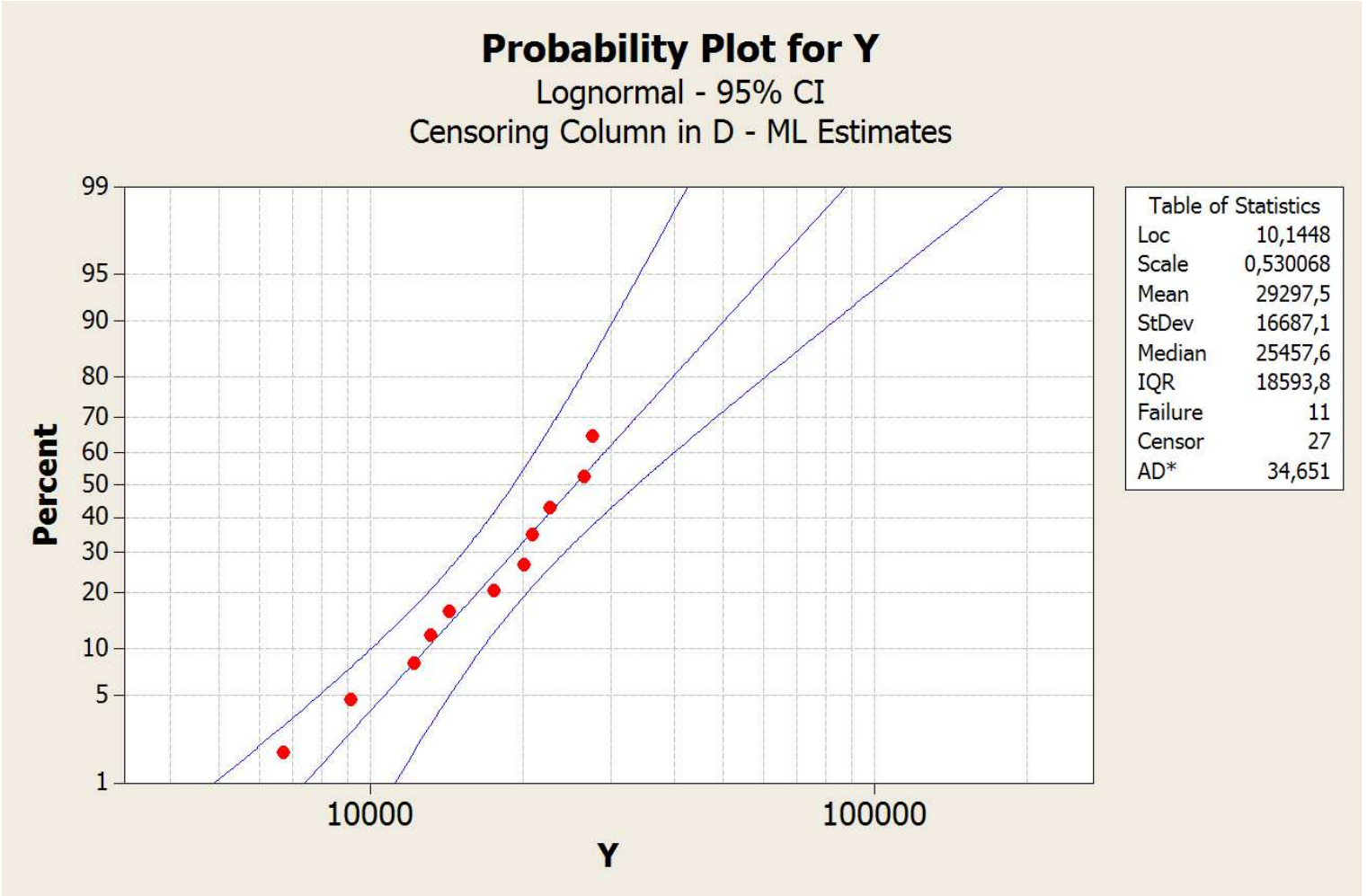
Log-Likelihood = -124,609

Goodness-of-Fit Anderson-Darling (adjusted) = 34,651

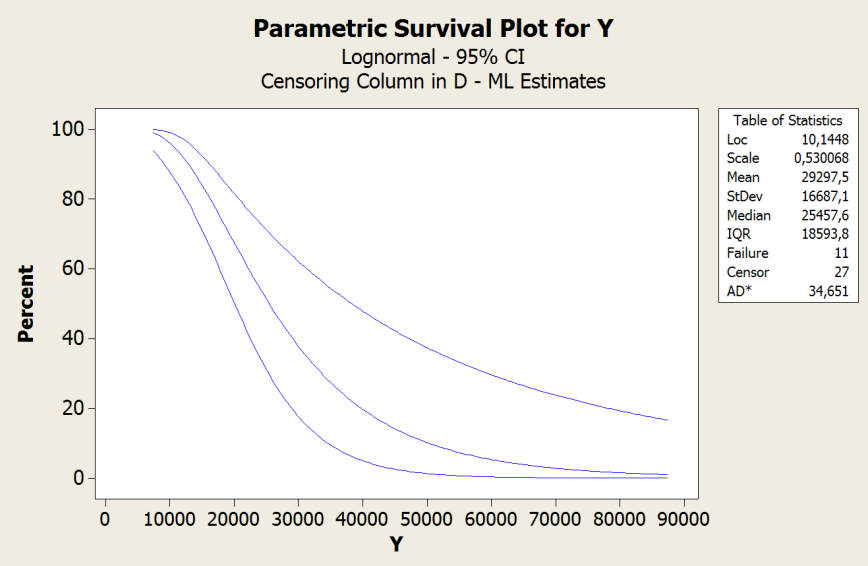
Characteristics of Distribution

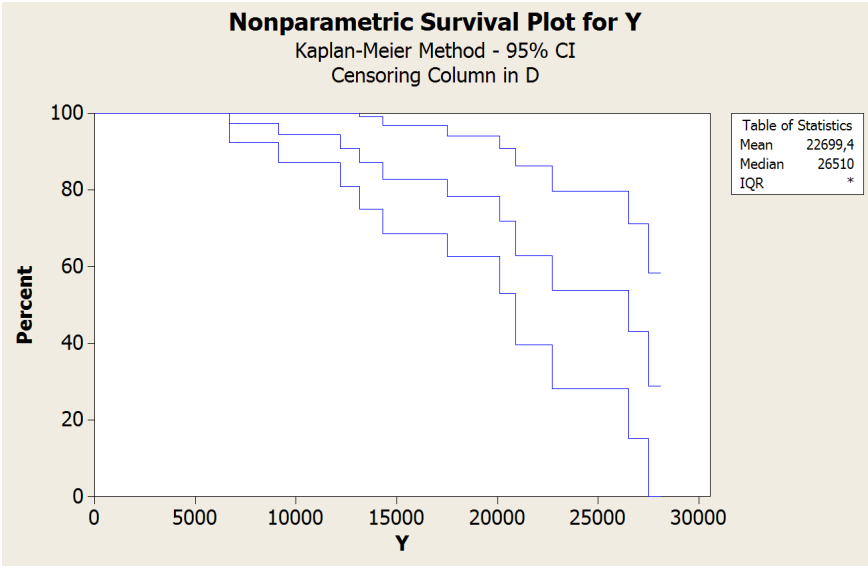
	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Mean(MTTF)	29297,5	5455,91	20338,3	42203,2
Standard Deviation	16687,1	6787,01	7519,35	37032,5
Median	25457,6	3670,36	19190,9	33770,7
First Quartile(Q1)	17805,2	2062,96	14188,1	22344,4
Third Quartile(Q3)	36399,0	7252,61	24631,2	53789,0
Interquartile Range(IQR)	18593,8	6115,60	9758,96	35426,9

Shock absorber data

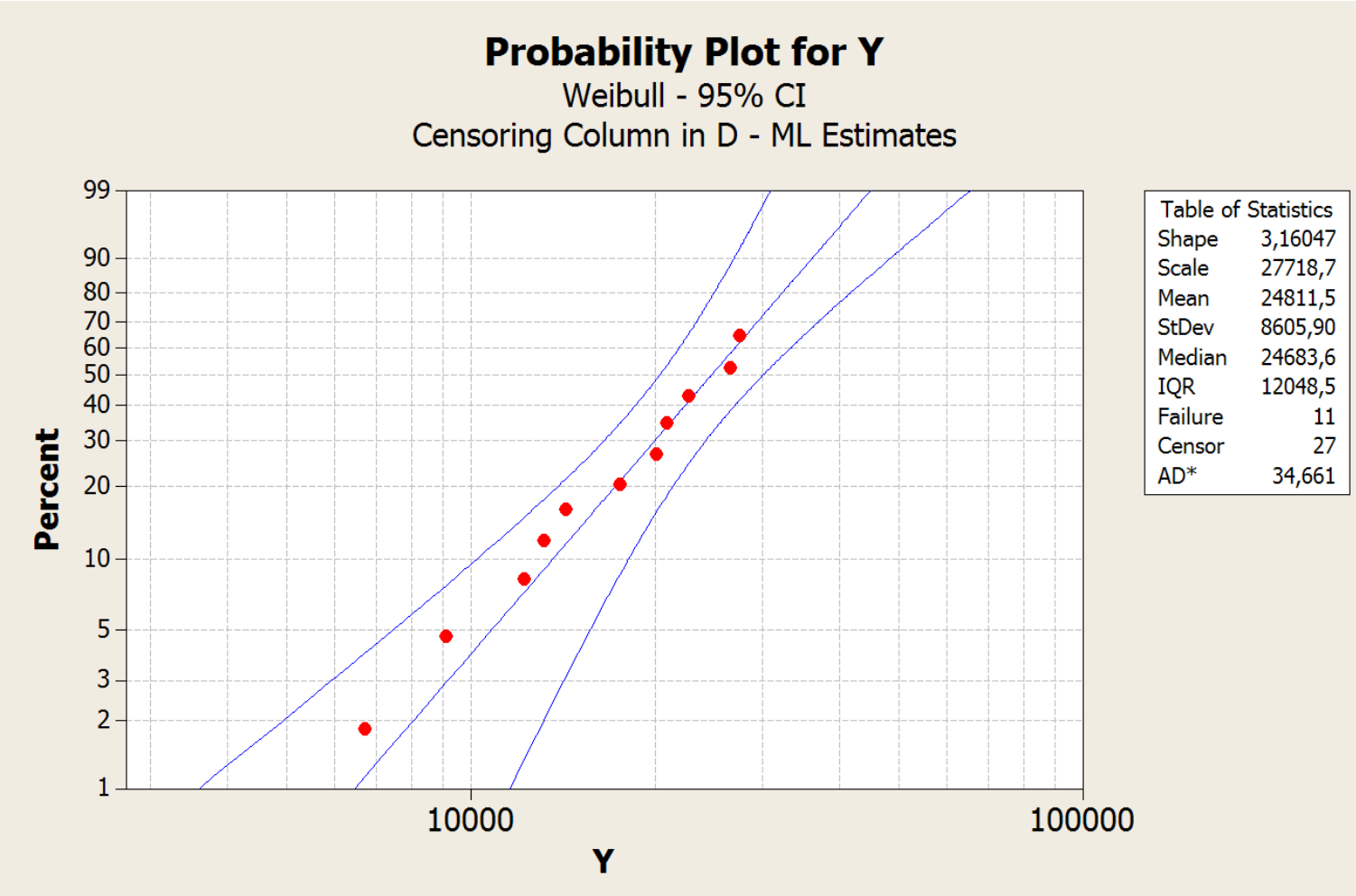


Shock absorber data

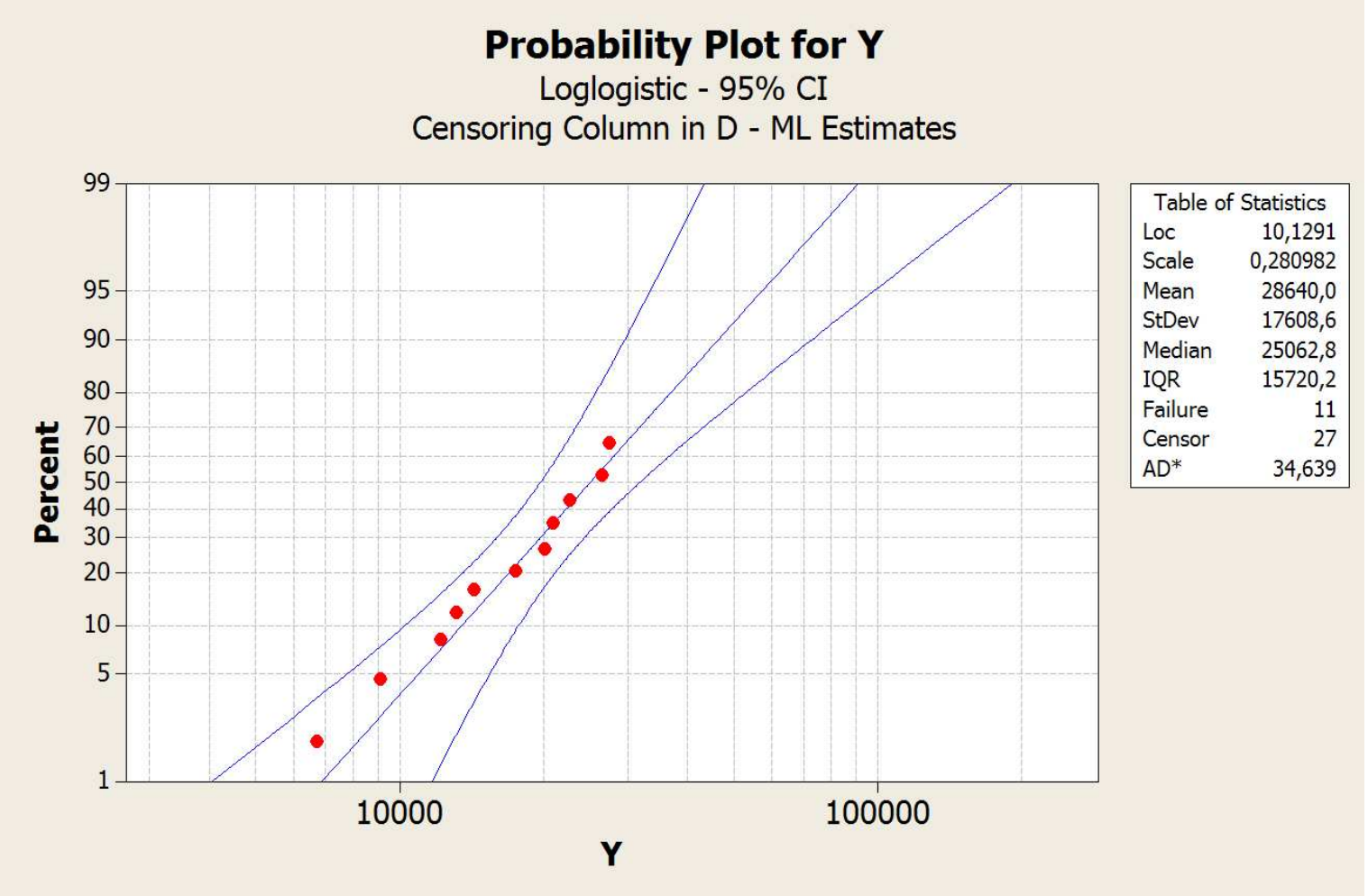




Shock absorber data



Shock absorber data



Shock absorber data:

Results for loglogistic (left), lognormal (middle), Weibull (right)

Table of Statistics	
Loc	10,1291
Scale	0,280982
Mean	28640,0
StDev	17608,6
Median	25062,8
IQR	15720,2
Failure	11
Censor	27
AD*	34,639

Table of Statistics	
Loc	10,1448
Scale	0,530068
Mean	29297,5
StDev	16687,1
Median	25457,6
IQR	18593,8
Failure	11
Censor	27
AD*	34,651

Table of Statistics	
Shape	3,16047
Scale	27718,7
Mean	24811,5
StDev	8605,90
Median	24683,6
IQR	12048,5
Failure	11
Censor	27
AD*	34,661

Confidence Interval for the Mean Life of a New Insulating Material

- A life test for a new insulating material used 25 specimens which were tested simultaneously at a high voltage of 30 kV.
- The test was run until 15 of the specimens failed.
- The 15 failure times (hours) were recorded as:

  1.08, 12.20, 17.80, 19.10, 26.00, 27.90, 28.20, 32.20, 35.90, 43.50, 44.00, 45.20, 45.70, 46.30, 47.80

  Then $TT = 1.08 + \cdots + 47.80 + 10 \times 47.80 = 950.88$ hours.

- The ML estimate of θ and a 95% confidence interval are:

$$\begin{aligned} \widehat{\theta} &= 950.88/15 = 63.392 \text{ hours} \\ \left[\underline{\theta}, \tilde{\theta}\right] &= \left[\frac{2(950.88)}{\chi^2_{(.975;30)}}, \frac{2(950.88)}{\chi^2_{(.025;30)}}\right] = \left[\frac{1901.76}{46.98}, \frac{1901.76}{16.79}\right] \\ &= [40.48, \quad 113.26]. \end{aligned}$$

Pike (1966) cancer data for rats

Row	Y	D
1	143	1
2	164	1
3	188	1
4	188	1
5	190	1
6	192	1
7	206	1
8	209	1
9	213	1
10	216	1
11	220	1
12	227	1
13	230	1
14	234	1
15	246	1
16	265	1
17	304	1
18	216	0
19	244	0

Pike (1966) cancer data for rats: 3-parameter Weibull

Distribution Analysis: C1

Variable: C1

Censoring Information	Count
Uncensored value	17
Right censored value	2

Censoring value: C2 = 0

Estimation Method: Maximum Likelihood

Distribution: 3-Parameter Weibull

Parameter Estimates

Parameter	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Shape	2,71148	1,05876	1,26135	5,82878
Scale	108,383	32,5734	60,1367	195,335
Threshold	122,026	28,6924	65,7898	178,262

Log-Likelihood = -87,324

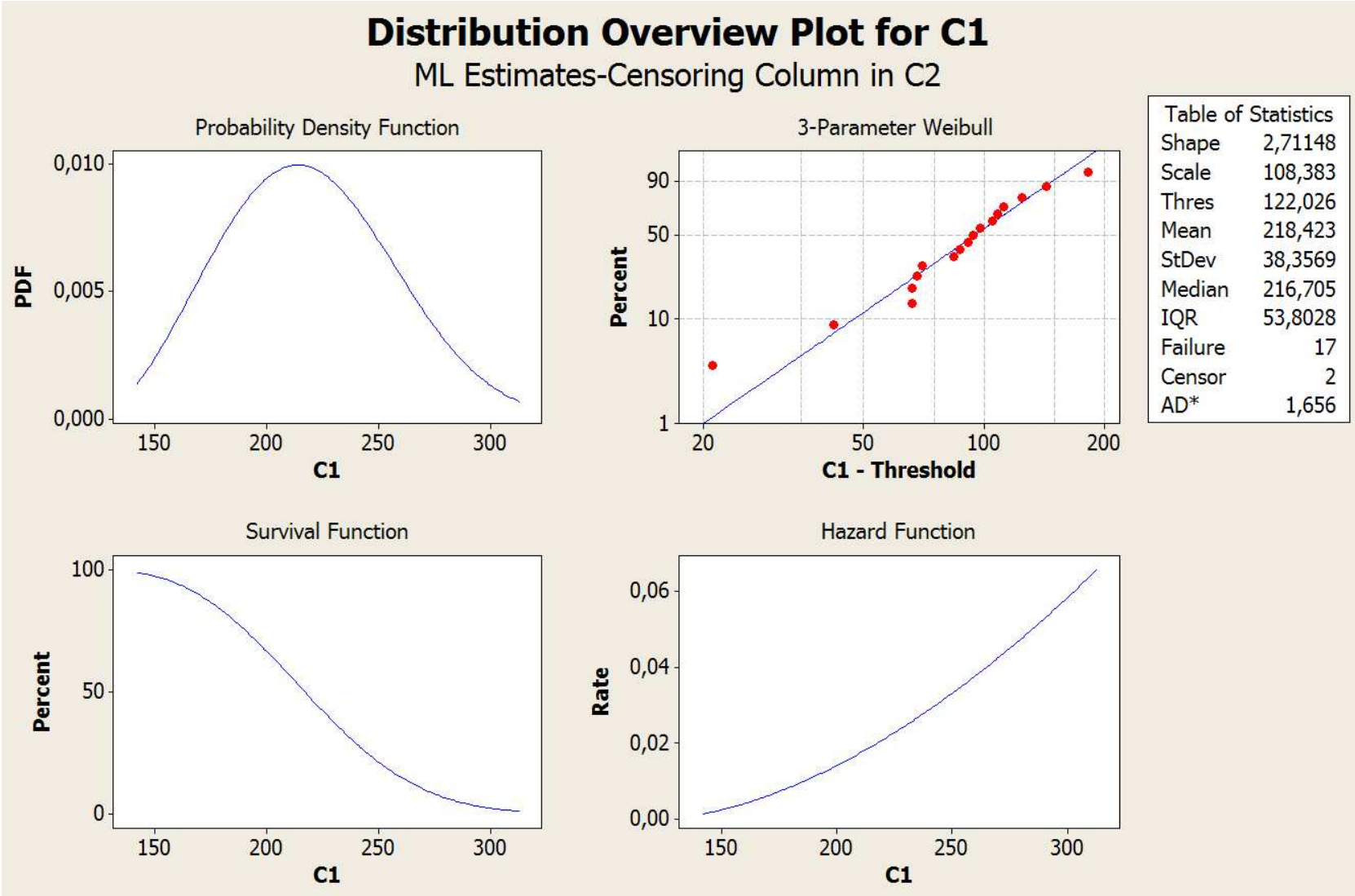
Goodness-of-Fit
Anderson-Darling (adjusted) = 1,656

Characteristics of Distribution

	Estimate	Standard Error	95,0% Normal CI Lower	Upper
Mean(MTTF)	218,423	8,99156	201,492	236,777
Standard Deviation	38,3569	6,41597	27,6352	53,2383
Median	216,705	9,89384	198,156	236,991
First Quartile(Q1)	190,481	9,63934	172,495	210,342
Third Quartile(Q3)	244,284	11,0118	223,627	266,849
Interquartile Range(IQR)	53,8028	8,97770	38,7945	74,6172

Pike (1966) cancer data for rats

3-parameter Weibull



Pike 3-parameter Weibull: Profile log likelihood for γ

γ	$\hat{\theta}(\gamma)$	$\hat{\alpha}(\gamma)$	$\tilde{l}(\gamma)$
0	234.3	6.08	−88.233
60	173.2	4.49	−87.831
100	131.8	3.38	−87.467
110	121.2	3.08	−87.381
120	110.6	2.78	−87.327
122	108.4	2.71	−87.324
125	105.2	2.61	−87.330
130	99.7	2.44	−87.382
135	94.0	2.24	−87.542
140	88.0	1.99	−88.064
142	85.2	1.80	−88.773
143	81.1	1.00	−91.718