Spring 2009

TMA4275 LIFETIME ANALYSIS

Bo Lindqvist

Department of Mathematical Sciences

NTNU

bo@math.ntnu.no

http://www.math.ntnu.no/~bo/

GOALS

After finishing this course you should ....

- **know the most common concepts and distributions from lifetime modeling**
- **be able to use graphical methods for description and comparison of lifetime data**
- **be able to use statistical methods for statistical inference (estimation, confidence interval, hypothesis testing) of lifetime data**
- **be able to analyze lifetime data by using computer software (MINITAB)**

About the course

The course gives an introduction to stochastic modelling and statistical methods for use in lifetime data analysis, with particular view to applications in reliability analysis and medicine.

The lectures are based on knowledge from TMA4240/TMA4245 Statistics or equivalent. It will be an advantage to have taken one of the courses TPK4120 Industrial safety and reliability, TMA4260 Industrial statistics, or TMA4255 Experimental design and applied statistical methods.

Contents: Basic concepts in lifetime modelling. Censored observations. Nonparametric estimation and graphical plotting for lifetime data (Kaplan-Meier, Nelson-plot). Estimation and testing in parametric lifetime distributions. Analysis of lifetimes with covariates. (Cox-regression, accelerated lifetime testing). Modelling and analysis of recurrent events. Nonhomogeneous Poisson-processes. Nelson-Aalen estimators. Bayesian lifetime analysis.

- Weekly hours: Spring: 4F+1Ø+7S = 7,5 SP
- Course type: Lectures and exercises with the use of a computer (MINITAB). Lectures may be given in English. Portfolio assessment is the basis for the grade awarded in the course. This portfolio comprises a written final examination 80% and selected parts of the exercises 20%. The results for the constituent parts are to be given in %-points, while the grade for the whole portfolio (course grade) is given by the letter grading system. Retake of examination may be given as an oral examination.

Lecturer

Professor Bo Lindqvist, room 1129, Sentralbygg II. Tlf. (735)93532
Office hours: To be announced.
Email: bo@math.ntnu.no

Exercise lab teacher	Research assistant Rupali Akerkar, room 1124, Sentralbygg II. Tlf. (735)92021 Office hours: To be announced. Email: Rupali.Akerkar@math.ntnu.no
Course book	The main source will be the book Rausand & Høyland: System Reliability Theory: Models, Statistical Methods, and Applications, 2nd Edition. Wiley 2004. Notes/copies about certain topics will be handed out. Foils from the lectures can be downloaded as pdf-files from this website.
Curriculum	PRELIMINARY CURRICULUM can be found here.
Lectures	Thursdays 12.15-14.00 in room F6. Fridays 10.15-12.00 in room F2. First lecture is January 15.
Exercises	Mondays 15.15-16.00 in room F2. First time: January 26. Link to exercise website will come here. Some exercises (including the obligatory ones) require use of the statistics computer package MINITAB, see http://www.ntnu.no/adm/it/brukerstotte/programvare/minitab. NTNU has an unlimited site licence for Windows and Macintosh for installation of MINITAB on NTNUs area and on private machines of students and staff. MINITAB is also available on several computer labs.
Final exam:	May 18, 2009. Written. 4 hours (9:00-13:00). **Permitted aids:** B - All printed and handwritten aids permitted, approved simple calculator.

PRELIMINARY CURRICULUM AND LECTURE PLAN

Week No.	Literature R & H	Topic	Comment R & H
3	2.1-2.14, 2.17, 2.20 Notes on "Log-location-scale…" .	Probability distributions for lifetimes. Fundamental properties. Important distributions and properties.	Only main results in 2.17 are covered.
4-5	11.1-11.3, 11.5 The note "About the exponential distribution …"	Lifetime data. Censoring. Nonparametric methods. Plotting (TTT, Kaplan-Meier, Nelson-Aalen.)	
6-8	11.4 Notes on "Likelihood construction" and "On parametric inference…."	Parametric estimation and testing. Maximum likelihood. Information matrix. Confidence intervals. Probability plots (MINITAB).	
9-10		No lectures	
11-13	Ch. 12. Notes on "Survival regression" (Lecture week 8) and "Medical study"	Regression methods. Covariates. Weibull regression. Cox-regression. Accellerated lifetime testing.	Example 12.2 page 532 and rest of Ch. 12 are not covered.

14 and 16-17	Ch. 7	Point processes. Recurrent events. Repairable systems. Poisson processes and renewal. Modelling and statistical analysis of data. Likelihood-methods.	The following is covered (not always in detail): 7.1 only to 7.1.3; most of 7.2, but only to 7.2.6; 7.3 to 7.3.4; but in addition 7.3.8; most of 7.4. Pooled versions of Laplace and Mil-Hdbk tests (as used by MINITAB).
15		Easter vacation	
18	Ch. 13	Bayesian lifetime analysis	Selected parts of 13.1-13.5 are covered.
19		Review	

RELIABILITY

Common technical definition of reliability:

The probability that a system or a component will perform its intended task, under given operational conditions, for a specified time period.

LIFE TIMES (SURVIVAL DATA)

- **Time to failure of a component or a system**
- **Number of cycles to failure (fatigue testing)**
- **Time to next epileptic seizure for patient**
- **Times of failure and repair of a machine**

WHY COLLECT AND ANALYZE LIFETIME/SURVIVAL/RELIABILITY DATA?

- **Assess reliability of a system/component/product**
- **Compare two or more products with respect to reliability**
- **Predict product reliability in the design phase**
- **Predict warranty claims for a product in the market**

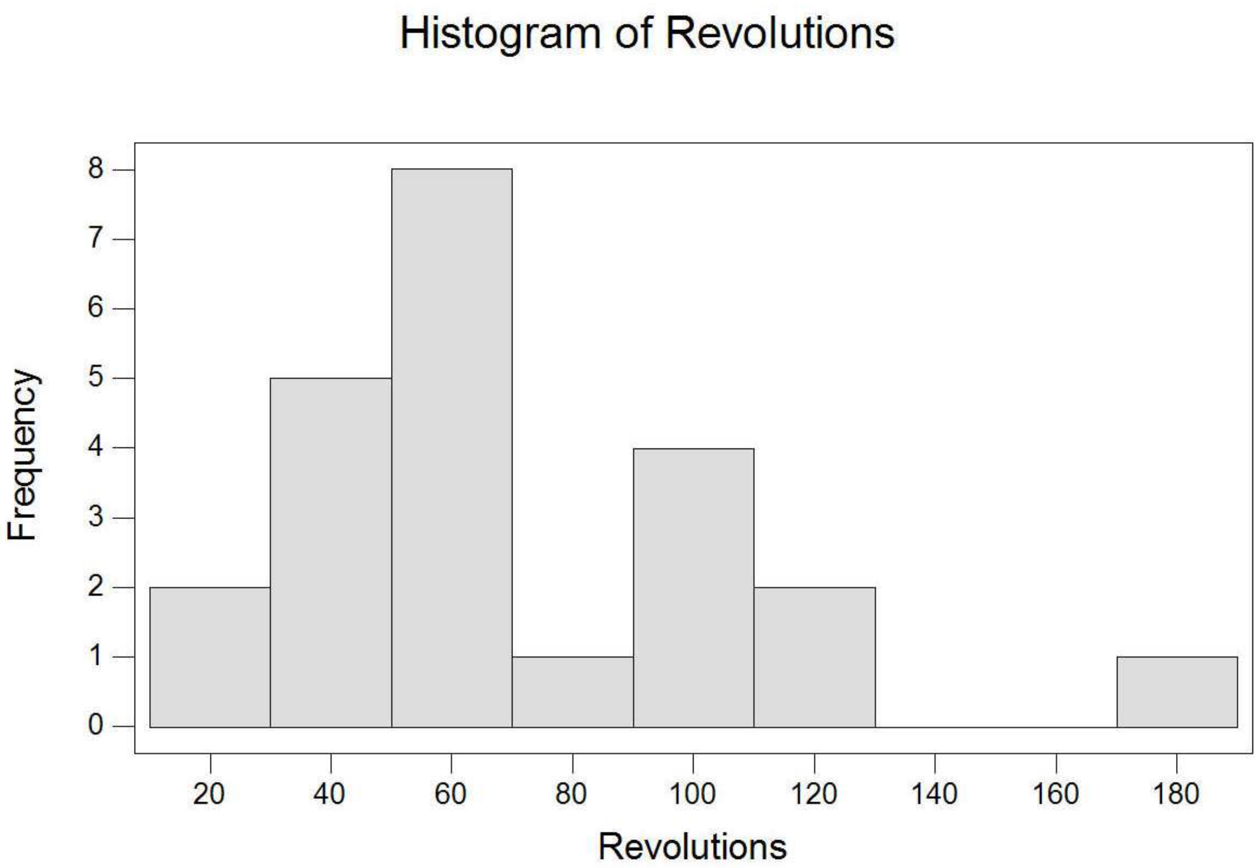
SPECIAL ASPECTS OF LIFETIME ANALYSIS

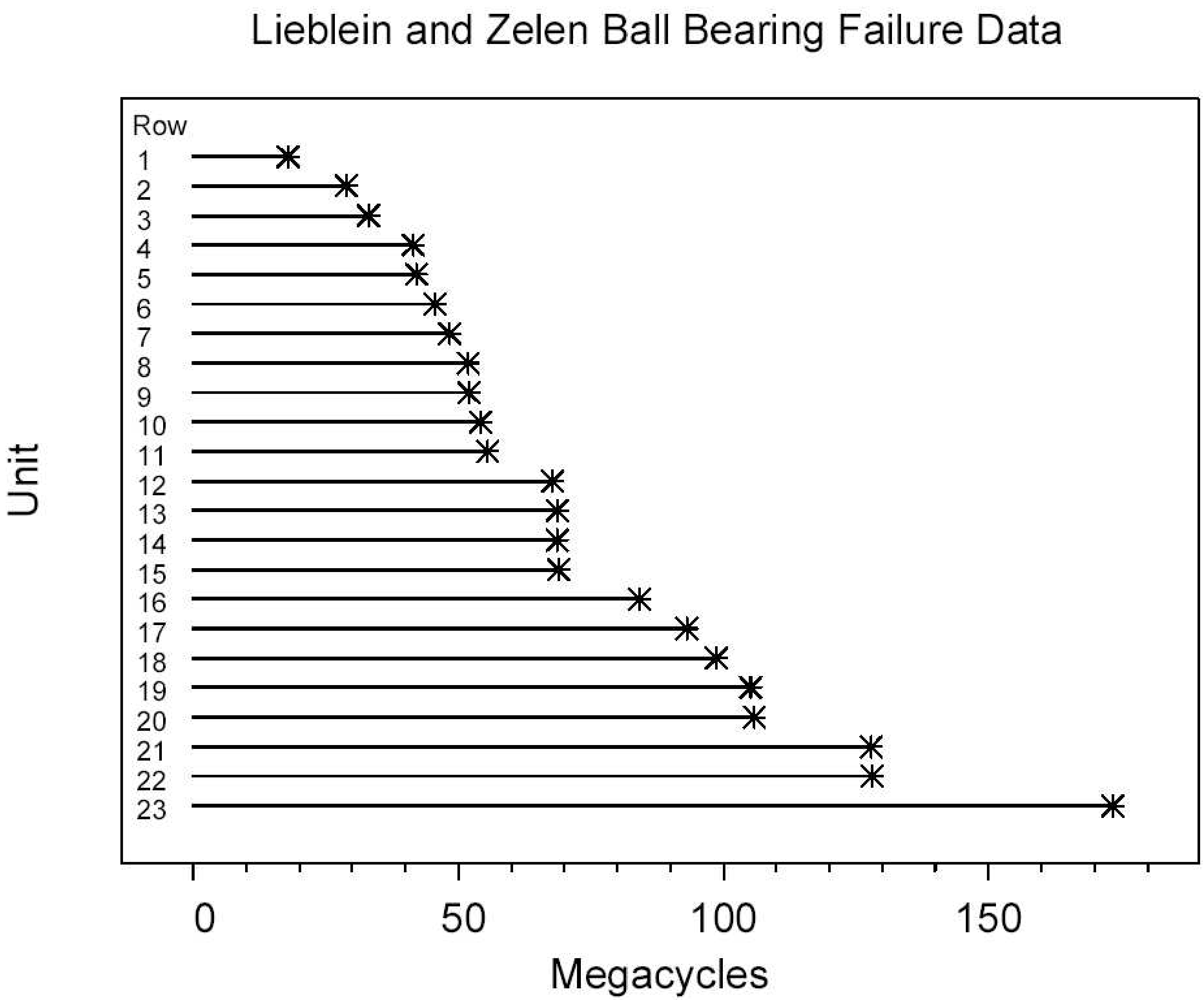
- Definition of starting time and failure time are difficult
- Definition of time scale (operation time, calendar time, number of cycles)
- Censored data (what do we do with units that have not failed within the observation period?)
- Effect of environmental conditions
- What if a unit fails of another cause than the one we would like to study? ("competing risks")
- Recurrent events – what if the system can fail several times; how to analyze recurring stages of a disease?

BALL BEARINGS FAILURE DATA

Data: Millions of revolutions to fatigue failure for 23 units

17,88	28,92	33,00	41,52	42,12	45,60	48,40	51,84
51,96	54,12	55,56	67,80	68,64	68,64	68,88	84,12
93,12	98,64	105,12	105,84	127,92	128,04	173,40	




Lieblein and Zelen Ball Bearing Failure Data
Row
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Unit
0
50
100
150
Megacycles

re Tools

IC Data (Meeker, 1987)

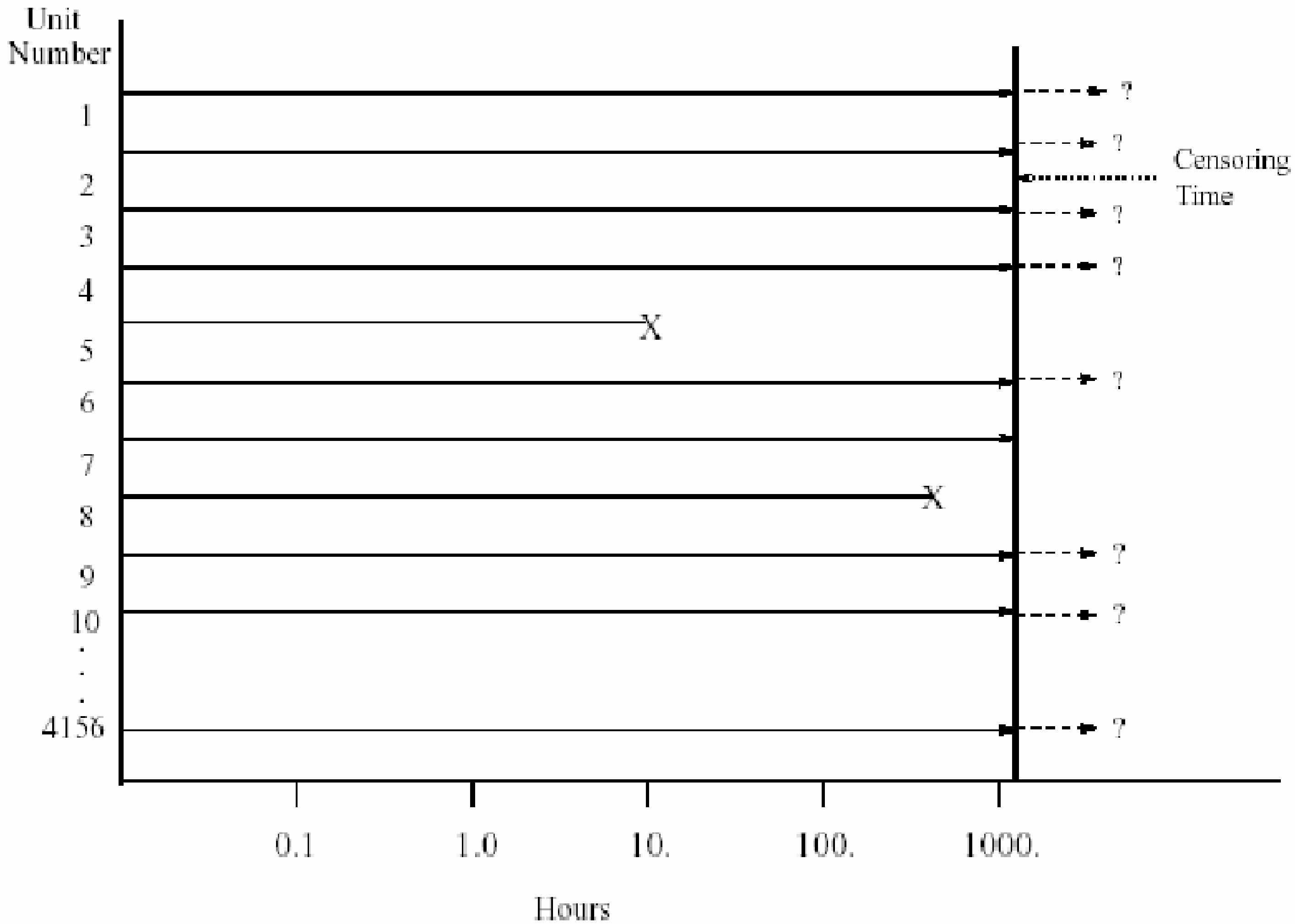
- Integrated circuit failure times in hours
  - n = 4156 ICs tested for 1,370 hours at 80° C and 80% relative humidity
  - There were 28 failures
  - When the test ended at 1,370 hours, 4128 units were still running

.10	.10	.15	.60	.80	.80
1.20	2.5	3.0	4.0	4.0	6.0
10.0	10.0	12.5	20.	20.	43.
43.	48.	48.	54.	74.	84.
94.	168.	263.	593.		

TYPICAL PROBLEMS:

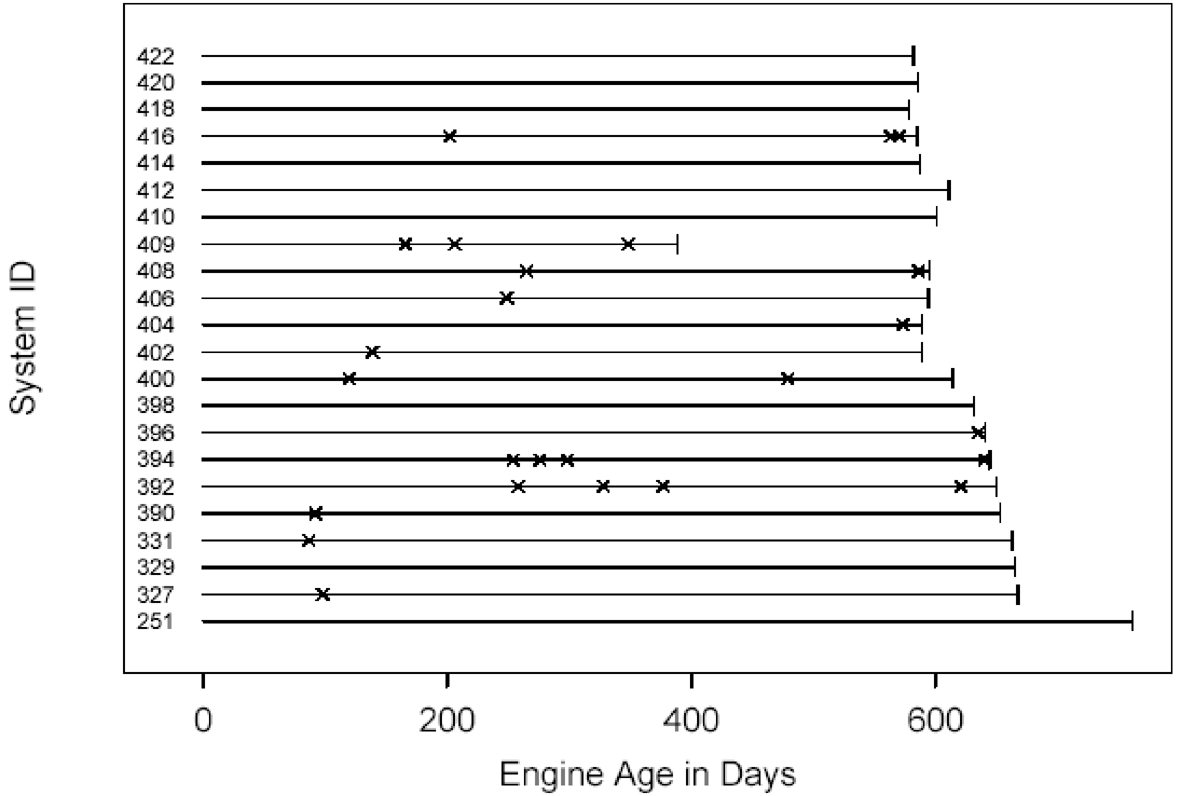
- **How to estimate the distribution of the failure time when there are censored observations?**
- **Probability of failure before 100 hours?**
- **Failure rate by 100 hours?**
- **Proportion failed after 10^5 hours?**

IC Data Failure Pattern



RECURRENT EVENTS/REPAIRABLE SYSTEMS

Valve Seat Replacement Times Event Plot
(Nelson and Doganaksoy 1989)

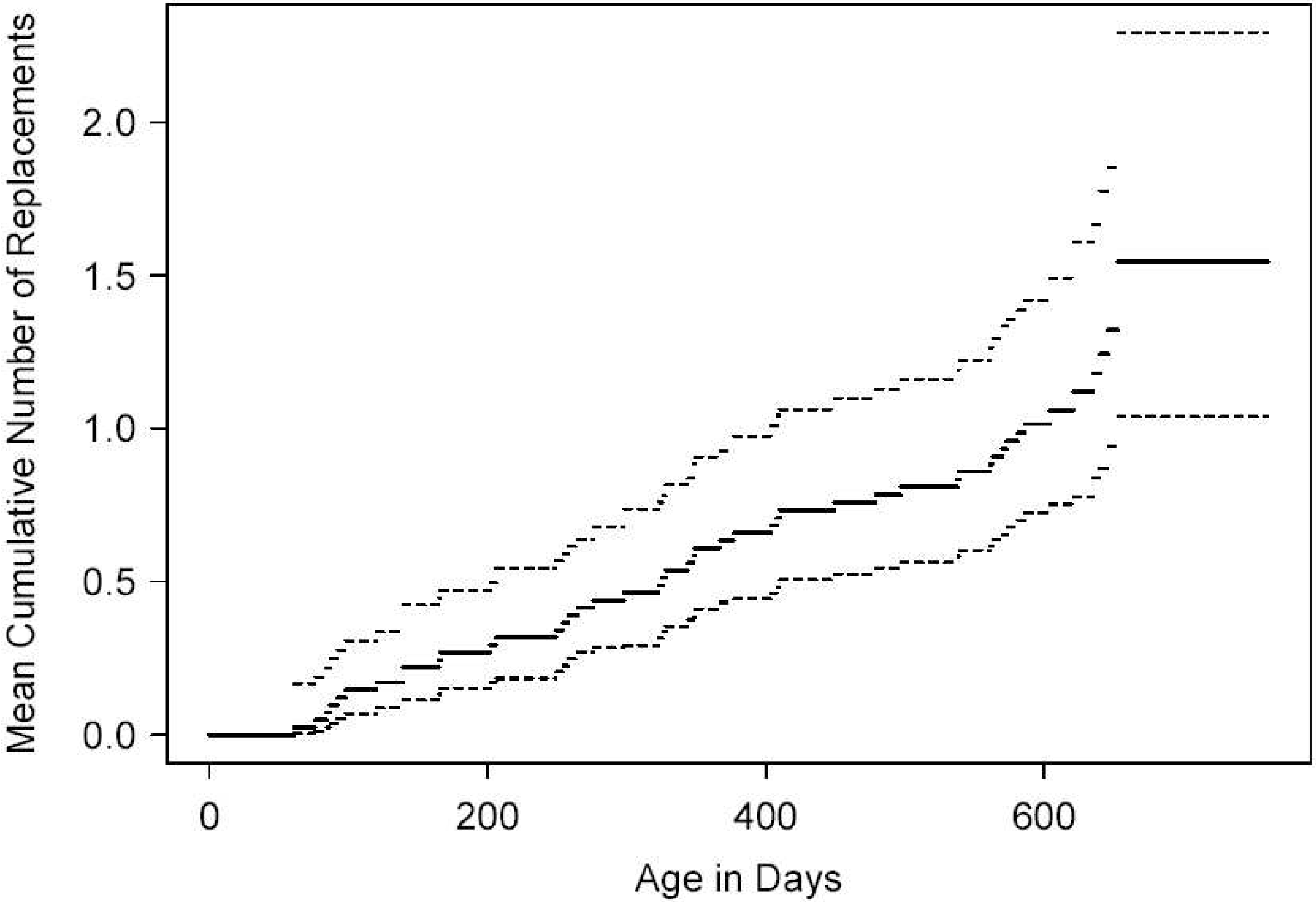


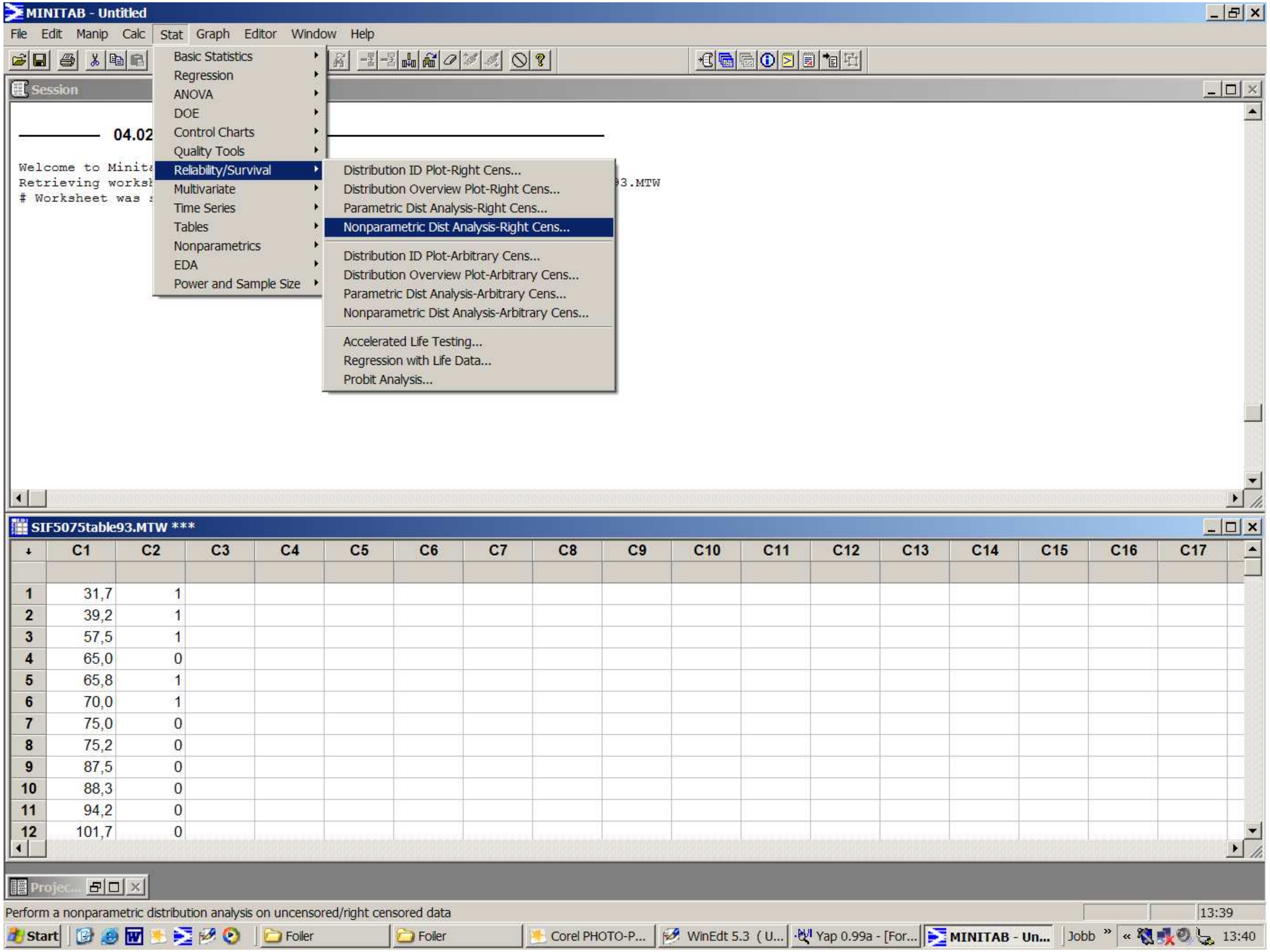
Valve Seat Replacement Times
(Nelson and Doganaksoy 1989)

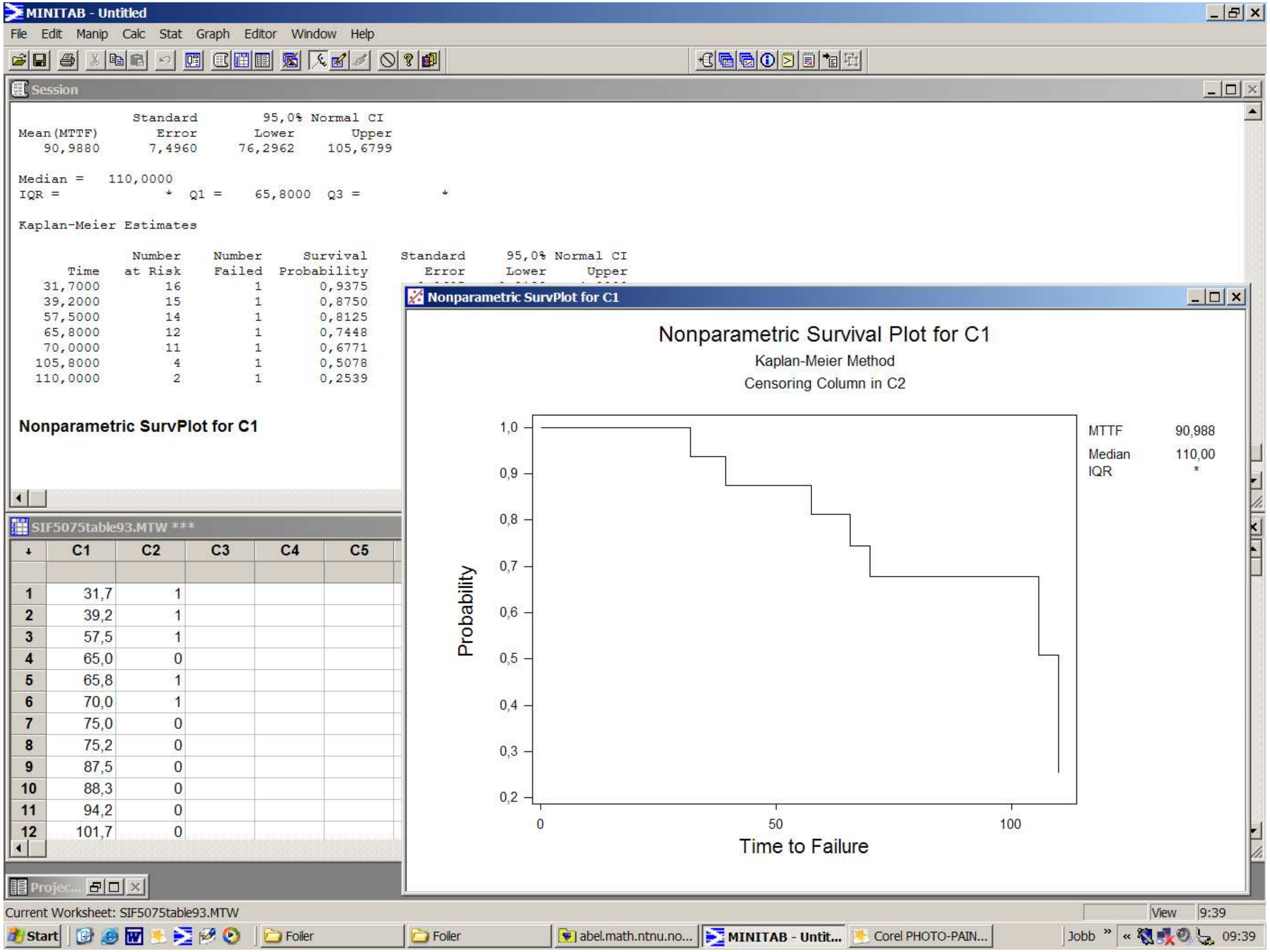
Data collected from valve seats from a fleet of 41 diesel engines (days of operation)

- Each engine has 16 valves
- Does the replacement rate increase with age?
- How many replacement valves will be needed in the future?
- Can valve life in these systems be modeled as a renewal process?

Estimate of Number of Valve Seat $\mu(t)$



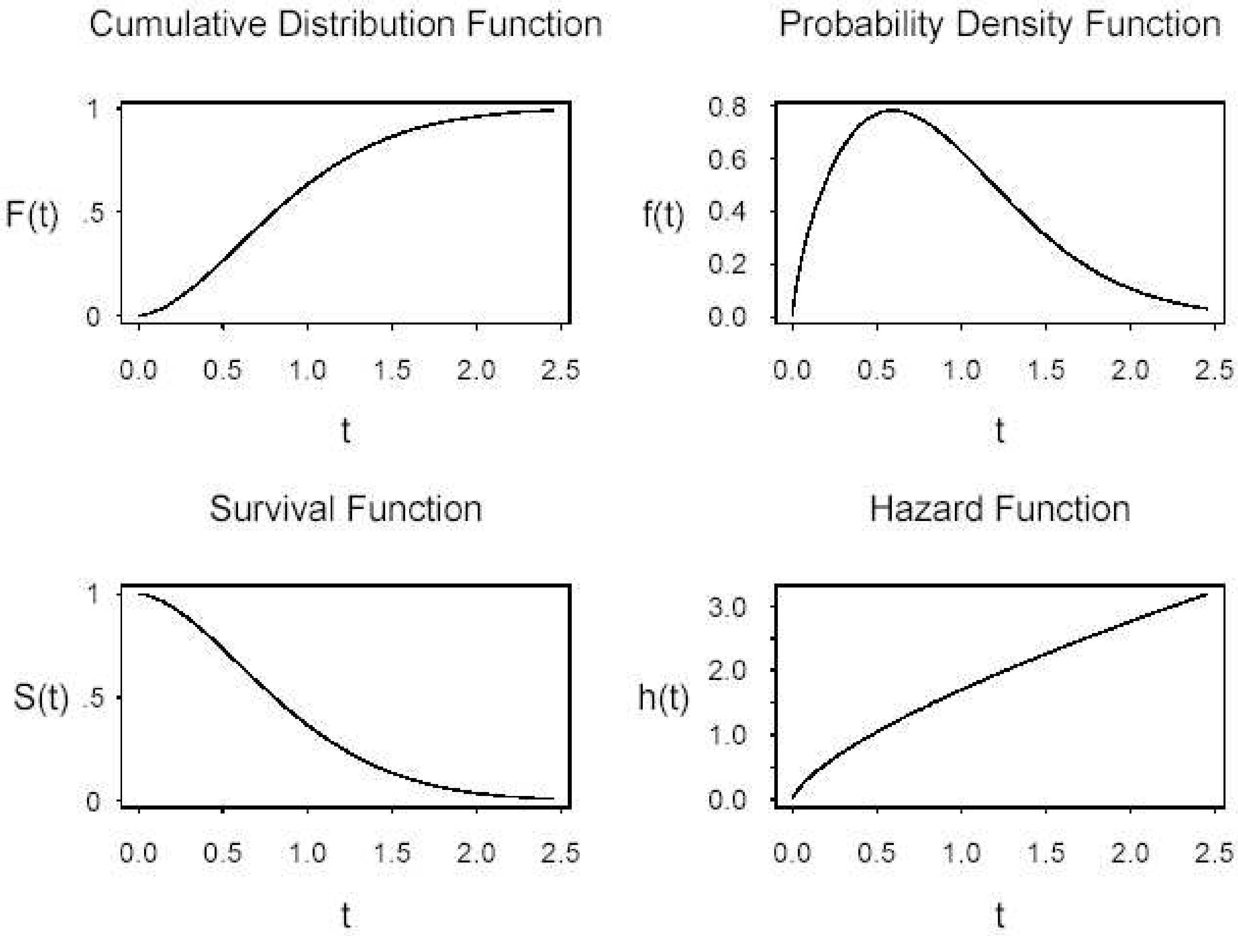

MINITAB - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Session
04.02
Welcome to Minit
Retrieving worksh
Worksheet was
3.MTW
Basic Statistics
Regression
ANOVA
DOE
Control Charts
Quality Tools
Reliability/Survival
Multivariate
Time Series
Tables
Nonparametrics
EDA
Power and Sample Size
Distribution ID Plot-Right Cens...
Distribution Overview Plot-Right Cens...
Parametric Dist Analysis-Right Cens...
Nonparametric Dist Analysis-Right Cens...
Distribution ID Plot-Arbitrary Cens...
Distribution Overview Plot-Arbitrary Cens...
Parametric Dist Analysis-Arbitrary Cens...
Nonparametric Dist Analysis-Arbitrary Cens...
Accelerated Life Testing...
Regression with Life Data...
Probit Analysis...
SIF5075table93.MTW ***
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15 C16 C17
1 31,7 1
2 39,2 1
3 57,5 1
4 65,0 0
5 65,8 1
6 70,0 1
7 75,0 0
8 75,2 0
9 87,5 0
10 88,3 0
11 94,2 0
12 101,7 0
Projec...
Perform a nonparametric distribution analysis on uncensored/right censored data
13:39
Start
Foler
Foler
Corel PHOTO-P...
WinEdt 5.3 (U...
Yap 0.99a - [For...
MINITAB - Un...
Jobb
13:40


MINITAB - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Session
Standard 95,0% Normal CI
Mean (MTTF) Error Lower Upper
90,9880 7,4960 76,2962 105,6799
Median = 110,0000
IQR = * Q1 = 65,8000 Q3 = *
Kaplan-Meier Estimates
Time Number at Risk Number Failed Survival Probability Standard Error 95,0% Normal CI Lower Upper
31,7000 16 1 0,9375
39,2000 15 1 0,8750
57,5000 14 1 0,8125
65,8000 12 1 0,7448
70,0000 11 1 0,6771
105,8000 4 1 0,5078
110,0000 2 1 0,2539
Nonparametric SurvPlot for C1
SIF5075table93.MTW ***
C1 C2 C3 C4 C5
1 31,7 1
2 39,2 1
3 57,5 1
4 65,0 0
5 65,8 1
6 70,0 1
7 75,0 0
8 75,2 0
9 87,5 0
10 88,3 0
11 94,2 0
12 101,7 0
Nonparametric SurvPlot for C1
Nonparametric Survival Plot for C1
Kaplan-Meier Method
Censoring Column in C2
Probability
Time to Failure
MTTF 90,988
Median 110,00
IQR *
Current Worksheet: SIF5075table93.MTW

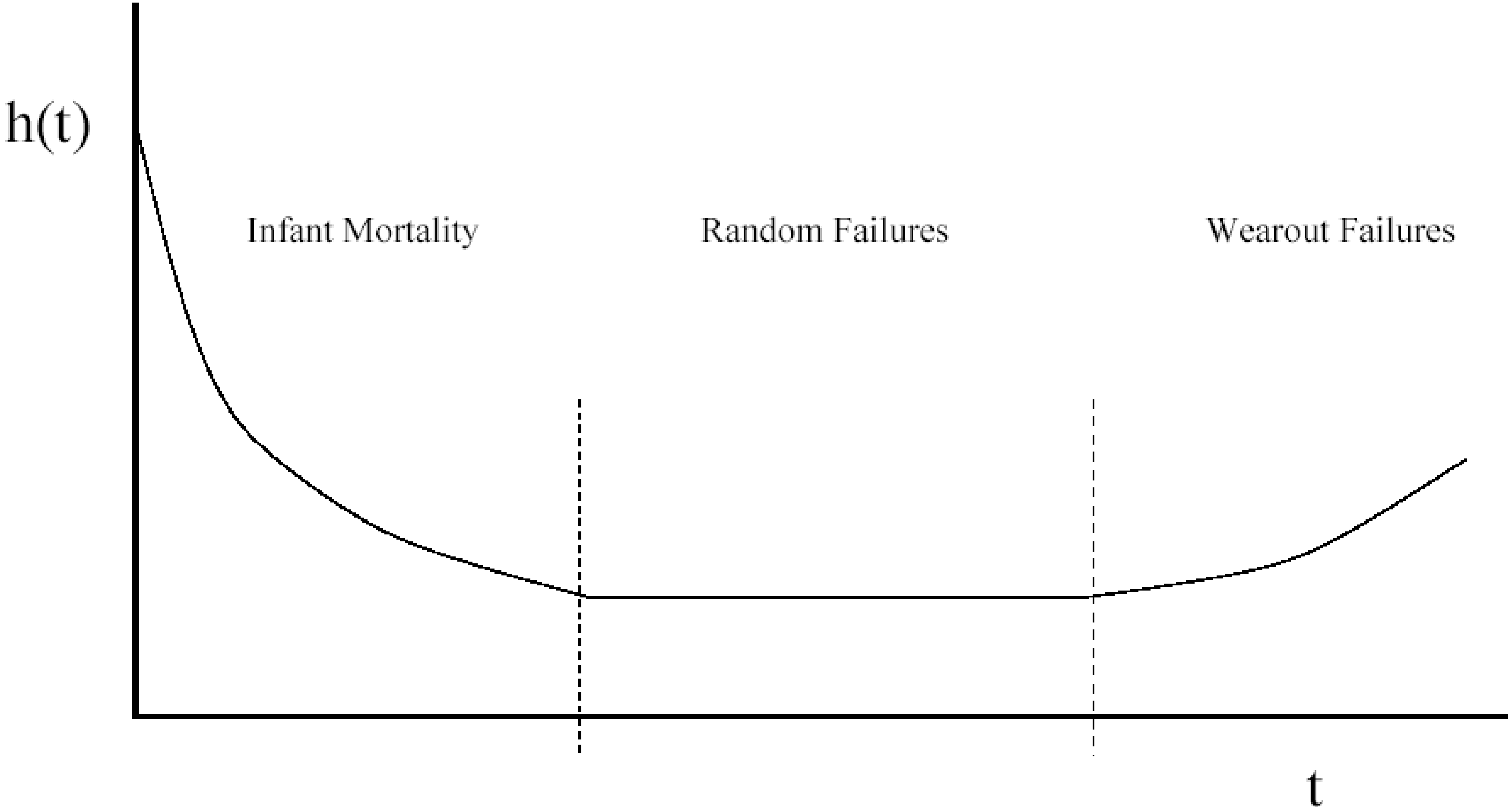
Typical Failure-time cdf, pdf, hf, and sf

$$F(t) = 1 - \exp(-t^{1.7}); \quad f(t) = 1.7 \times t^{.7} \times \exp(-t^{1.7})$$

$$S(t) = \exp(-t^{1.7}); \quad h(t) = 1.7 \times t^{.7}$$



Bathtub Curve Hazard Function

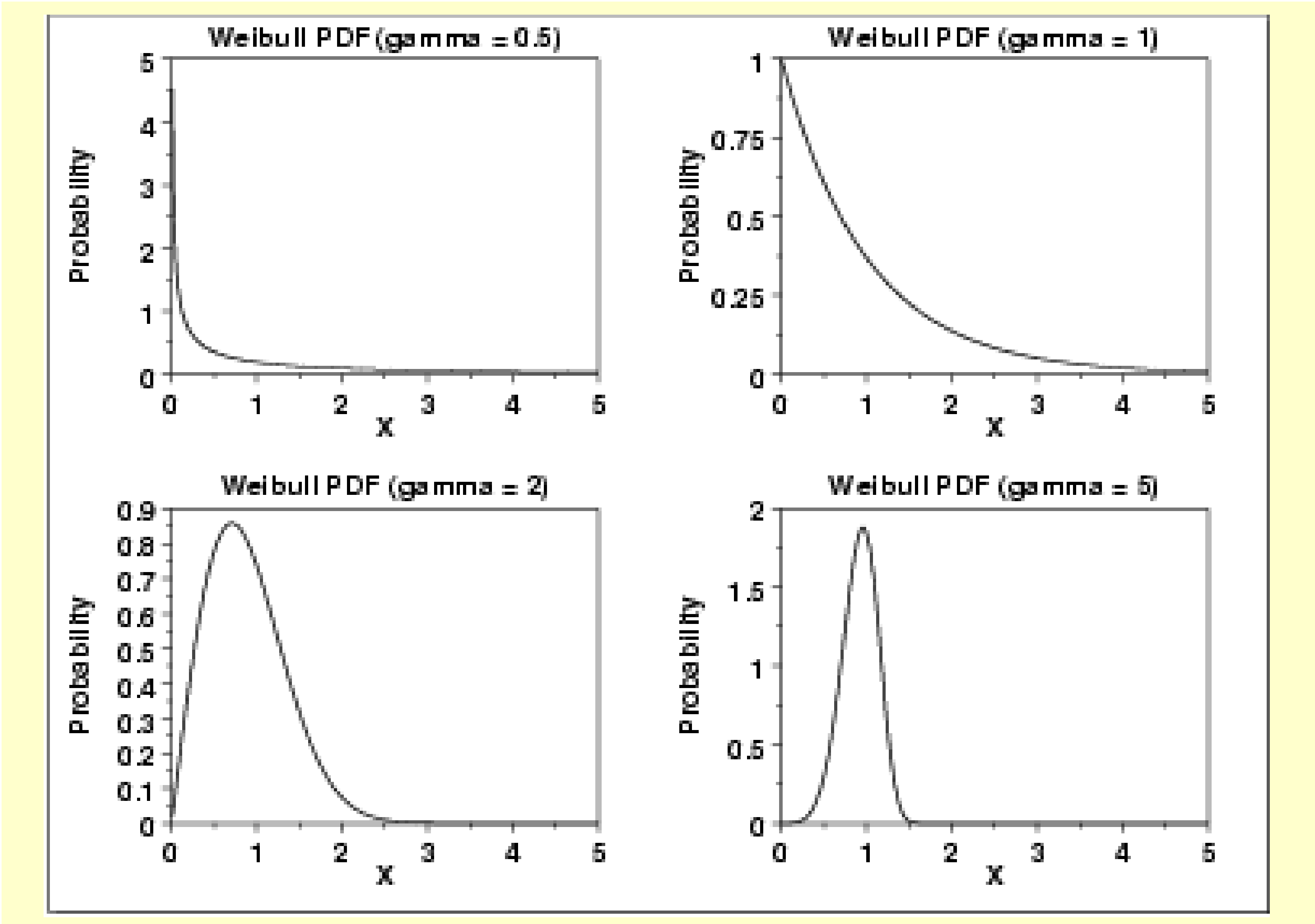


10 Dødelighetstabeller [1]. 2001

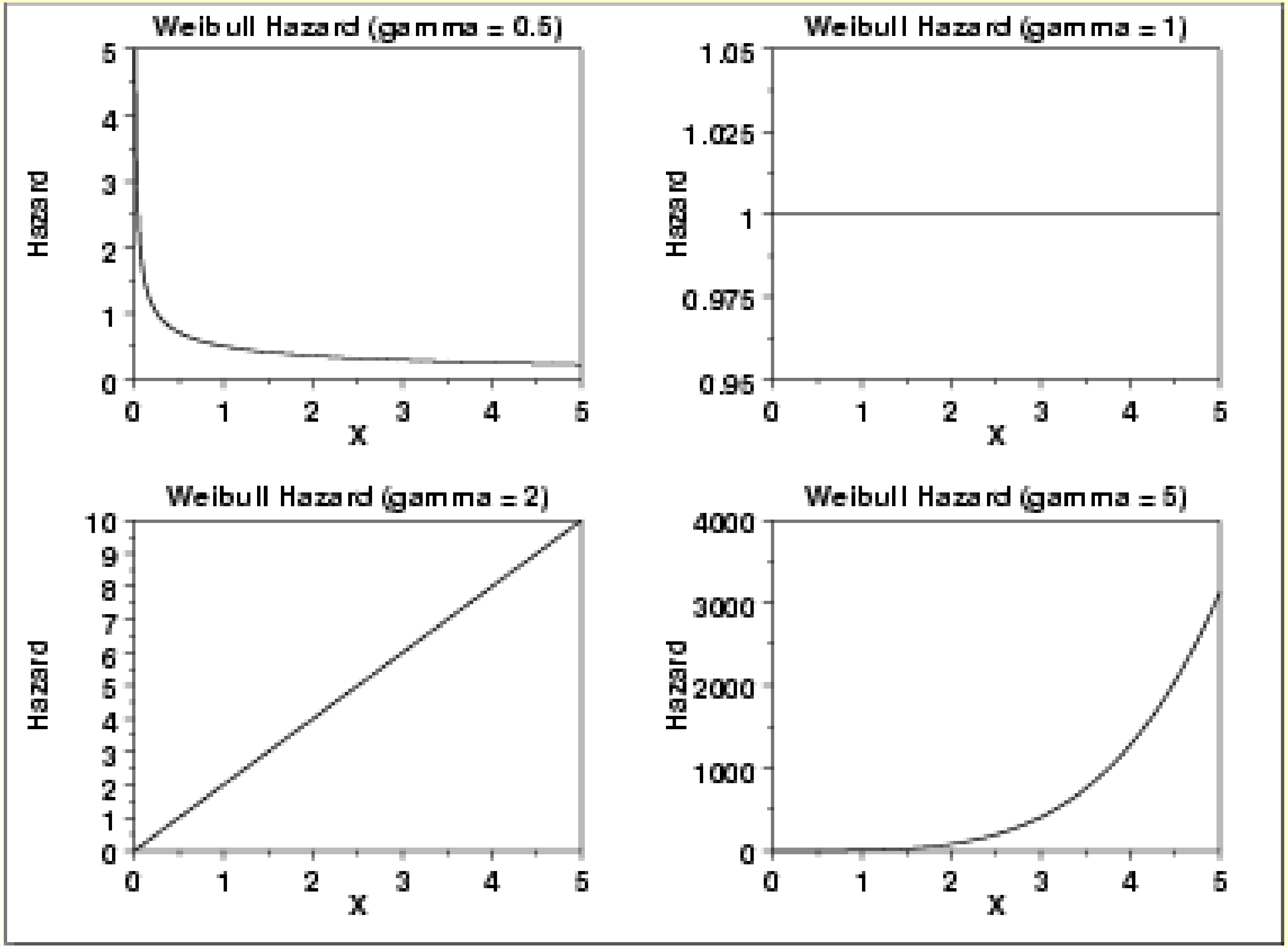
Alder x	Levende ved alder x		Døde i alder x til x+1		Forventet gjenstående levetid ved alder x		Dødssannsynlighet for alder x, Promille, (Uglattet).	
	lx		dx		e0x		qx	
	Menn	Kvinner	Menn	Kvinner	Menn	Kvinner	Menn	Kvinner
0	100 000	100 000	424	343	76,21	81,53	4,24	3,43
1	99 576	99 657	39	27	75,53	80,81	0,39	0,28
2	99 537	99 630	33	7	74,56	79,83	0,33	0,07
3	99 504	99 623	32	17	73,58	78,84	0,33	0,17
4	99 472	99 606	13	10	72,61	77,85	0,13	0,10
5	99 459	99 596	6	10	71,62	76,86	0,06	0,10
6	99 453	99 586	22	17	70,62	75,87	0,22	0,17
7	99 431	99 569	10	3	69,64	74,88	0,10	0,03
8	99 422	99 566	9	13	68,64	73,88	0,09	0,13
9	99 412	99 552	9	3	67,65	72,89	0,09	0,03
10	99 403	99 549	12	3	66,66	71,89	0,12	0,03
11	99 390	99 546	3	10	65,66	70,89	0,03	0,10
12	99 387	99 536	16	3	64,67	69,90	0,16	0,03
13	99 371	99 532	10	11	63,68	68,90	0,10	0,11
14	99 361	99 522	7	7	62,68	67,91	0,07	0,07
15	99 354	99 514	32	11	61,69	66,92	0,32	0,11
16	99 322	99 503	33	23	60,71	65,92	0,33	0,23
17	99 289	99 480	77	39	59,73	64,94	0,77	0,39
18	99 212	99 441	90	35	58,77	63,96	0,91	0,35
19	99 122	99 407	123	34	57,83	62,99	1,24	0,34
20	98 999	99 373	155	60	56,90	62,01	1,57	0,60
21	98 844	99 313	142	15	55,99	61,04	1,44	0,15

40	96 600	98 433	147	85	38,05	42,49	1,53	0,86
41	96 453	98 348	144	110	37,11	41,53	1,49	1,12
42	96 309	98 239	208	100	36,16	40,58	2,16	1,02
43	96 101	98 138	181	110	35,24	39,62	1,89	1,12
44	95 919	98 029	205	112	34,31	38,66	2,14	1,15
45	95 715	97 916	190	153	33,38	37,70	1,98	1,57
46	95 525	97 763	256	172	32,44	36,76	2,68	1,76
47	95 268	97 591	256	160	31,53	35,83	2,68	1,64
48	95 013	97 431	324	191	30,61	34,88	3,41	1,96
49	94 689	97 240	310	197	29,72	33,95	3,28	2,03
50	94 379	97 042	324	233	28,81	33,02	3,43	2,40
51	94 055	96 810	387	265	27,91	32,10	4,11	2,74
52	93 668	96 545	332	255	27,02	31,18	3,54	2,64
53	93 336	96 290	461	293	26,12	30,27	4,94	3,04
54	92 875	95 997	504	343	25,25	29,36	5,42	3,58
55	92 371	95 653	546	342	24,38	28,46	5,91	3,57
56	91 825	95 311	583	362	23,52	27,56	6,35	3,80
57	91 242	94 949	647	400	22,67	26,66	7,09	4,22
58	90 595	94 549	593	435	21,83	25,77	6,55	4,60
59	90 002	94 115	713	554	20,97	24,89	7,92	5,89
60	89 289	93 560	797	543	20,13	24,04	8,93	5,81
61	88 492	93 017	853	543	19,31	23,17	9,64	5,83
62	87 639	92 475	911	626	18,49	22,31	10,39	6,77
63	86 728	91 848	1 200	781	17,68	21,45	13,84	8,50
64	85 528	91 068	1 359	795	16,92	20,63	15,89	8,73
65	84 168	90 273	1 356	763	16,19	19,81	16,11	8,45
66	82 812	89 509	1 349	883	15,44	18,98	16,29	9,86
67	81 463	88 627	1 572	897	14,69	18,16	19,30	10,12
68	79 891	87 730	1 746	1 070	13,97	17,34	21,86	12,19
69	78 145	86 660	1 869	1 056	13,27	16,55	23,91	12,19

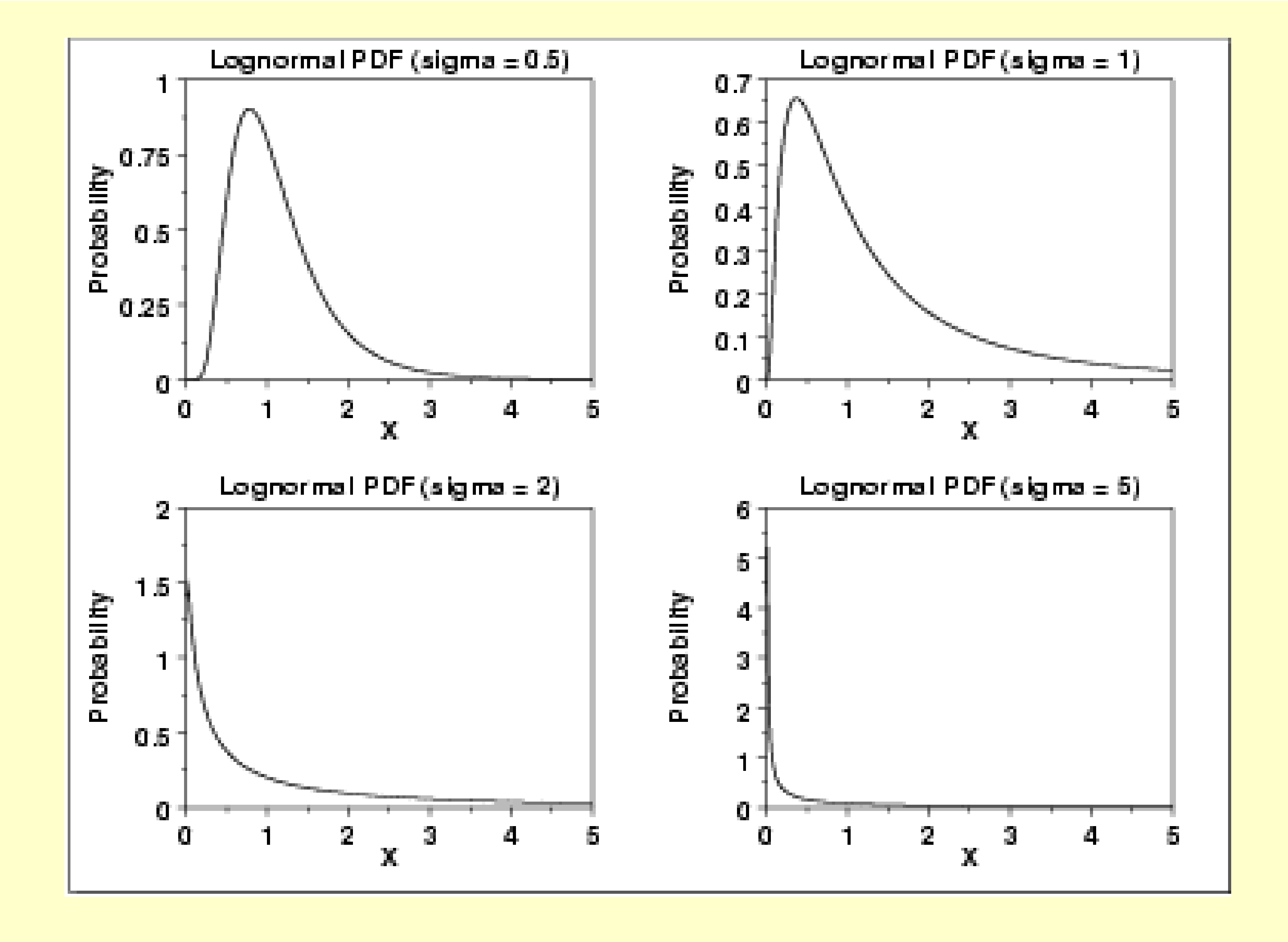
WEIBULL-DENSITIES ($\alpha = \gamma$, $\lambda = 1$ in terminology of book)



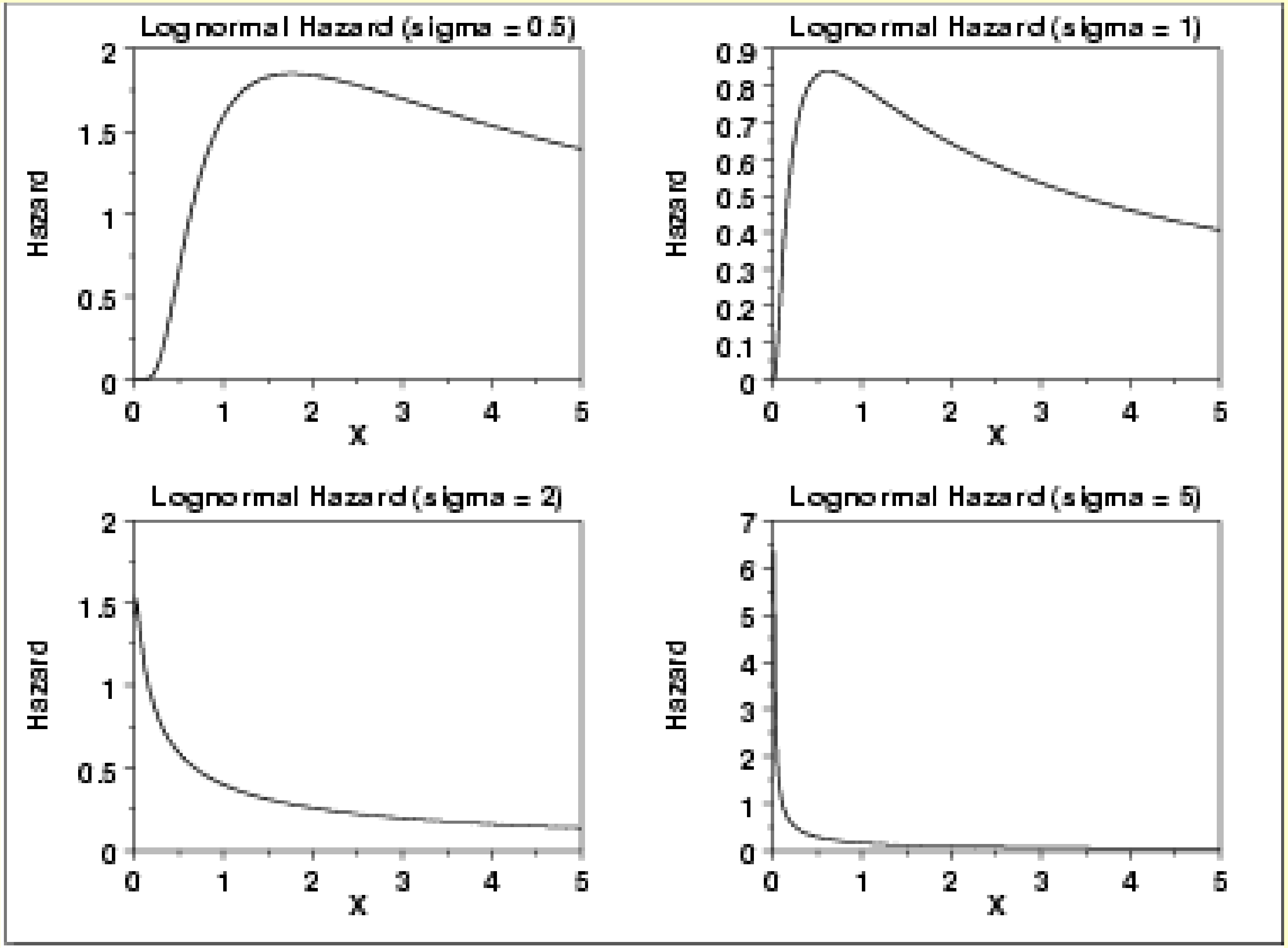
WEIBULL-HAZARDS ($\alpha = \gamma$, $\lambda = 1$ in terminology of book)



LOGNORMAL DENSITIES ($\tau = \sigma$, $\nu = 0$ in terminology of book)



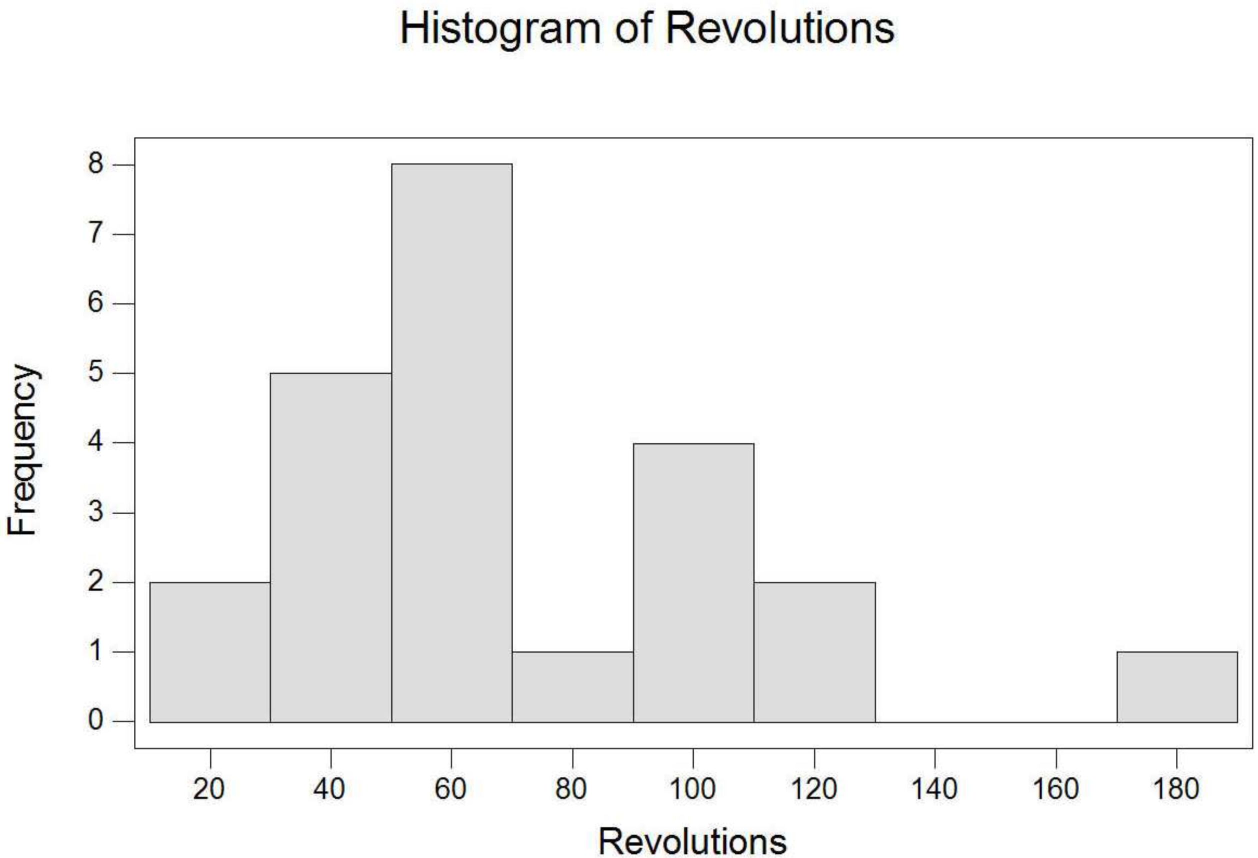
LOGNORMAL HAZARDS ($\tau = \sigma$, $\nu = 0$ in terminology of book)



RECALL BALL BEARINGS FAILURE DATA

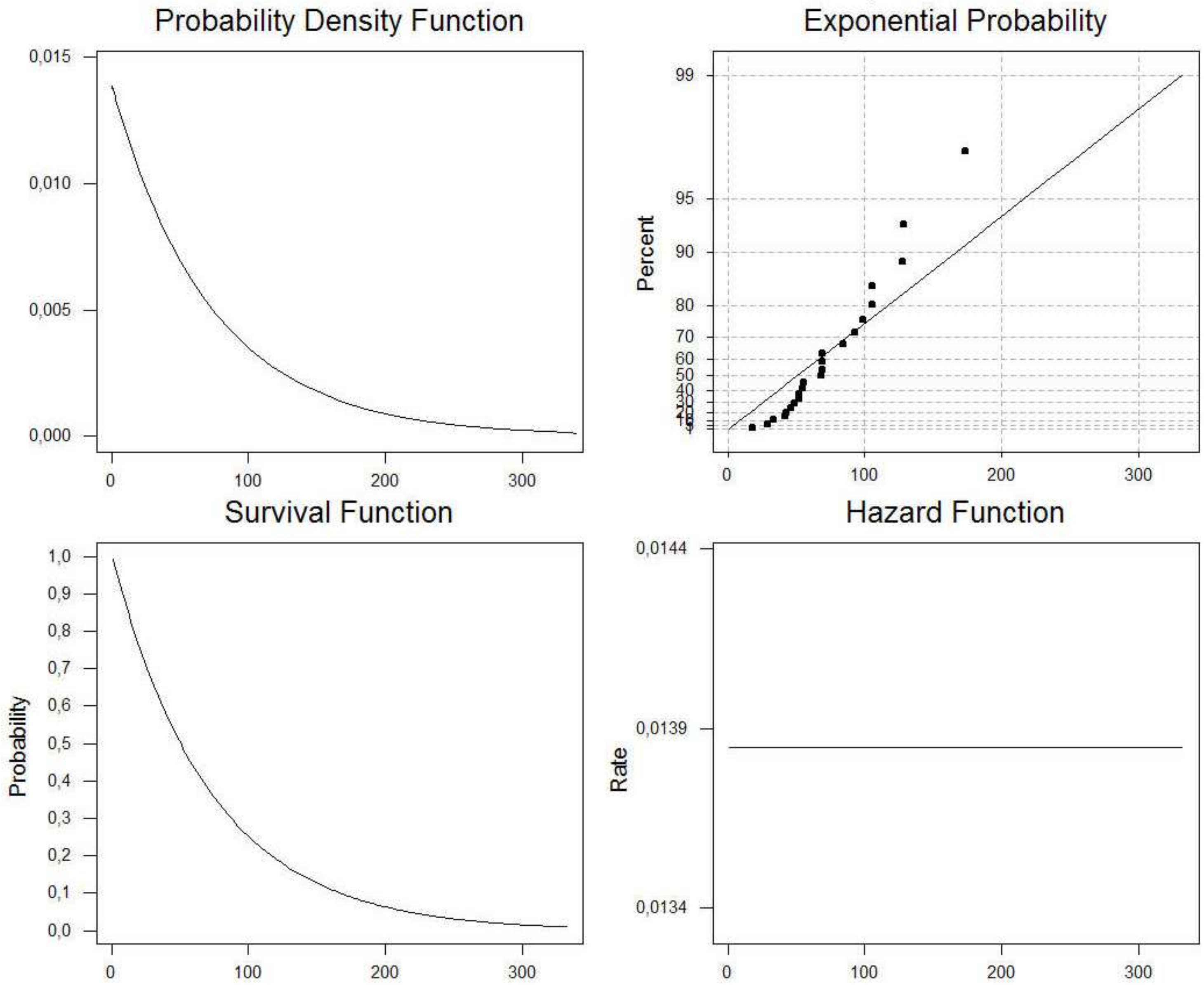
Data: Millions of revolutions to fatigue failure for 23 units

17,88	28,92	33,00	41,52	42,12	45,60	48,40	51,84
51,96	54,12	55,56	67,80	68,64	68,64	68,88	84,12
93,12	98,64	105,12	105,84	127,92	128,04	173,40	



Ball Bearings Failure Data

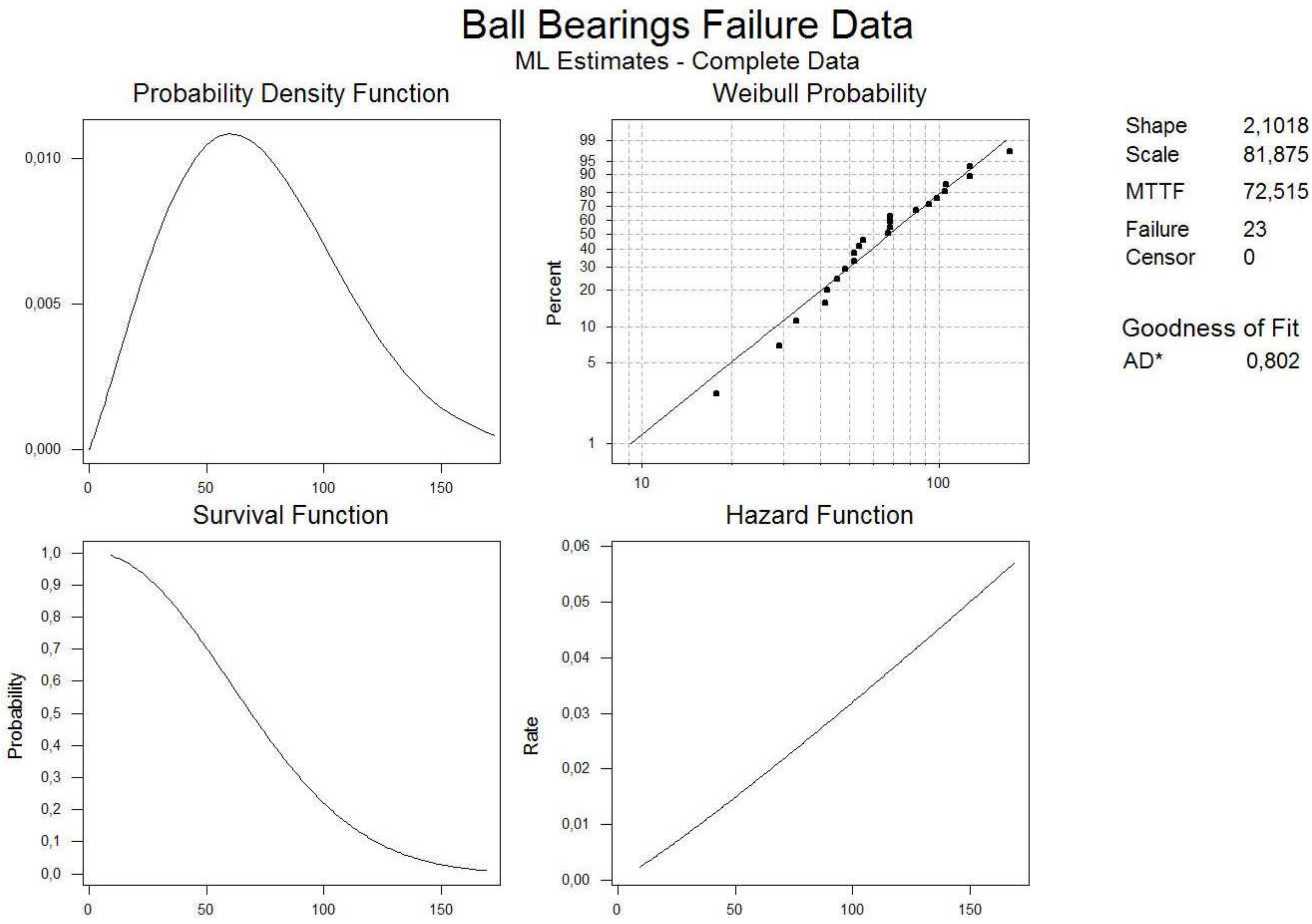
ML Estimates - Complete Data

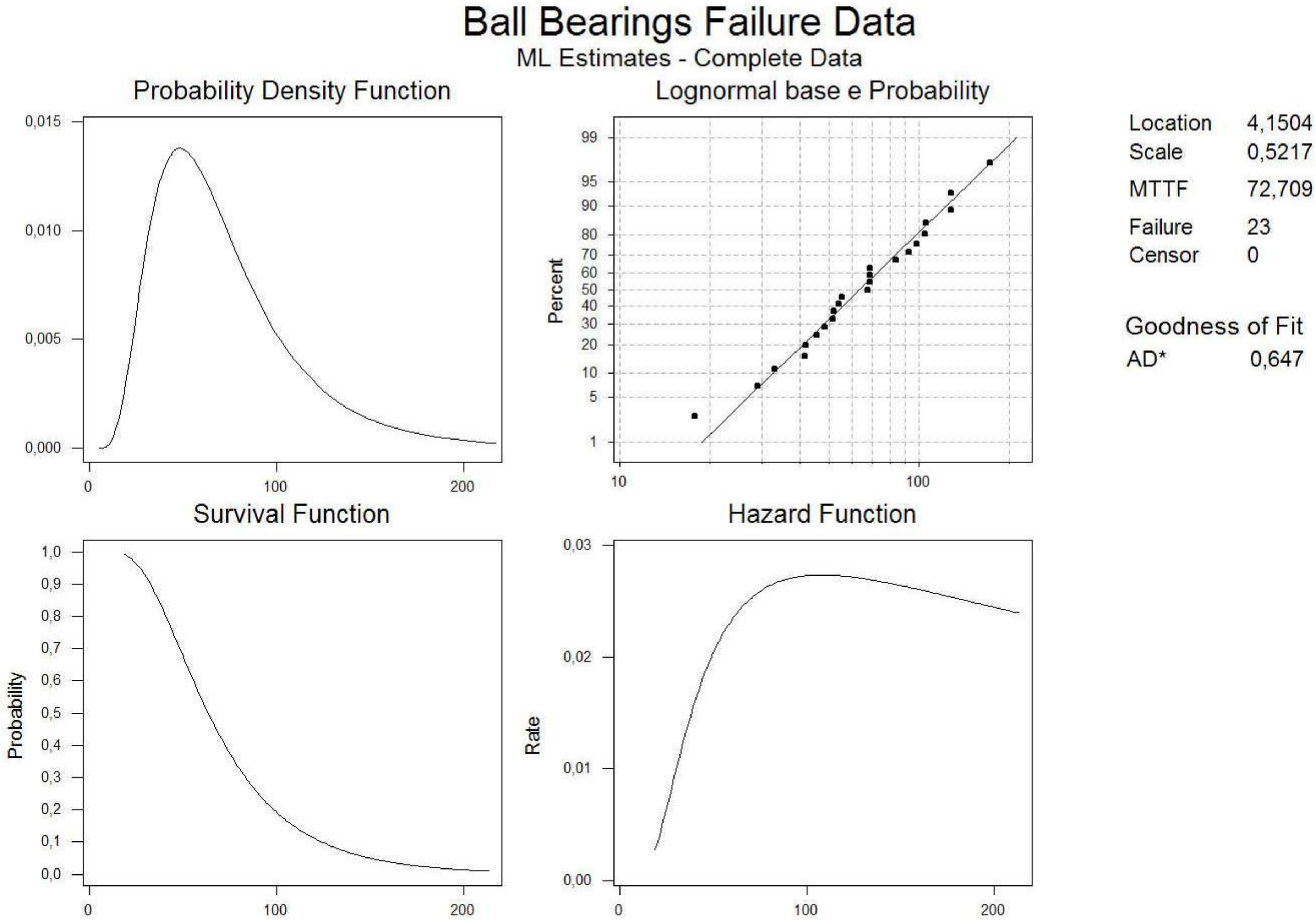


Shape	1
Scale	72,221
MTTF	72,221
Failure	23
Censor	0

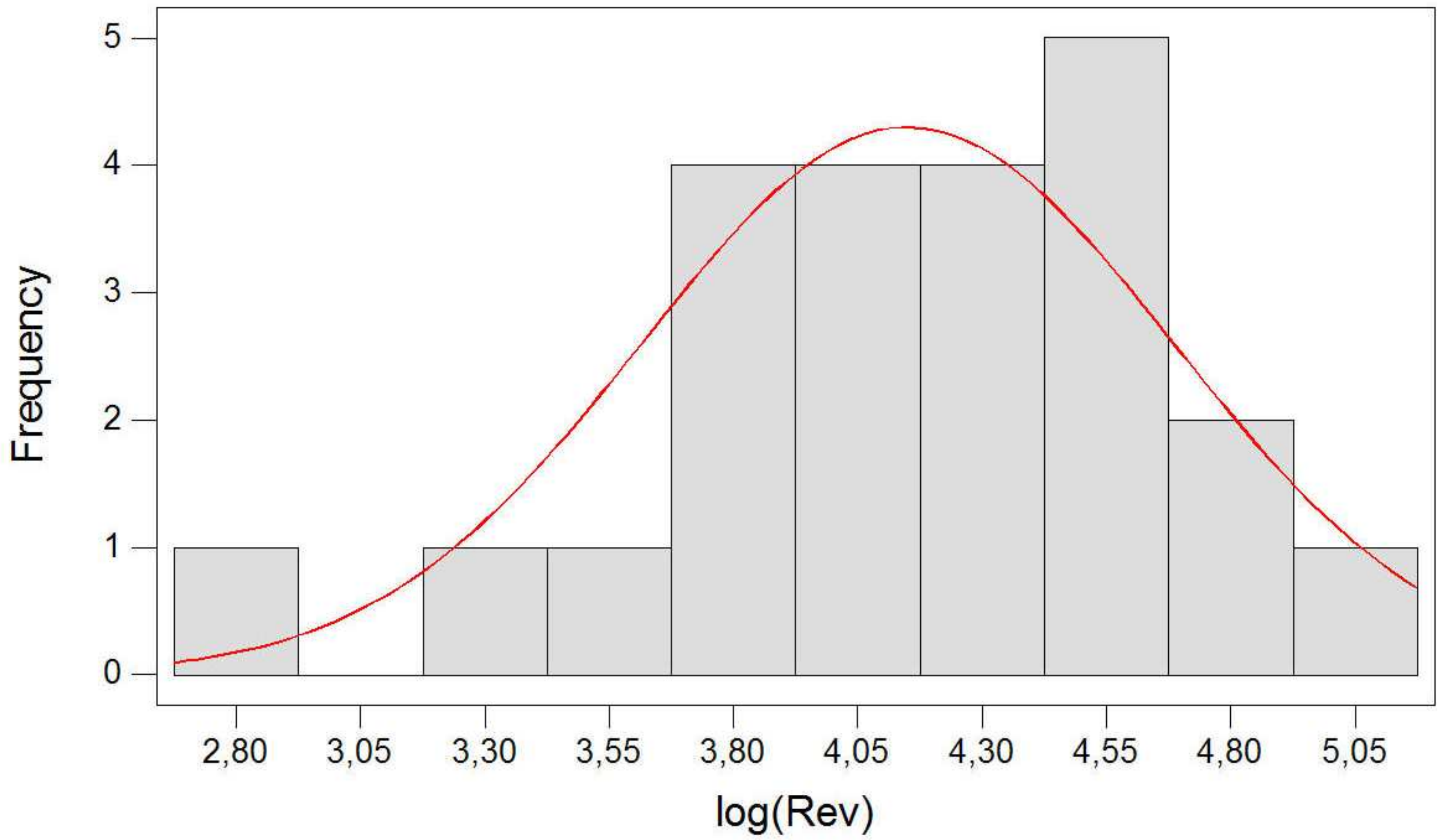
Goodness of Fit

AD*	3,341


Ball Bearings Failure Data
ML Estimates - Complete Data
Probability Density Function
Weibull Probability
Percent
Survival Function
Probability
Hazard Function
Rate
Shape 2,1018
Scale 81,875
MTTF 72,515
Failure 23
Censor 0
Goodness of Fit
AD* 0,802

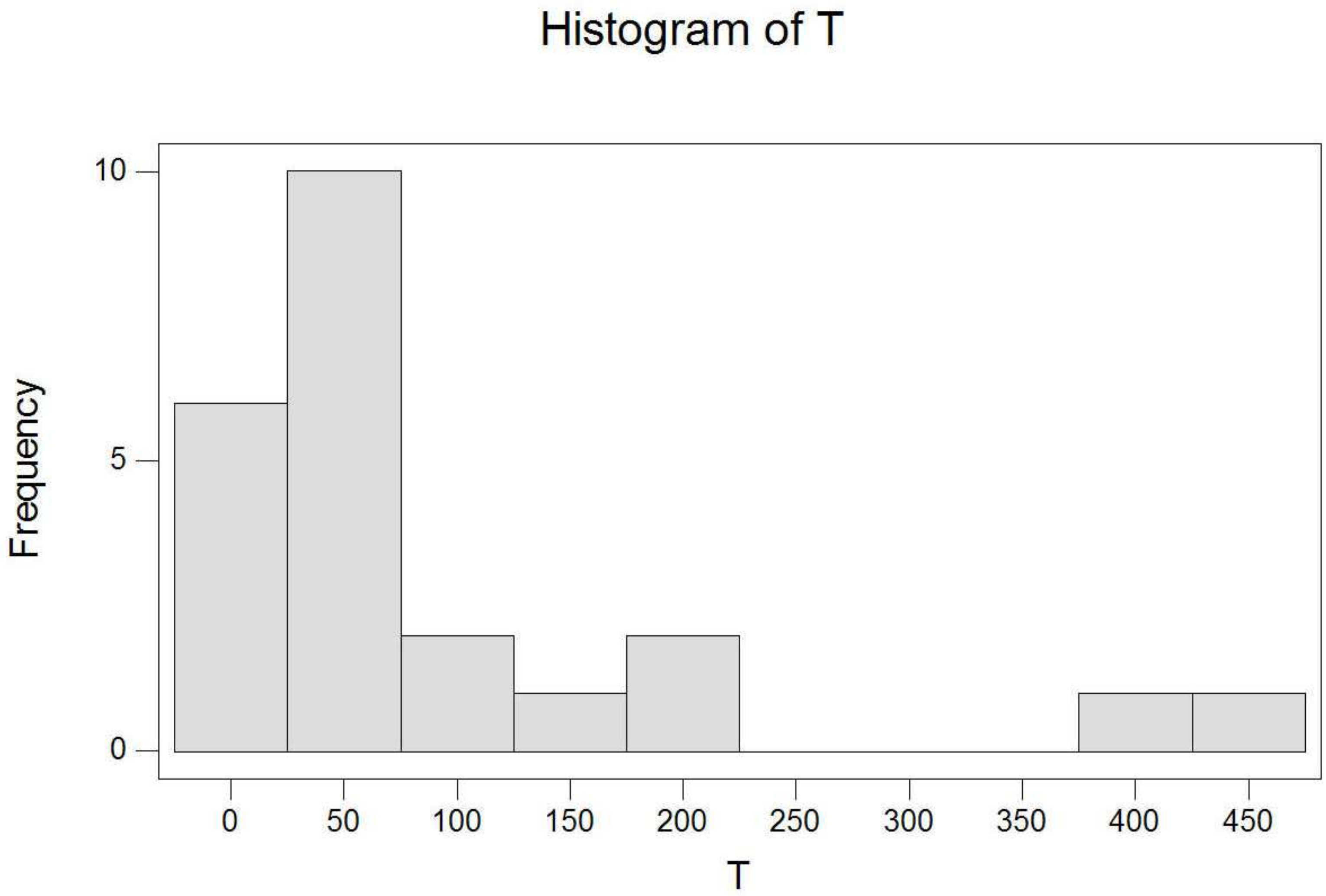

Ball Bearings Failure Data
ML Estimates - Complete Data
Probability Density Function
Lognormal base e Probability
Survival Function
Hazard Function
Percent
Probability
Rate
Location 4,1504
Scale 0,5217
MTTF 72,709
Failure 23
Censor 0
Goodness of Fit
AD* 0,647

Histogram of log(Rev), with Normal Curve

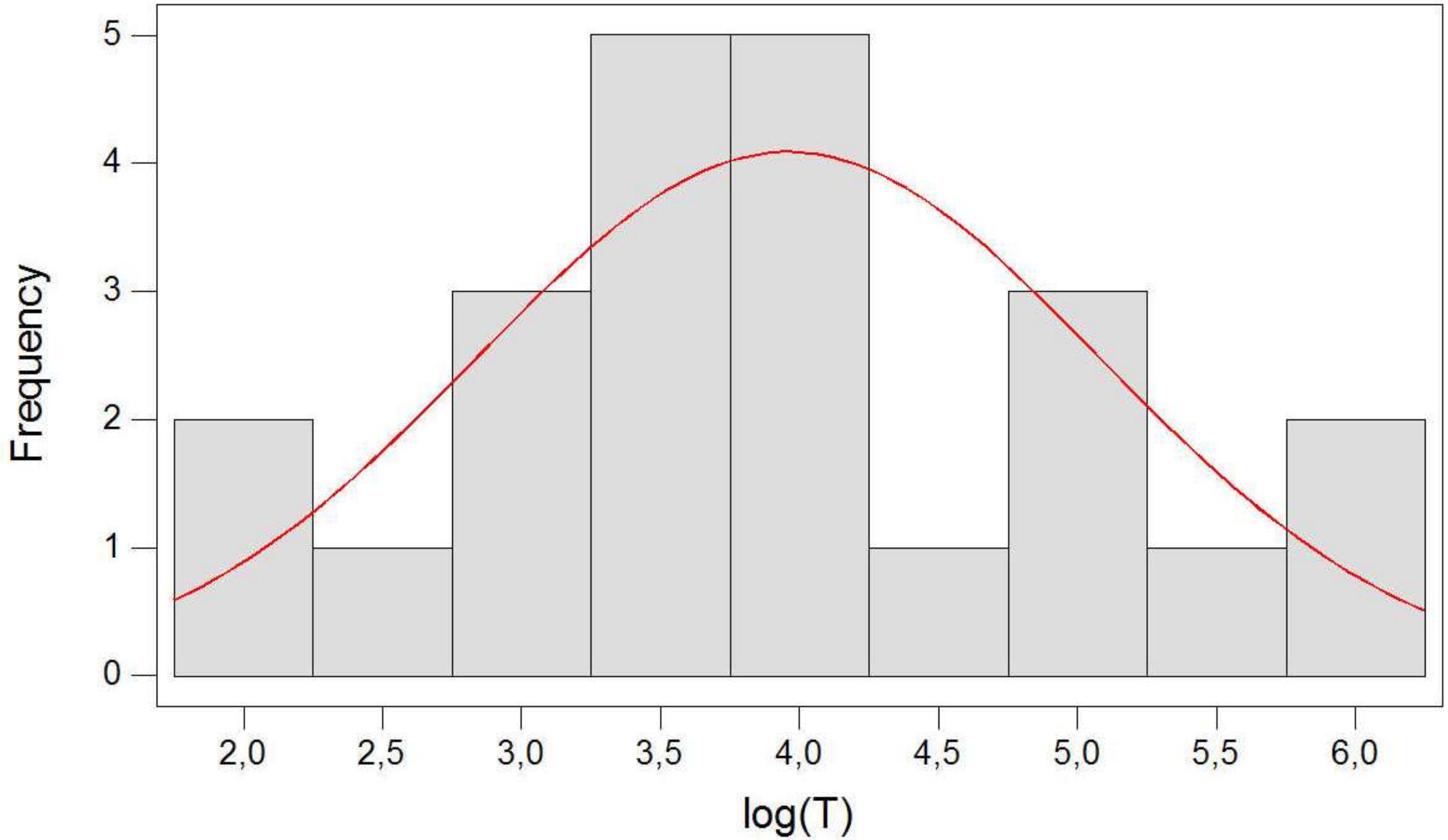


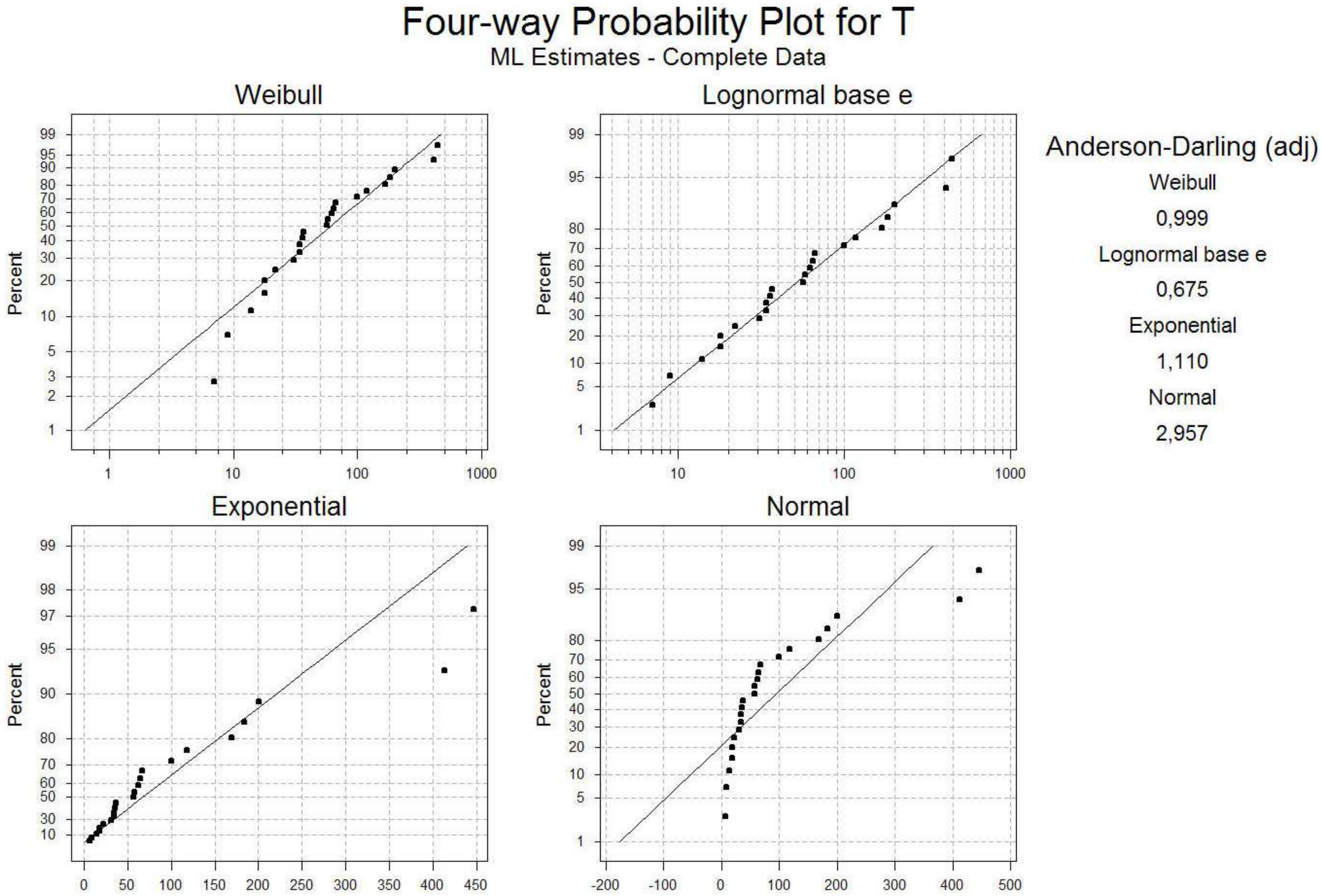
AIRCONDITION FAILURES ON BOEING AIRPLANES (PROSCHAN, 1963)

	C1 T	C2 log(T)
1	413	6,02345
2	14	2,63906
3	58	4,06044
4	37	3,61092
5	100	4,60517
6	65	4,17439
7	9	2,19722
8	169	5,12990
9	447	6,10256
10	184	5,21494
11	36	3,58352
12	201	5,30330
13	118	4,77068
14	34	3,52636
15	31	3,43399
16	18	2,89037
17	18	2,89037
18	67	4,20469
19	57	4,04305
20	62	4,12713
21	7	1,94591
22	22	3,09104
23	34	3,52636


Histogram of T
Frequency
10
5
0
0
50
100
150
200
250
300
350
400
450
T

Histogram of log(T), with Normal Curve




Four-way Probability Plot for T
ML Estimates - Complete Data
Weibull
Lognormal base e
Exponential
Normal
Percent
Anderson-Darling (adj)
Weibull
0,999
Lognormal base e
0,675
Exponential
1,110
Normal
2,957

Motivation for the exponential distribution

- Simplest distribution used in the analysis of reliability data.
- Has the important characteristic that its hazard function is constant (does not depend on time t).
- Popular distribution for some kinds of electronic components (e.g., capacitors or robust, high-quality integrated circuits).
- This distribution would not be appropriate for a population of electronic components having failure-causing quality-defects.
- Might be useful to describe failure times for components that exhibit physical wearout only after expected technological life of the system in which the component would be installed.

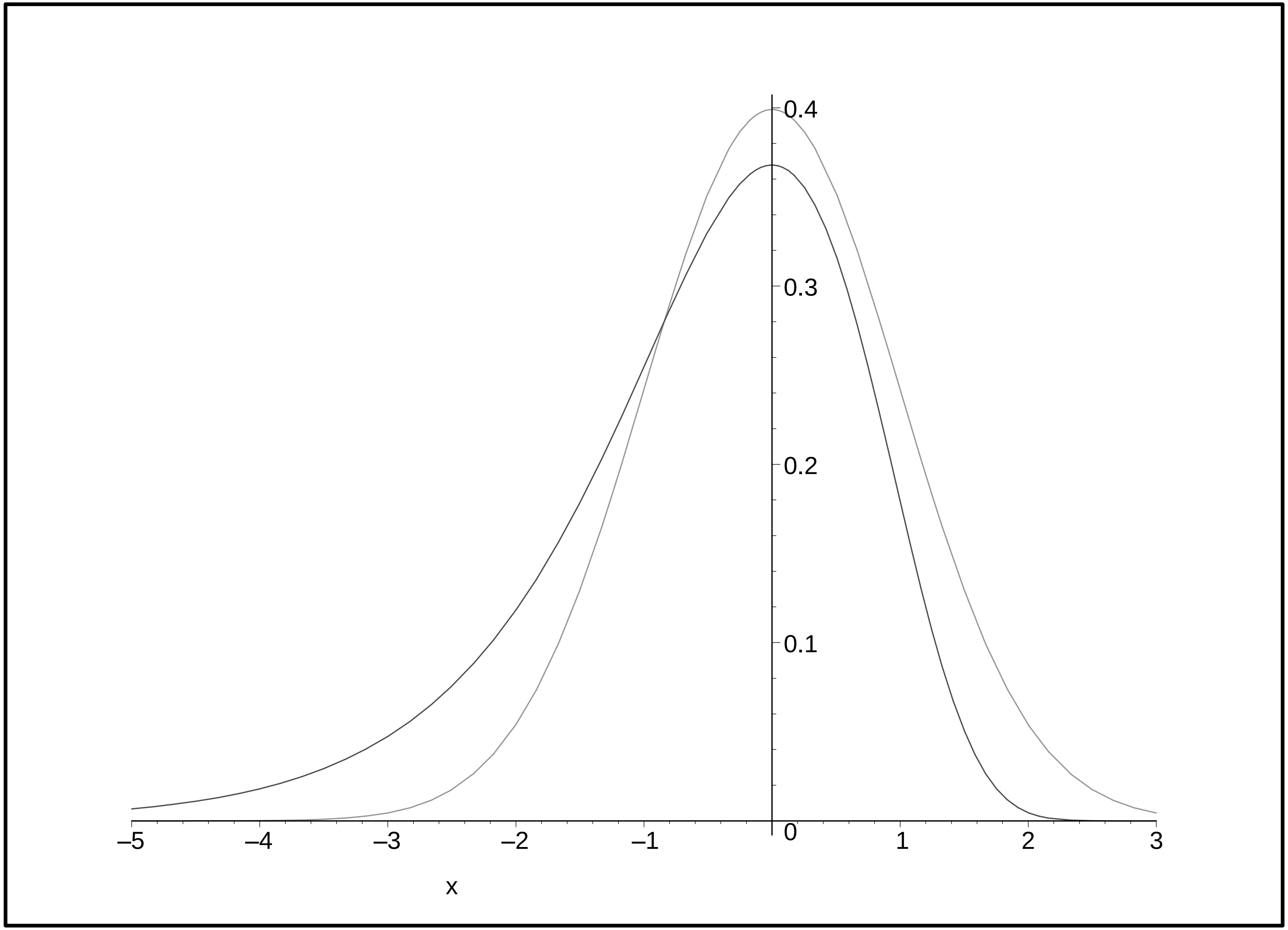
Motivation for the Weibull distribution

- The theory of extreme values shows that the Weibull distribution can be used to model the minimum of a large number of independent positive random variables from a certain class of distributions.
  - Failure of the weakest link in a chain with many links with failure mechanisms (e.g. fatigue) in each link acting approximately independent.
  - Failure of a system with a large number of components in series and with approximately independent failure mechanisms in each component.
- The more common justification for its use is empirical: the Weibull distribution can be used to model failure-time data with a decreasing or an increasing hazard function.

Motivation for lognormal distribution

- The lognormal distribution is a common model for failure times.
- It can be justified for a random variable that arises from the product of a number of identically distributed independent positive random quantities (remember central limit theorem for sum of normals).
- It has been suggested as an appropriate model for failure time caused by a degradation process with combinations of random rates that combine multiplicatively.
- Widely used to describe time to fracture from fatigue crack growth in metals.
- Useful in modeling failure time of a population electronic components with a decreasing hazard function (due to a small proportion of defects in the population).
- Useful for describing the failure-time distribution of certain degradation processes.

STANDARD GUMBEL AND NORMAL DISTRIBUTIONS $(\mu = 0, \sigma = 1)$



STANDARD LOGISTIC AND NORMAL DISTRIBUTIONS
$(\mu = 0, \sigma = 1)$