

第4回 続高橋セミナー

寿命試験データの統計解析

2015年4月10日

高橋 行雄

BioStat 研究所(株)

要約： 工業製品の通常的环境下での寿命を予測することは、長い時間かかるために過酷な条件下で製品が故障するまでの時間から、通常の使用状況下での製品寿命を推定することになる。加速(過酷)寿命試験では、事前に設定した試験期間になった場合に、対象製品が故障していなくとも試験を終了することになる。通常の使用条件下では、全く故障が起きない場合もあり、製品の寿命を予測する場合には、ある時間まで故障しなかった打ち切りデータを含めた統計解析が求められている。本書は、Excel および統計ソフト JMP を用いた最尤法によるワイブル分布のあてはめを導入し、さらに、ワイブル回帰の詳細を解説し、各種の信頼区間のために必須となる最小極値分布についても解説する。また、合成分散の一般式(デルタ法)についても各種の適用事例を示す。

目次

1. はじめに	1
2. 寿命試験データの例示	4
3. ワイブル分布のあてはめ	13
4. 最尤法による加速試験データの統計解析	35
5. JMP の寿命の二変量を用いた加速試験の解析	48
6. 寿命時間の対数変換による最小極値分布の活用	61
7. NEWTON-RAPHSON 法による尤度の最大化	77
8. 回帰分析の拡張	85
9. 多因子による寿命試験の事例	95
10. 合成分散の一般式(デルタ法)	102
11. プロファイル尤度を用いた95%信頼区間	109
12. JMP による対数尤度関数の偏微分	121
文献	126
索引	127
Excel, JMP ファイル 一覧	132

非売品, 無断複製を禁ずる

連絡先: 高橋 行雄, 〒105-0014 東京都 港区 芝 1-12-3

takahashi.stat@nifty.com, FAX : 03-342-8035

目 次

1. はじめに	1
2. 寿命データの例示	4
2.1 ディーゼル発電機のファンの故障データ	4
2.2 デバイス A に対する加速試験データ	7
2.3 段階的な学習	11
3. ワイブル分布のあてはめ	13
3.1 ワイブル分布	13
3.2 最尤法によるワイブル分布のパラメータ推定	15
3.3 最尤法による正規分布のパラメータ推定	17
3.4 打ち切りデータを考慮した生存率と故障率の推定	20
3.5 打ち切りデータを考慮した最尤法	24
3.6 JMP による解析	28
4. 最尤法による加速試験データの統計解析	35
4.1 加速試験データの統計解析の歴史的背景	35
4.2 多群の寿命データの統計解析	37
4.3 群ごとに異なるワイブル分布のあてはめ	38
4.5 WEIBULL β を共通とするワイブル分布のあてはめ	42
4.5 回帰パラメータを用いたワイブル分布のあてはめ	44
4.6 アレニウス変換温度	46
5. JMP の寿命の二変量を用いた加速試験の解析	48
5.1 寿命の二変量の基礎	48
5.2 温度を名義尺度にしたワイブル分布のあてはめ	50
5.3 温度を名義尺度だが WEIBULL β を共通	52
4.4 温度を連続尺度にしたワイブル回帰	54
5.5 アレニウス変換温度によるワイブル回帰	56
5.6 各種の推定	59
6. 寿命時間の対数変換による最小極値分布の活用	61
6.1 ワイブル分布を最小極値分布へ変換	61

6.2	最小極値分布を用いた最尤法	64
6.3	パラメータの分散共分散行列	67
6.4	パラメータの 95%信頼区間	70
6.5	時間 T における故障確率の 95%信頼区間	71
6.6	故障率に対する時間 T の 95%信頼区間	75
<hr/>		
7.	NEWTON-RAPHSON 法による尤度の最大化	77
7.1	WEIBULL パラメータの分散共分散行列の直接推定	77
7.2	NEWTON-RAPHSON 法によるワイブル分布での最尤法	80
7.3	NEWTON-RAPHSON 法による正規分布での最尤法	82
<hr/>		
8.	回帰分析への拡張	86
8.1	最尤法による加速試験データに対する回帰分析	86
8.2	EXCEL を用いた最尤法による回帰分析	87
8.3	JMP による解析	90
8.3	EXCEL による分散共分散の推定	92
<hr/>		
9.	多因子による寿命試験の事例	95
9.1	接着剤の寿命試験	95
9.2	接着剤の寿命試験に対する新たな解析法	97
9.3	今後の課題	101
<hr/>		
10.	合成分散の一般式 (デルタ法)	102
10.1	最小 2 乗法での場合	102
	パラメータに関して線形の場合	102
	パラメータに関して非線形の場合のデルタ法	103
	パラメータに関して線形の場合へのデルタ法の適用	104
10.2	最尤法における合成分散の一般式	105
	パラメータが 1 つの場合	105
	パラメータが複数の場合	105
10.3	知る人ぞ知るデルタ法	106
	寿命データ解析では	106
	回帰直線の推定値の分散での例	107
	デルタ法の統計検定 1 級での扱い	107
<hr/>		
11.	プロファイル尤度を用いた 95%信頼区間	109
11.1	プロファイル尤度による 95%信頼区間	109

11.2	パラメータの格子点での対数尤度	113
11.3	カイ 2 乗分布の下側確率によるプロファイル尤度	117
11.4	ワイブル分布での下側確率の等高線図	119
12.	JMP による対数尤度関数の偏微分	121
12.1	微分した式	121
12.2	微分した式による数値計算	122
12.3	JMP の微分関数と最大化関数	124
	文献	126
	索引	127
	EXCEL, JMP ファイル 一覧	132

1. はじめに

工業製品の通常的环境下での寿命を予測することは、長い時間かかるために過酷な条件下で製品が故障するまでの時間から、通常の使用状況下での製品寿命を推定することになる。加速(過酷)寿命試験では、事前に設定した試験期間になった場合に、対象製品が故障していても試験を終了することになる。通常の使用条件下では、全く故障が起きない場合もあり、製品の寿命を予測する場合には、ある時間まで故障しなかった打ち切りデータを含めた統計解析が求められている。

がん患者を対象にした臨床試験では、少しでも延命できる新たな治療法の確立を目指して数多くの臨床試験が実施されている。治療開始から死亡するまで、あるいは増悪するまでなどの時間に着目し、従来の治療法に代わる新しい治療法が優れていることを統計的に証明したい。がんの臨床試験では、試験期間が長期になることから試験の途中で追跡ができなくなる症例、あらかじめ定められた試験期間内に生存はしているが、強制的に観察を打ち切らねばならない症例もある。このような途中打ち切りを伴う生存時間データに対して各種の統計手法が開発され、多くの統計ソフトで手軽に統計解析が行えるようになってきた。

臨床試験における生存時間 (Survival Time) データの統計解析では、標準の治療法に対して新たに提案された治療法との比較が主体となり、ノンパラメトリックなログランク検定などの検定手法、比例ハザードモデルなどのセミパラメトリックな手法が主体となっている。

他方、寿命データの統計解析では、過酷な条件下での故障データから通常の使用条件下での製品寿命の推定が主体となるので、パラメトリックな生存時間解析が必須となる。歴史的には、寿命データ (Life Data) あるいは故障データ (Failure Data) 解析とも言われてきた。

多くの生存時間データの統計解析に関する教科書は、統計ソフトを使うことを前提としている。しかしながら、日本における寿命データの統計解析は、ワイブル確率紙、ワイブル型累積ハザード紙の使用を前提とした統計解析が主体となっている。この違いは何故なのだろうか。

寿命データの場合にはを前提としたパラメトリックな統計モデルによる解析が主体であるのに対し、生存時間データの統計解析は、ノンパラメトリックな Kaplan-Meier 法による生存曲線の推定、ノンパラメトリック検定、比例ハザードモデルによるセミパラメトリックな統計解析が主体となっている。

JMP などの統計ソフトなど標準的に使用されている最尤法による寿命データの統計解析は、日本ではマイナーな存在のようである。大橋靖雄・浜田知久馬 (1995)、の第 4.1 節「加速モ

デルによる解析 (LIFEREG プロシジャ)」にワイブル分布を仮定した解説があるが、工業分野における寿命データでの例示はない。

統計ソフト SAS における、寿命データの統計解析は、SAS/QC 関連のパッケージの中にある RELIABILITY プロシジャが主体となるが、これまで日本の SAS ユーザー総会で利用法が紹介されてことはない。

統計ソフト JMP の最新のバージョンは、品質管理・信頼性データについての統計解析機能の充実が著しく、多彩な寿命データの統計解析が可能となっている。ただし、日本語に翻訳されている 440 ページにもなる「品質管理および信頼性 / 生存時間」操作マニュアルは、解析手順を主体としており、基本から学習をしようとする人たちにとって優しくない。そこで、ワイブル確率紙、ワイブル型累積ハザード紙を使い慣れた技術者達を想定し、JMP による現代的な寿命データの統計解析にチャレンジしようと思いたくなるような入門書を作成することにした。

寿命データの統計解析で直面するのは、データが正規分布に従うのではなくワイブル分布に従うことを前提にしている、慣れ親しんだ統計解析の前提が全く異なっている。さらに、寿命データの統計解析では、打ち切りデータを考慮した解析が必須であり、このためにデータの平均・標準偏差などの標準的な要約統計量の算出も不適切となってしまう。回帰分析あるいは 1 元配置分散分析などの誤差分散が正規分布に従うことを前提とする標準的な統計解析も適用することができない。

打ち切りデータを含む寿命データの統計解析では、標準的な最小 2 乗法に変えて最尤法を用いた統計解析が必須となる。しかしながら、多くの書物に最尤法の考え方についての記述はあるものの、数理統計学の専門家ではない読者が、実データに即した最尤法の学習をしようと思っても推奨できる入門書は全く見い出せない。このことが、最尤法による寿命データの統計解析を目指す技術者達の学習意欲を阻害する要因となっている。

統計理論を学習し実際の問題を解決するための応用力を付けるためには、基本的な問題に対して「手計算」をしつつ学習を重ねることが必須である。統計解析を統計ソフトに委ねていては、理論を土台にした応用力は身につかない。とはいえ、最尤法の学習には、電卓を想定した「手計算」は、実施不可能である。

統計ソフトの普及に伴い、多くの統計手法は、誤用され続けている。使う人も誤用とは気が付かないし、誤用であるとの指摘できる人も極めて少ない。このような負の連鎖を少しでも食い止めたいと願っている。「手計算」の手段として表計算ソフト Excel が優れていると筆者は思っている。寿命データの統計解析に欠かせないワイブル分布の確率密度関数の計算、累積分布関数の計算のための関数、基本的な行列計算のための関数が用意されていることも

嬉しい。

実データを用いて最尤法によるワイブル分布のパラメータ推定を行うことも Excel のソルバーを併用することにより瞬時に求めることができる。このような「手計算」は、新しい統計解析の方法論を学ぶために必要不可欠と考える。

Excel のソルバーは、「手計算」を効率的に行う手段と筆者は思っている。しかしながら、Excel のソルバーでの解析では、統計的推測のために不可欠な推定したパラメータに関する分散共分散行列の出力がないため、それを補わなければならない。また、ワイブル分布のパラメータ推定に使われている最尤法の本質的な理解のためには、Excel の行列関数を用いた手作業によるニュートン・ラフソン法の逐次計算の経験が必要である。このことは、統計ソフト内での逐次計算の過程を真に理解することになる。

本書では、Excel での寿命データの解析を先に示し、その上で統計ソフト JMP による解析方法と結果を示すこととした。このように「手計算」による、寿命データの統計解析の基礎をしっかりと築くことが、現実の世界での応用力の根源となると確信している。

本書では、JMP の「品質管理および信頼性 / 生存時間」操作マニュアルで用いられているサンプルデータを用いている。これらのサンプルデータは、読者が JMP のサンプルデータファイルを容易に得ることができ、また引用した著書により、更なる継続的な学習を願っている。

パラメトリックな寿命データの統計解析の成書では、指数分布、正規分布、対数正規分布など多数の分布の使い分けについて論じられているが、本書では、実時間に対するワイブル分布、対数時間に対する最小極値分布に限定した。これは、指数分布がワイブル分布に含まれるためである

正規分布については、完全データについての最尤法に限定して導入したが、打ち切りデータにを含んだ解析には言及しなかった。これは、正規分布の累積分布関数が、積分の形式で定義されていて、扱い難いためであること、ワイブル分布に変えて積極的に使う現実的な例示が見い出せなかったたでもある。

2. 寿命データの例示

2.1 ディーゼル発電機のファンの故障データ

ディーゼル発電機のファンの保障期間を見直すために、これまで設置した 70 件のディーゼル発電機のファンの故障時間データを収集した結果を表 2.1 に示す。故障した記録があったのは 12 件であり、58 件のファンは故障していなかった。

これまでの保障期間は 8,000 時間としていたが、故障時間がワイブル分布に従うとして、その時間内に故障するファンの故障率と信頼区間の算出が要求された。さらに、このデータから、50%および 90%のファンが故障するまでの時間と信頼区間の算出も要求された。

表 2.1 ディーゼル発電機ファンの故障時間データ（単位は時間）

450	1850 +	2200 +	3750 +	4300 +	6100	7800 +	8750 +
460 +	1850 +	3000 +	4150 +	4600	6100 +	7800 +	8750 +
1150	1850 +	3000 +	4150 +	4850 +	6100 +	8100 +	9400 +
1150	2030 +	3000 +	4150 +	4850 +	6100 +	8100 +	9900 +
1560 +	2030 +	3000 +	4150 +	4850 +	6300 +	8200 +	10100 +
1600	2030 +	3100	4300 +	4850 +	6450 +	8500 +	10100 +
1660 +	2070	3200 +	4300 +	5000 +	6450 +	8500 +	10100 +
1850 +	2070	3450	4300 +	5000 +	6700 +	8500 +	11500 +
1850 +	2080	3750 +		5000 +	7450 +	8750	
+ 動作時間を示す		奥野忠一監訳（1988）、「寿命データの解析」, p293, 表8.1.1.					

最も早い故障は 450 時間で起こり、次の 460 時間のデータは故障ではなく、その時間まで動作（+）していることが確認されたが、観測が打切られたデータである。故障が確認されたファンは、450, 1150, 1150, 1600, 2070, 2070, 2080, 3100, 3450, 4600, 6100, 8750 時間の 12 件であり、最長の動作確認されたデータは 11,500 時間であった。

短い 460 時間、1560 時間などで打ち切りが発生するのは、ディーゼル発電機の設置が順次行われていて、ある観測時点から過去の設置記録を遡った時に 460 時間前に設置したディーゼル発電機が含まれていたことと解釈される。最長の 11,500 時間は、調査対象のディーゼル発電機が $11500/24=475$ 日前に設置され無故障で稼働（生存）していたことになる。

図 2.1 および 図 2.2 に示すように故障した 12 件の平均値は 3,047 時間、標準偏差は 2,399 時間、歪度は 1.43 であり右に裾を引いている。正規確率プロットでも故障データは、下に凸であり、正規分布を仮定することには、難点がある。故障していなかった 58 件の平均値は 5308 時間であり、故障した 12 件のファンの平均値より 2,260 時間よりも長い。

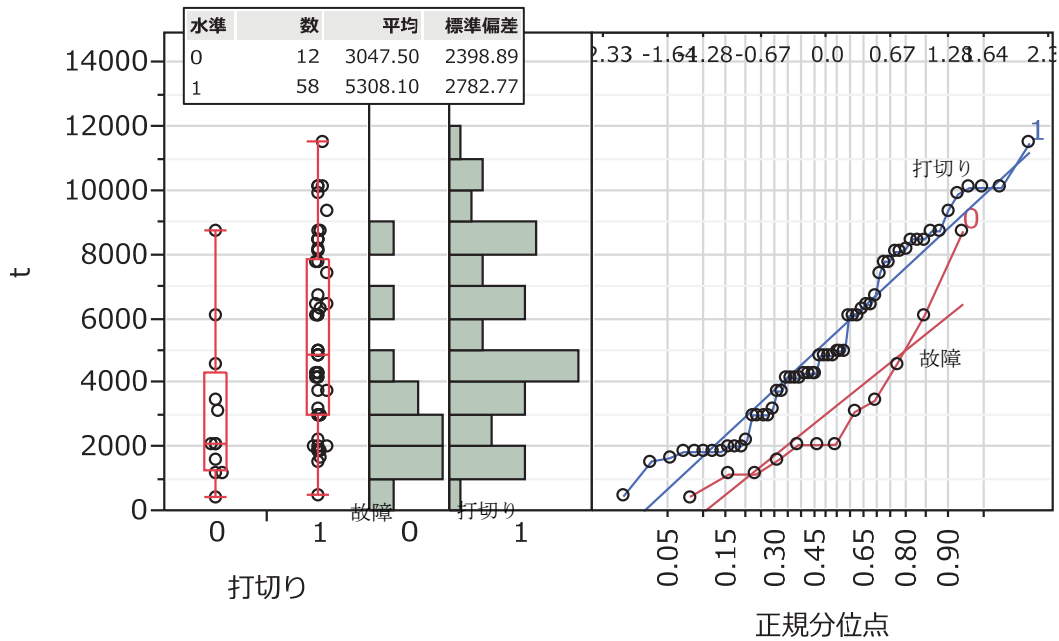


図 2.1 ディーゼル発電機ファンの故障データのプロット (JMP 二変量の関係)

もしも、70 件のファンが故障するまでの観察を続けていたとすれば、平均故障時間は、故障した 12 件のファンの平均値より大きくなることが明らかである。故障せずに観察が打ち切りとなったファンのデータも加味して寿命データの各種の統計量を算出するためには、どのような統計解析を行ったらよいのであろうか。

	故障 0	打ち切り 1
平均	3047.5	5308.1
標準偏差	2398.9	2782.8
N	12.0	58.0
歪度	1.4	0.3
尖度	1.8	-0.9

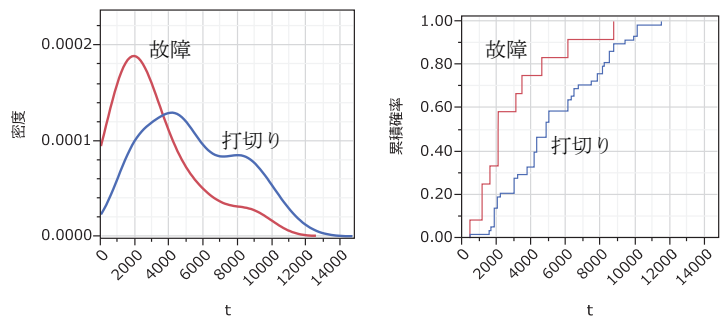


図 2.2 ディーゼル発電機ファンの故障データの確率密度と累積分布 (JMP 二変量の関係)

通常統計解析は、すべての観察データが得られている完全データにしか適用できない。打ち切りが含まれているような不完全データに対しては、異なる解析手法が必要となる。

奥野監訳本(1988)の309ページにワイブル型累積ハザード紙への12件の故障したファンのプロットに、最尤法によって得られた推定値を用いた直線および95%信頼区間が上書きされている。

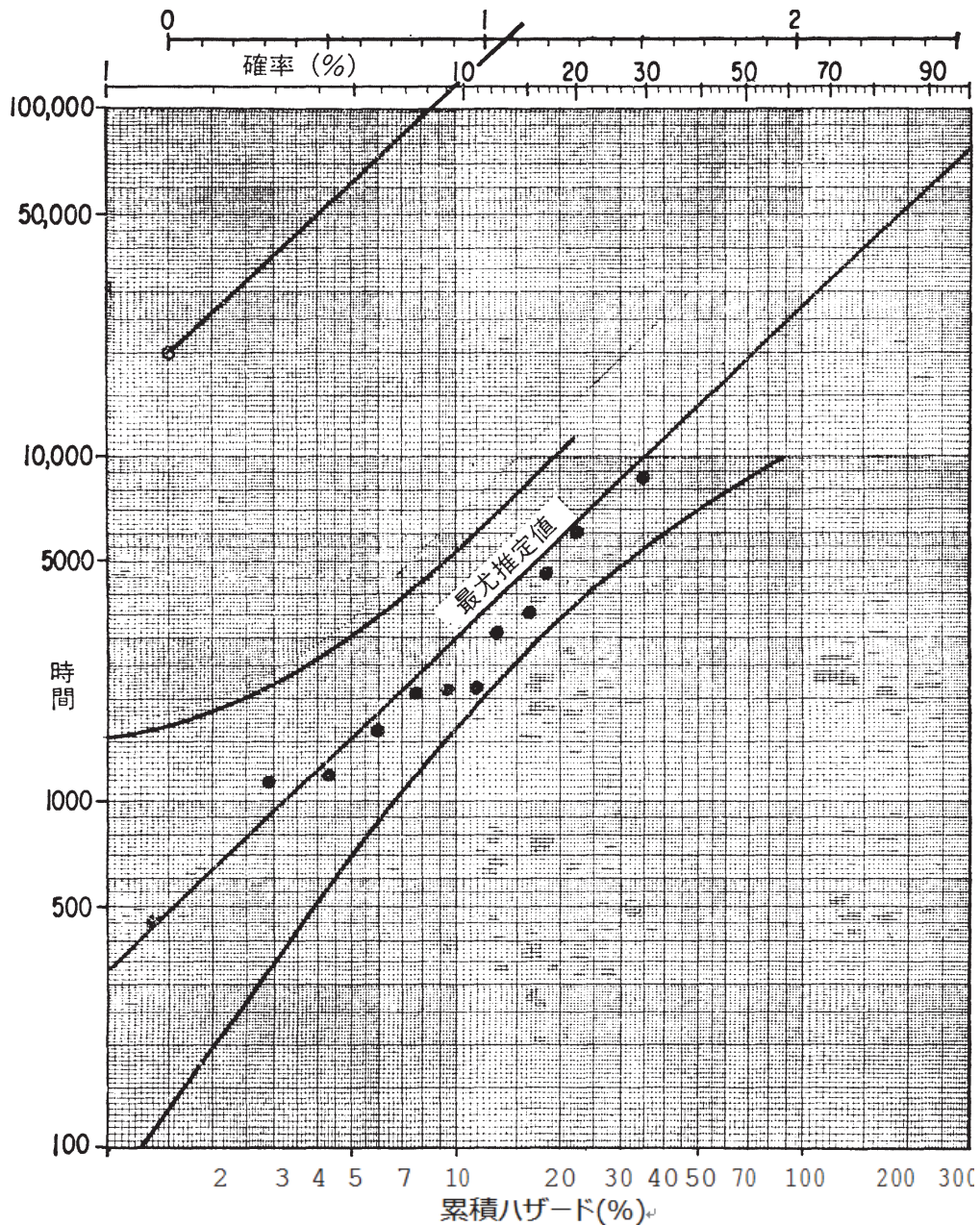


図 2.3 ワイブル型累積ハザード紙へのプロット (奥野監訳本. 図 8.3.1)

また、310 ページには、原著者 Nelson らが独自に開発した計算機プログラムによる最尤法によるパラメータの推定結果が示されている。

JMP の操作マニュアルの第 13 章 寿命の分布では、JMP のサンプル・ライブラリの Fun.jmp ファイルを用いて多くの紙面を割いた説明がされている。また、第 19 章の信頼性／生存時間でも多くの紙面が割かれている。ただし、JMP の操作マニュアルでは、多彩な JMP の解析機能の紹介を主体にしており、統計的な側面には、触れられていない。

2.2 デバイス A に対する加速試験データ

デバイス A の通常の使用条件である温度が 10°C における 10,000 時間および 30,000 時間での故障率を予測したい。試験期間の関係から 5,000 時間で観察を打ち切ることとなったので、高温状態での故障時間を測定し、10°C での寿命時間を予測することにした。

表 2.2 に示すように 10°C では、打ち切りの 5,000 時間で 30 件中すべてのデバイス A の故障は観測されなかった。40°C の条件下では 100 件中 10 件が 5,000 時間内で故障したが、残りの 90 件は故障しなかった。60°C では、20 件中 9 件、80°C では 15 件中 14 件が 5,000 時間内で故障し、1 件のみが故障せずに打ち切りとなった。

表 2.2 デバイス A の高温下での寿命試験結果（単位は時間）

温度			
10	40	60	80
5000 +	1298	581	283
(30件)	1390	925	361
	3187	1432	515
	3241	1586	638
	3261	2452	854
	3313	2734	1024
	4501	2772	1030
	4568	4106	1045
	4841	4674	1767
	4982	5000 +	1777
	5000 +	(11件)	1856
	(90件)		1951
			1964
			2884
			5000 +
			(1件)

Meeker and Escobar (1998), Chapter 19 Accelerated Life Test, p493.

図 2.4 に故障したデータについて各温度での分布の形状を示す。温度が 40°C の場合、5,000 時間の打ち切り以前に 100 件中 10 件が故障データし、左に裾を引く（値が小さい方）分布のようであり、温度が 60°C の場合は 20 件中 9 件が故障データで、わずかに右に裾を引いているものの正規確率プロットから正規分布に近く、温度が 80°C の場合は、15 件中 14 件の故障データではやや右に裾を引く分布となり、13 件は 2,000 時間以前に 1 故障している。

図 2.5 に打ち切りデータを無視した回帰直線を示す。デバイス A の温度が 10°C の場合には、おおよそ 5,000 時間と読み取れる。実際には、5,000 時間で 30 件中すべてのデバイス A の故障は観測されなかったのであるから、故障データのみでの解析は、極端な過小評価がおきてしまう。

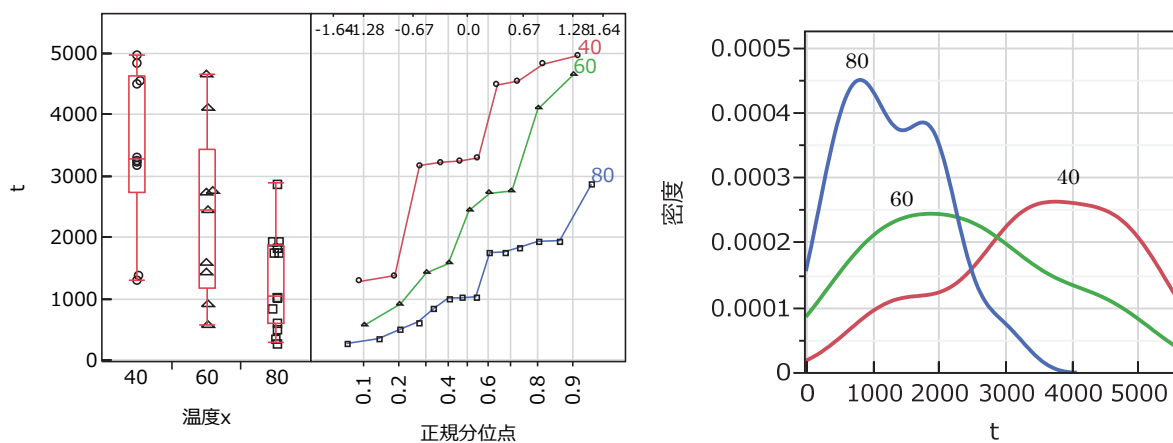


図 2.4 デバイス A の高温下での故障データのプロット (JMP 二変量の関係)

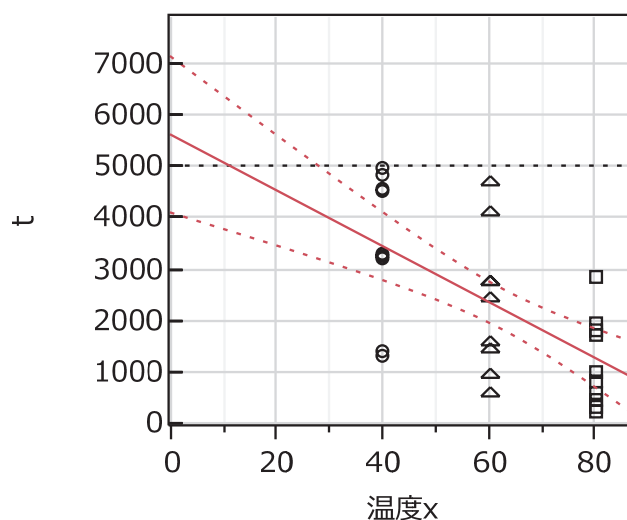


図 2.5 デバイス A の打ち切りを無視した回帰直線 (JMP 二変量の関係)

デバイス A の各温度で打ち切りデータを含めた故障率に対して、ワイブル確率紙へのプロットを 図 2.6(左) に示す. 温度 10°C のデータはすべて打ち切りデータなので, この図に示すことができない. 図(右) は, ノンパラメトリックな故障率曲線を示しており, 10°C の場合は, 故障率が 0.0 の水平線となっており. これらの方法では, 温度が 10°C の場合の寿命の予測は行えない.

加速試験データの統計解析は, 図 2.5 に示した回帰分析の拡張したものである. 一般的な回帰分析では, 回帰直線上から各々のデータ縦方向の偏差平方和が最小になるような回帰パラメータを推定 (いわゆる最小 2 乗法) している. ただし, この方法では, 打ち切りデータに対してまったく対応することができない.

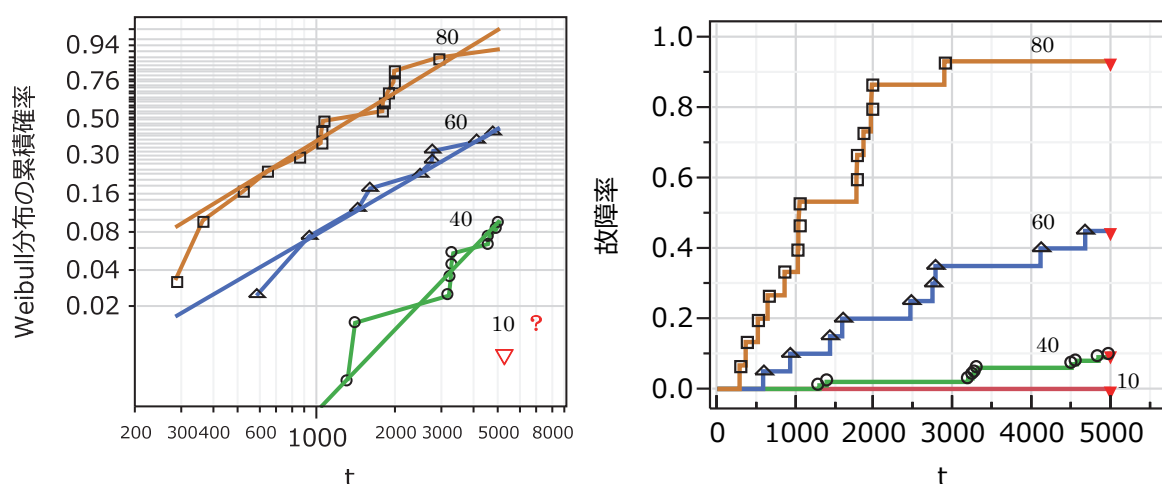


図 2.6 デバイス A のワイブルプロットおよび故障率プロット (JMP 生存時間分析)

寿命データの統計解析では、一般的な統計解析に対して、多くの拡張がなされている。第 1 の拡張は、最尤法の適用である。最尤法は、回帰パラメータの推定だけではなく、誤差分布が正規分布に従うとした場合には、正規分布の分散も平均値と同時に推定する方法での適用である。

第 2 の拡張は、打ち切りデータに対して、その時間における累積分布関数の上側確率を尤度とすることである。誤差が正規分布に従うとした場合に、データが故障データであれば、正規分布の確率密度関数 $f(t_i)$ を尤度とし、打ち切りデータであれば累積分布関数 $F(t_i)$ の上側確率 (生存関数) $S(t_i) = 1 - F(t_i)$ を尤度とする。そのために、全数が打ち切りの 10°C のデータも考慮した回帰直線より、10°C での製品寿命の予測が適切に行なえる。

図 2.7 では、下側確率が 0.10, 0.50, 0.90 の 3 本の直線が書き込まれている。全数が 5000 時で打切られた 10°C の場合に、下側確率 0.50 点は、12,500 時間と推定されている。このような、回帰分析が加速試験の統計解析で求められている。

第 3 の拡張は、時間 t をそのまま使うのではなく自然対数 $\ln(t)$ を使うことである。時間 t をそのまま使う場合には、表面上はワイブル分布を使っているように見える。ところが、内部の計算では、自然対数 $\ln(t)$ を用いた、最小極値分布が使われている。これは、ワイブル分布は、パラメータの大きさによって分布の形状が大きく異なり、信頼区間計算が複雑になること、また、べき乗が関数に含まれており、逐次計算の結果が発散しやすく不安定であるためである。

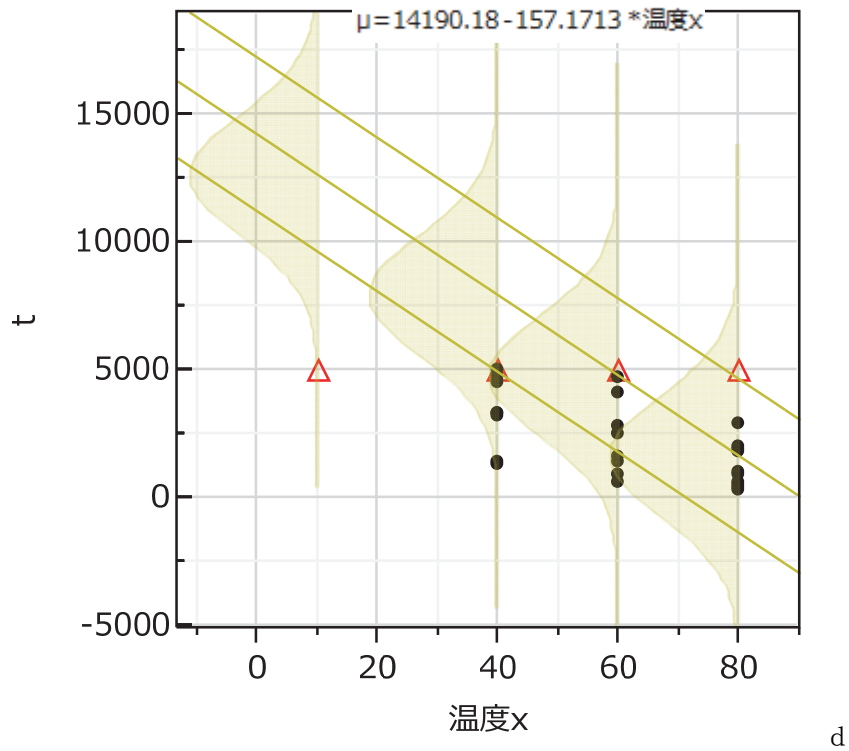


図 2.7 正規分布の確率密度関数と生存関数を用いた最尤法 (JMP 寿命の二変量)

最小極値分布は、正規分布と同様に一山型の分布で統計的に扱いやすからである。ワイブル分布と最小極値分布の関係は、正規分布と対数正規分布との関係に似ていて、互いに互換性があり、最小極値分布を用いた計算結果を、ワイブル分布の推定値に変換ができる。

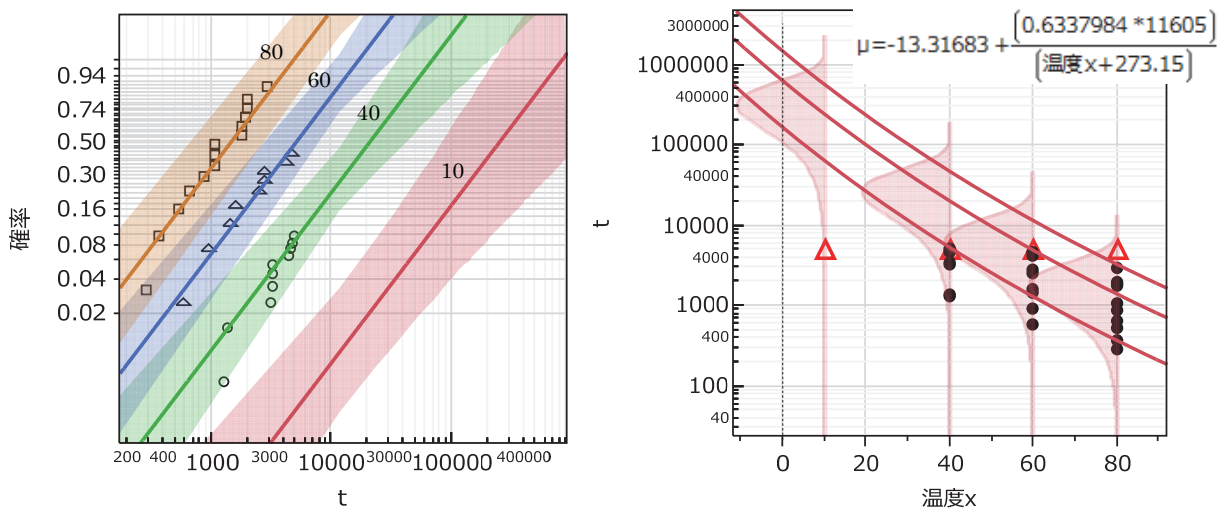


図 2.8 第 4 の拡張：アレニウス変換温度での解析 (JMP 寿命の二変量)

第4の拡張は、加速因子として温度の場合であれば、アレニウス変換が一般的に使われている。ただし、通常の摂氏温度での解析結果と全く異なるパラメータが推定されて、解釈がしにくくなる。図 2.8 (左) に示すように、摂氏温度を X 軸とした場合には、「直線」ではなく曲線となる。X 軸をアレニウス変換温度とすれば、直線関係となるが、解釈には摂氏温度への換算が必要となる。

第5の拡張は、加速因子が質的な場合、あるいは、加速因子が2つ以上ある場合、質的因子と量的因子が両方含まれている場合などへの拡張である。これは、1元配置、2元配置、直交表などの拡張の考え方と同様である。この場合には、JMP の「生存時間 (パラメトリック) のあてはめ」でワイブル分布を使った解析が必要となる。

2.3 段階的な学習

加速試験データの統計解析は、通常の統計の入門書では取り扱わない方法を多数含んでいる。本書は、回帰分析の使用経験があり、これから寿命データの統計解析にチャレンジしたい技術者を想定している。また、ワイブル確率紙を用いた寿命データの解析を行っているが、打ち切りデータを含めた最尤法について学習をしたい技術者にも、段階的な学習ができるよう配慮した。

臨床試験の統計解析では、統計ソフトを用いて比例ハザードモデルによる計算が、日常的に行なわれ、結果の解釈を中心とした教科書が数多く出版されている。ところが、その理論的背景を理解しているつもりではあるが、統計ソフトにたよらず最尤法による比例ハザードモデルを理解したいと思っても適切な教科書が見つからないので、手が付けようがない。などと、思っている臨床統計の関係者も想定している。

筆者も統計ソフトを日常的に使用してきたが、同じような悩みを持っていた。最初のチャレンジは、統計ソフトの行列計算で2値データに対するロジスティック回帰分析を、統計ソフト SAS の行列言語 IML (対話式行列計算言語) を用いて、ニュートン・ラフソン法による最尤法の計算方法を学習した。

第2のチャレンジは、SAS の MIXED プロシジャでの計算原理である制限付き最尤法 (REML) について、JMP の行列計算スクリプトの maximize 関数を用いて、複数誤差を持つ線形混合モデルについて計算原理について理解を深めた。

第3のチャレンジは、Excel の行列関数とソルバーを組み合わせ、実データでの Cox の比例ハザードモデルを Excel の1枚のシート上で実現したことである。この時でも、加速試験データの解析を Excel で実施することはできなかった。

第4のチャレンジは、第2番目の制限付き最尤法を Excel シート上で実現することであった。理論的背景を精査し、データの尤度とパラメータの尤度の比を最大化することに気が付き解決したことであった。

第5のチャレンジは、最尤解を求めるためのニュートン・ラフソン法などの逐次計算による最適解をソルバーの力を借りることなく Excel で実施することであった。最尤法の前に、対数尤度の1階の偏微分式で済む非線形最小2乗法を Excel で実現した。この経験に基づき、対数尤度の2階の偏微分式を用いるニュートン・ラフソン法を Excel によりようやく実現できた。ただし、最初は、実データを用いて、正規分布の「平均」と「分散」を最尤法で同時推定ことから始め、次いで、ワイブル分布の2つのパラメータの同時推定を実現することであった。

第6のチャレンジが、本書である。JMP の「寿命の二変量」で最小極値分布を誤差分散とする回帰分析の理論について、Excel で例示することであった。最後まで残ったのは、プロファイル尤度を用いた95%信頼区間を、自ら算出することであった。ただし、Excel の基本機能では、現在の筆者の力量では手が届かないことを認識し、それに代わる JMP の計算機能の助けを借りて、提示することにした。

このように、筆者も段階的な自己学習を継続し、ようやく加速試験データの統計解析を Excel でも実施できるようになった。この際に参考になったのが統計ソフト JMP の「寿命の二変量」であった。加速因子の温度をアレニウス変換しなければならないとの要求をどのように実データで示すかは難解であった。これについては、廣野(2000)が参考になった。

本書では、統計ソフト JMP の操作マニュアル「品質管理および信頼性 / 生存時間」の「寿命の一変量」で取り上げられている Nelson のディーゼル発電機のファンデータ (Fan.jmp)、
「寿命の二変量」で取り上げられている Meeker のデバイス A のデータ (Daval.Jmp) を用いた。統計解析の理論的な背景を簡潔に示し、実際の計算を Excel シート上で実現し、読者が追試できるように配慮した。これは、統計ソフト JMP を持っていない技術者に対しても段階的な学習ができるように配慮である。もちろん、JMP による解析が適切に行えるようにも配慮した。

3. ワイブル分布のあてはめ

3.1 ワイブル分布

ワイブル分布の累積分布関数 $F(t)$ は, t を寿命時間とし,

$$F(t) = 1 - \exp\left\{-\left(\frac{t}{\alpha}\right)^\beta\right\} \quad (3.1)$$

で与えられる, 尺度 (scale) パラメータを α , 形状 (shape) パラメータを β といい, 共に正である. 確率密度関数 $f(t)$ は, 累積分布関数 $F(t)$ を t で微分して

$$f(t) = \frac{\beta}{\alpha^\beta} t^{\beta-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^\beta\right\} \quad (3.2)$$

が得られる. 累積分布関数 $F(t)$ に $t = \alpha$ を代入すると β にかかわらず,

$$F(t = \alpha) = 1 - e^{-1} = 0.632 \quad (3.3)$$

と計算される. したがって, ワイブル分布の尺度パラメータ α は, 下側確率 0.632 となり, おおよその “平均” を示す.

図 3.1 に尺度パラメータを $\alpha = 10$ と固定し, 形状パラメータを $\beta = 0.5, 1, 2, 4, 6$ と変化させたときのワイブル分布の確率密度関数 (左) と累積分布関数 (右) を Excel の Weibull 関数を用いて計算し, Excel により作成した線グラフを示す.

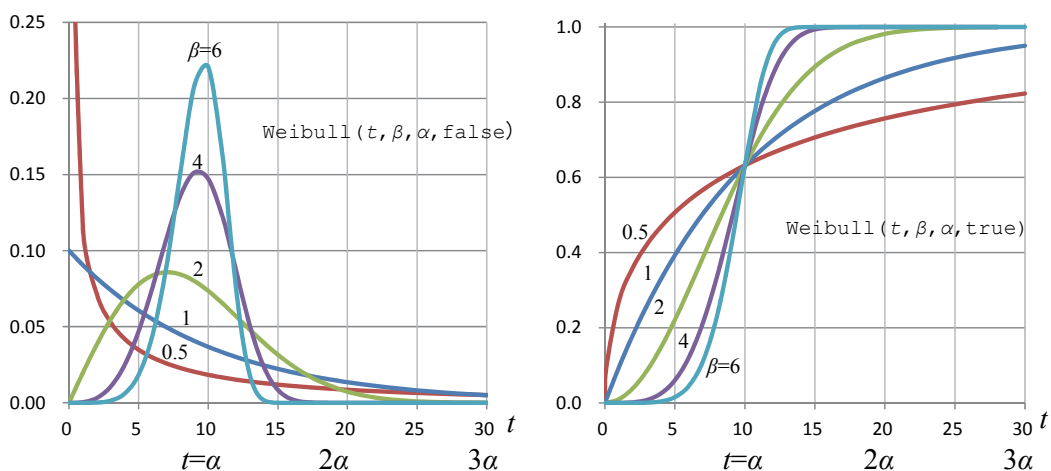


図 3.1 ワイブル分布の確率密度関数 (左) と累積分布関数 (右)

ワイブル分布の尺度パラメータを $\alpha=10$ としたので、累積分布関数では $t=10$ で形状パラメータが $\beta=0.5, 1, 2, 4, 6$ と異なっても同じ下側確率 0.632 を通ることが図（右）から確認される。確率密度関数の場合は、 $\beta \leq 1$ 以下の場合に単調減少の形状となり、 $\beta > 1$ の場合には一山型の形状で、山のピークは右に移動する。形状パラメータが $\beta=4$ の場合には正規分布に近く、 $\beta > 4$ の場合には、確率密度関数の山がさらに高く、左に裾を引くようになり、幅も狭まる。

本節のワイブル分布の 2 つのパラメータに対して、尺度パラメータを α 、形状パラメータを β と表記した。しかしながら、多くの別表記が混在している。日本の信頼性関係の書物では、

$$\text{尺度パラメータを } \eta = \alpha,$$

$$\text{形状パラメータを } m = \beta$$

と表記している。Excel の関数では、

$$A = \beta, \quad B = \alpha$$

としている。あるいは、別の書物では、

$$\alpha = \beta, \quad \beta = \alpha$$

と逆の表記法もあり、混沌とした状況である。

JMP では、ワイブル分布の 2 つのパラメータに対して、Weibull α 、Weibull β と表記して、（尺度、形状）のを使うことを避けている。寿命データ t の自然対数 $y = \ln(t)$ を取って最小極値分布にした場合には、Weibull パラメータ α も自然対数を取って極値パラメータ

$$\lambda = \ln(\alpha) \tag{3.4}$$

とし、Weibull β は逆数を取り、最小極値パラメータとして

$$\delta = 1/\beta \tag{3.5}$$

を用いている。

JMP に新たに加わった「寿命の二変量」では、

$$\text{位置パラメータ : } \mu = \lambda = \ln(\alpha) \tag{3.6}$$

$$\text{尺度パラメータ : } \sigma = \delta = 1/\beta \tag{3.7}$$

としている。本節で尺度パラメータ α を使ったのであるが、「寿命の二変量」では β の逆数を尺度パラメータとして使っている。これらの多彩な表記法は、統計解析の歴史的な発展過程の産物である。そのために、読者が混乱しないように、JMP の表記法に合わせて Weibull α 、Weibull β を基本とした記述とする。

3.2 最尤法によるワイブル分布のパラメータ推定

ある n 個の寿命データ t_i から平均値と不偏分散を求める場合に、最初に算術平均 $\hat{\mu}$ を求め、 $\hat{\mu}$ からの偏差平方和を自由度 $n-1$ で割って不偏分散を求める。ディーゼル発電機のファンが故障したのは、450, 1150, 1150, 1600, 2070, 2070, 2080, 3100, 3450, 4600, 6100, 8750時間の12件である。算術平均は $\hat{\mu}=3047$ となり、算術平均からの偏差の平方は63,301,625、不偏分散は5,754,693となり、その平方根から標準偏差が2,399と計算される。

さて、ファンの故障データがワイブル分布に従うと仮定し、その Weibull α , Weibull β をどのように推定するのであろうか。歴史的には、ワイブル確率紙に累積故障率をプロットし、目の子で回帰直線をあてはめてワイブル分布のパラメータを推定していた。

統計ソフト JMP では、ワイブル確率紙へのプロット、回帰直線をあてはめたグラフ表示を行ってはいるが、内部では最尤法によるワイブル分布のパラメータを直接推定し、その結果をグラフ表示に用いている。

故障せずに打ち切りとなった58件のファンのデータを含めた最尤法は、3.4節で示すことにし、ここでは故障した12件のみを用いてワイブル分布のパラメータを最尤法で推定する方法を示す。

故障した12件のファンがワイブル分布に従うとし、その確率密度関数 $f(t_i)$, $i = 1, 2, \dots, r$ の積を尤度 L

$$L = \prod_{i=1}^r L_i = \prod_{i=1}^r f(t_i) = \prod_{i=1}^r \frac{\beta}{\alpha^\beta} t_i^{\beta-1} \exp\left\{-\left(\frac{t_i}{\alpha}\right)^\beta\right\} \quad (3.8)$$

とする。まず、故障した r 個のデータから目の子で推定される Weibull パラメータ α と β の初期値とする。Weibull α は、下側確率が0.632となる時間 t なので、 $12 \times 0.632 = 7.58 \approx 8$ となり、小さい方から8番目の $\hat{\alpha} = 3100$ 時間を初期値とする。Weibull β は、初期故障が散見され、右に裾を引いているので 図 3.1 を参考に $\hat{\beta} = 1.2$ とする。

図 3.2 に示すように

$$\begin{aligned} L_1 &= f(t_1) = 0.000238, \\ L_2 &= f(t_2) = 0.000234, \\ &\vdots \\ L_{12} &= f(t_{12}) = 0.000015 \end{aligned}$$

となり、それらの積 L は

$$L = \prod_{i=1}^{12} L_i = \prod_{i=1}^{12} \left\{ \frac{1.2}{3100^{1.2}} t_i^{0.2} \exp\left[-\left(\frac{t_i}{3100}\right)^{1.2}\right] \right\} = 2.110 \times 10^{-47}$$

と計算される。なお、この計算には、Excel のワイブル分布の確率密度関数 WEIBULL () を用いている。データ数が多いと指数部分がさらに小さくなり、コンピュータ内部で数値として取り扱えなくなる。したがって、積として定義されている尤度 L は対数を取ることで積ではなく和となり数値計算が容易になる。

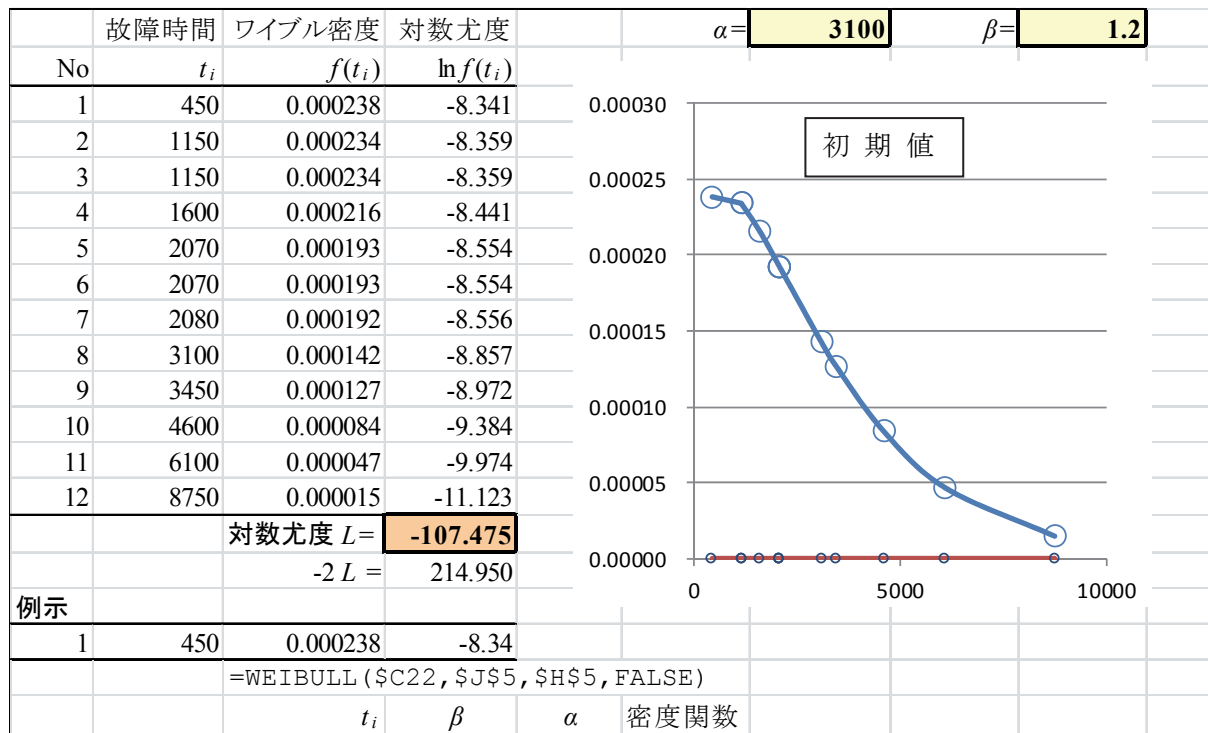


図 3.2 ワイブル分布の初期値に対する確率密度の計算と線グラフ

最尤法では、式 (3.9) で示す対数尤度 $\ln L$ が標準的に用いられている。

$$\ln L = \sum_{i=1}^{12} \ln[L_i] = \sum_{i=1}^{12} \ln[f(t_i)] = \sum_{i=1}^{12} \ln \left\{ \frac{\beta}{\alpha^\beta} t_i^{\beta-1} \exp \left[-\left(\frac{t_i}{\alpha} \right)^\beta \right] \right\} \quad (3.9)$$

実際に計算すると $\ln f(t_1) = -8.341$, $\ln f(t_2) = -8.359, \dots, \ln f(t_{12}) = -11.123$ となるので対数尤度は、

$$\ln L = \sum_{i=1}^{12} \ln[f(t_i)] = -107.48$$

と扱いやすい数値となり、データ数が 100 倍になったとしても、コンピュータの内部計算の支障とはならない。

計算された対数尤度 $\ln L = -107.475$ を最大にするように、試行錯誤的に Weibull パラメータ α と β を逐次的に変化させる必要がある。この初期値 $\hat{\alpha} = 3100$, $\hat{\beta} = 1.2$ を前後に動かしながら、 $\ln L = -107.475$ が最大になるよう探索する。Excel シート上で結果は省略するが、 $\hat{\beta} = 1.4$ と大きくすると $\ln L = -107.282$ のように 0.193 大きくなる。次に Weibull $\hat{\alpha} = 3400$ と大きくす

ると $\ln L = -107.206$ のようにさらに 0.076 大きくなる.

注) このような逐次的な作業を代行する機能が, Excel のアドインとして提供されているソルバーである. ソルバーを用いた最尤法による統計解析の詳細は, 芳賀敏郎 (2004), 「最小 2 乗法, 最尤法, 線形モデル, 非線形モデル」 <http://www.yukms.com/biostat/haga/download/archive/likelihood/Likelihood.pdf> が詳しいので参照してもらいたい. 筆者も, この文献によりソルバーの有用性を認識し, 愛用している.

Excel のソルバーを用いて, 対数尤度 $\ln L$ が最大になるように初期値 $\hat{\alpha} = 3100$, $\hat{\beta} = 1.2$ を変化させると, 図 3.3 に示すように $\hat{\alpha} = 3370$, $\hat{\beta} = 1.415$ が得られる. 手作業での $\hat{\alpha} = 3400$ と $\hat{\beta} = 1.4$ の組合せに比べて, さらに $\ln L = -107.20$ と 0.01 大きくなっていることがわかる.

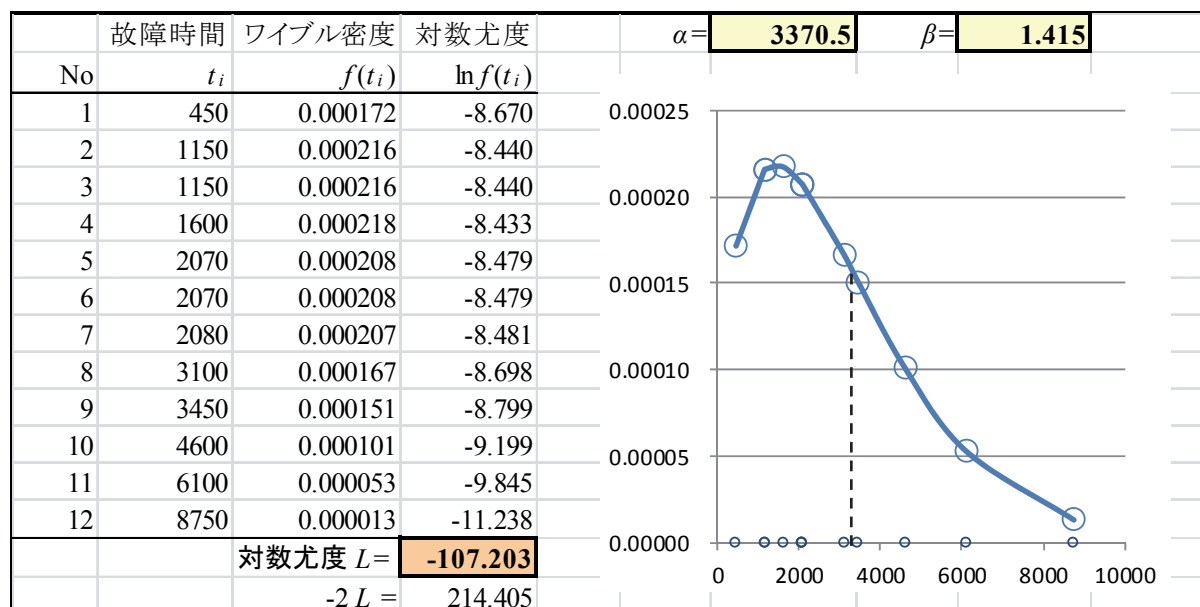


図 3.3 対数尤度を最大化した結果

図 3.3(右)は, X 軸を時間 t , Y 軸をワイブル分布の確率密度 $f(t_i)$ の点を Excel の散布図で描き, 平滑化曲線により線を加えたものである. 図中の縦の点線は, X 軸の $\hat{\alpha} = 3370$ 点であり, 下側確率は 0.632, $\hat{\beta} = 1.415$ と 1.0 よりも大きいので右に裾を長く引く形状となっている.

3.3 最尤法による正規分布のパラメータ推定

前節では, 最尤法によるワイブル分布のパラメータ推定をいきなり行った. 段階的な学習のためには, データが正規分布に従うと仮定する場合から始めることが望ましい. 得られたデータから平均値と標準偏差を推定する場合に, データの分布がどのようなものであっても,

平均値 μ は、算術平均として推定する。標準偏差の推定値 $\hat{\sigma}$ は、推定平均 $\hat{\mu}$ を用いて計算された不偏分散の平方根として計算する。このことは、得られたデータが正規分布に従うことがわかっているにもかかわらず、その情報を全く使用することがない。正規確率プロットなどで、事後的に得られたデータが正規分布に従うことを事後的に確認するだけである。

最尤法を用いる場合は、データが正規分布に従うことを前提に、データの確率密度から、平均値と標準偏差を同時に推定する。最尤法を用いることによりワイブル分布にかぎらず、他の多くの分布のパラメータを直接推定することができる。最尤法は、計算が煩雑なのでほとんどの書物に例示がないが、多くの統計ソフトでは内部の計算に多用されている。そこで、データが正規分布に従うとした場合に、平均 μ と標準偏差 σ を最尤法により同時に推定する手順を示す。

図 3.4 は、ワイブル分布についての図 3.3 の Excel シートを正規分布用に変更したものである。変更は、 $f(t_i)$ の計算をワイブル分布の確率密度関数から正規分布の確率密度関数とし、 $\hat{\alpha}$ に代えて $\hat{\mu}$ 、 $\hat{\beta}$ に代えて $\hat{\sigma}$ とした点である。

実際に、対数尤度 $\ln L$ を最大化するように、 $\hat{\mu}$ と $\hat{\sigma}$ を変化させると、対数尤度は $\ln L = -109.898$ となり、図 3.3 に示したワイブル分布の最尤解での対数尤度 -107.203 に比べて 2.695 小さく、あてはまりが良くないことが示されている。

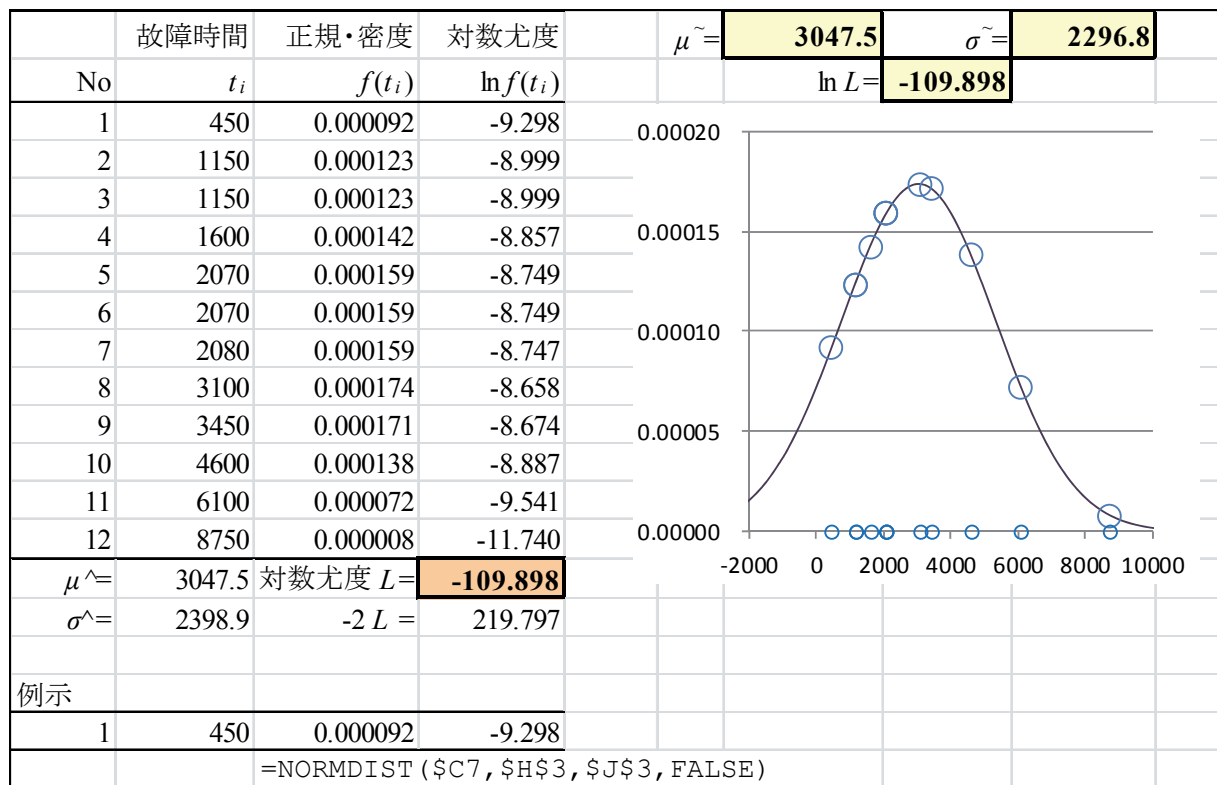


図 3.4 正規分布のパラメータの最尤法による同時推定

最尤法で推定されたパラメータ $\hat{\mu}=3047.5$ は、算術平均 $\hat{\mu}$ に等しいが、標準偏差 $\hat{\sigma}=2296.8$ は、不偏分散の平方根として求められた $\hat{\sigma}=2398.9$ とは異なる。これは、不偏分散の計算に際して、算術平均自体の揺らぎを考慮するために偏差平方和を $r-1=11$ で除しているためである。最尤法での標準偏差 $\hat{\sigma}=2296.8$ は、データ数 r で除した結果となっている。

最尤法では正規分布の確率密度関数 $f(t_i)$ を用いているのに、なぜ「尤度」と別の用語とするのだろうか。これは、 $t_1=450$ の場合の確率密度は $f(t_1)=0.000092$ と計算されているが、平均値 $\hat{\mu}$ と標準偏差 $\hat{\sigma}$ が変われば、 $f(t_i)$ の値も連続的に変化するので確率密度ではなく likelihood (尤度) と区別して使われている。

図 3.5(左) に、最尤推定量 ($\hat{\mu}=3047.5$, $\hat{\sigma}=2296.8$) の場合の $f(t_1)=0.000092$ が示されていて対数尤度は -109.898 である。図の中央は、 $\hat{\mu}=4000$ と変化させた場合で $f(t_1)=0.000053$ と小さくなり、全体の対数尤度は -110.930 と小さくなり、図の右で $\sigma=3000$ とすると $f(t_1)=0.000066$ と 0.000092 に比べて小さく、対数尤度は -111.225 とさらに小さくなる。最尤法は、個々のデータに対する尤度ではなく r 個のデータ全体の尤度の積 (対数尤度の和) を最大にするような分布のパラメータを求める方法である。

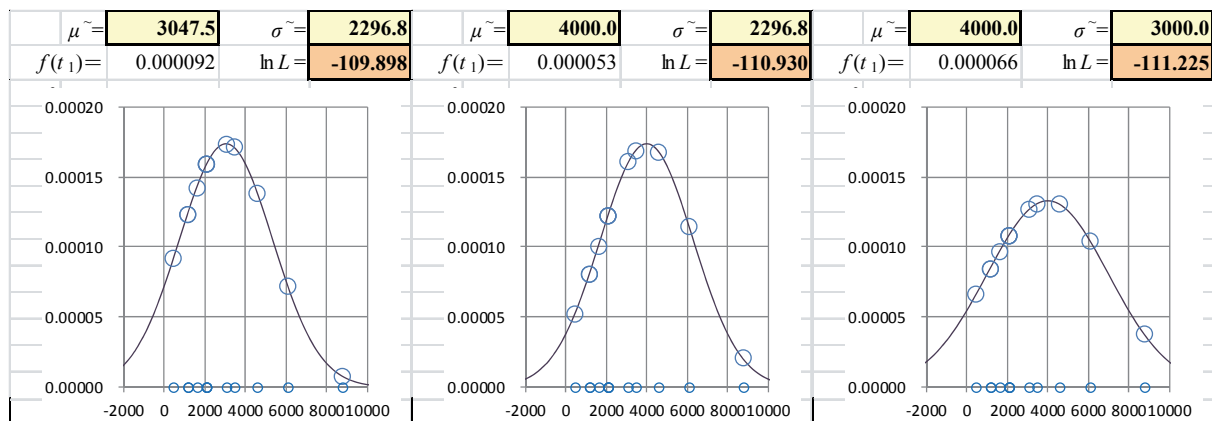


図 3.5 正規分布のパラメータを変化させたときの対数尤度の変化

JMP の「寿命の一変量」によっても表 3.1 に示すように図 3.4 と同様に Excel と同様に最尤法により正規分布のあてはめが行える。「寿命の一変量」には、多彩な分析が可能であり、

表 3.1 正規分布を選択した場合のパラメータの推定 (JMP 寿命の一変量)

パラメータ	推定値	標準誤差	下側95%	上側95%	規準	
位置 $\hat{\mu}$	3047.50	663.02	1748.01	4346.99	(-2)*対数尤度	219.7967
尺度 $\hat{\sigma}$	2296.77	468.83	1377.89	3215.65	AICc	225.1300
					BIC	224.7665

その選択および結果の解釈を適切に行うために、Excel による分析での基礎を作っておくことが肝要である。パラメータの欄の「位置」は $\tilde{\mu}$ であり、「尺度」は $\tilde{\sigma}$ である。標準誤差、信

頼区間の欄についての Excel による計算方法については、第 7 章を参照してもらいたい。基準の欄の (-2) 対数尤度 = 219.7967 となっているのは、図 3.5 の対数尤度の -2 倍で、 $-2\ln(L) = 219.947$ に対応する。

図 3.6 に JMP で正規分布の位置 $\tilde{\mu} = 4000$ 、尺度 $\tilde{\sigma} = 3000$ と変化させたときの確率密度関数 $f(t_i)$ と対数尤度を示した。JMP でも図 3.5 と同様に、ある特定の値をパラメータとして設定した場合の確率密度関数、パラメータに関する対数尤度の図および対数尤度の推定も可能である。さらなる応用は、第 11 章で示す。

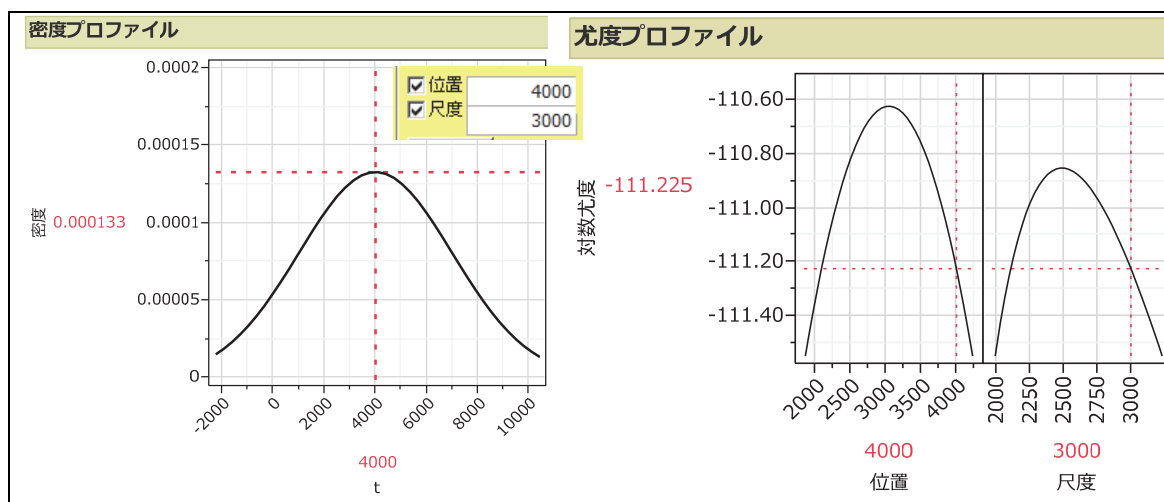


図 3.6 $\tilde{\mu} = 4000$ ， $\tilde{\sigma} = 3000$ の場合の各種尤度曲線 (JMP 寿命の一変量)

図 3.6 (左) は、平均値 (位置) を 4,000、標準偏差 (尺度) を 3,000 としたとき、時間を左右に振った時の正規分布の確率密度曲線であり、図 3.5 (右) と同等に平均が 4,000 の場合に密度が 0.000133 となることが示されている。右は、平均 (位置) および標準偏差 (尺度) を変化させた場合の対数尤度のプロファイル曲線であり、平均が 4,000、標準偏差が 3,000 の場合に対数尤度が -111.225 となることが表示されている。

3.4 打ち切りデータを考慮した生存率と故障率の推定

故障した $r=12$ のデータについてワイブル分布の確率密度関数のパラメータ推定を最尤法により示し結果は、故障しないで打ち切りとなったデータを全く考慮していない。したがって、推定された Weibull パラメータ $\hat{\alpha} = 3,370$ は“平均的”な故障時間よりも極端に短くなっていることが推測される。

最初の故障は表 3.2 に示すように 450 時間であった。この時間直前に稼働していたファンは全体の 70 件であるので、450 時間の故障率は、 $1/70=0.014$ と推定され直後の生存率（稼働率）は $1-0.014=0.986$ と推定される。

次の 460 時間は観察が中止され打ち切りデータであるが、その後も故障しないで稼働していると推測すると生存率は 0.986 のままとするのが妥当である。

表 3.2 故障率および生存率の推定

No	t +	故障	瞬間 故障率	生存率 稼働率	故障率	No	t +	故障	瞬間 故障率	生存率 稼働率	故障率
1	450	1 / 70	0.014	0.986	0.014	36	4300 +			0.852	
2	460 +			0.986		37	4600	1 / 34	0.029	0.827	0.173
3	1150	1 / 68	0.015	0.971	0.029	38	4850 +			0.827	
4	1150	1 / 67	0.015	0.957	0.043	39	4850 +			0.827	
5	1560 +			0.957		40	4850 +			0.827	
6	1600	1 / 65	0.015	0.942	0.058	41	4850 +			0.827	
7	1660 +			0.942		42	5000 +			0.827	
8	1850 +			0.942		43	5000 +			0.827	
9	1850 +			0.942		44	5000 +			0.827	
10	1850 +			0.942		45	6100	1 / 26	0.0385	0.795	0.205
11	1850 +			0.942		46	6100 +			0.795	
12	1850 +			0.942		47	6100 +			0.795	
13	2030 +			0.942		48	6100 +			0.795	
14	2030 +			0.942		49	6300 +			0.795	
15	2030 +			0.942		50	6450 +			0.795	
16	2070	1 / 55	0.018	0.925	0.075	51	6450 +			0.795	
17	2070	1 / 54	0.019	0.908	0.092	52	6700 +			0.795	
18	2080	1 / 53	0.019	0.891	0.109	53	7450 +			0.795	
19	2200 +			0.891		54	7800 +			0.795	
20	3000 +			0.891		55	7800 +			0.795	
21	3000 +			0.891		56	8100 +			0.795	
22	3000 +			0.891		57	8100 +			0.795	
23	3000 +			0.891		58	8200 +			0.795	
24	3100	1 / 47	0.021	0.872	0.128	59	8500 +			0.795	
25	3200 +			0.872		60	8500 +			0.795	
26	3450	1 / 45	0.022	0.852	0.148	61	8500 +			0.795	
27	3750 +			0.852		62	8750	1 / 9	0.111	0.707	0.293
28	3750 +			0.852		63	8750 +			0.707	
29	4150 +			0.852		64	8750 +			0.707	
30	4150 +			0.852		65	9400 +			0.707	
31	4150 +			0.852		66	9900 +			0.707	
32	4150 +			0.852		67	10100 +			0.707	
33	4300 +			0.852		68	10100 +			0.707	
34	4300 +			0.852		69	10100 +			0.707	
35	4300 +			0.852		70	11500 +			0.707	

第 3 番目は 1150 時間で故障したが、稼働しているのは 68 件となっている。生存率は、直前の生存率が 0.986 なので、その時点での生存率 $1-0.014=0.986$ を掛けて

第3番目 1150時間の生存率： $0.986 \times (1 - 1/68) = 0.971$

と推定するのが妥当である。第4番目も1150時間で故障したので、同様に

第4番目 1150時間の生存率： $0.971 \times (1 - 1/67) = 0.957$

と推定される。順次計算を繰り返して、8750時間での故障時の生存率は、

7850時間の生存率： $0.795 \times (1 - 1/9) = 0.707$

と推測される。最後の70件目は11500時間まで稼働が確認されているが、その時点までの生存率0.707が継続していると推測される。

信頼性工学の分野でも同様に、鈴木ら(2009)によれば、瞬間故障率をハザードとして、それを累積したハザードに着目して、ワイブル分布のパラメータの推定のために「ワイブル型累積ハザード紙」が活用されている。

このように、生存率に着目して故障があった場合に、直前の生存率に、その後の生存率を掛け合わせる方法は、生存率のノンパラメトリックな推定法として知られている Kaplan・マイヤー法に一致している。70件のファンが順次すべて故障したとするならば、69件目での生存率は、

$$\begin{aligned} F(t_{69}) &= 1 \times \left(1 - \frac{1}{70}\right) \times \left(1 - \frac{1}{69}\right) \times \left(1 - \frac{1}{68}\right) \times \cdots \times \left(1 - \frac{1}{3}\right) \times \left(1 - \frac{1}{2}\right) \\ &= 1 \times \frac{69}{70} \times \frac{68}{69} \times \frac{67}{68} \times \cdots \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{70} \end{aligned}$$

となり、常識的な生存率に一致し、70件目が故障すれば、

$$F(t_{70}) = \frac{1}{70} \times \left(1 - \frac{1}{1}\right) = 0$$

と全て故障したことになり、故障が発生した直前の生存率に直後の生存率を掛け合わせて生存率を求める考え方は、途中打ち切りのあるデータに対して統計モデルを使わないノンパラメトリックな方法として標準的に使われている。

ここでは、累積分布関数の下側確率を故障率 $F(t)$ とし、上側確率を生存率 $1 - F(t)$ としたのであるが、信頼性工学では“信頼度”としている。これは、統計での95%信頼区間の信頼とも混同しやすいので、本書では、JMPの表記法に合わせ $F(t)$ を故障率とし、 $S(t) = 1 - F(t)$ を生存率とする。

図 3.8 にノンパラメトリックな方法により推定され故障率と生存率について階段状の曲線（通常 Kaplan・マイヤー曲線といわれている）を示す。

- 1) 生存率は最初の0時間目の生存率を1.0とし、第1番目の故障が起きる450時間まで1.0が続くとし、水平線を引く。

- 2) 故障が起きた 450 時間で生存率が 0.986 まで階段状に下げ、次の故障が起きるまで 0.986 が続くとみなして水平線を引く。
- 3) 打ち切りとなった 460 時間のデータは、“|” などの何らかの記号で、水平線の上に髭を付ける。
- 4) 順次、故障が起きるたびに、階段状に生存率を落とし、水平線を次の故障が起きるまで引く。
- 5) 最後は、最後の故障 8,750 時間より 11,500 時間の打ち切りまで故障率 0.707 が継続するとして水平線を引く。
- 6) 打ち切りがあったすべてのデータを ○ 印としてすべて積み上げて故障率曲線の上側にプロットする。

Kaplan-Meier 曲線を作成するためには、表 3.2 の生存率の計算シートを元に、Excel の散布図を使い、作図に必要なデータを追加すればよい。

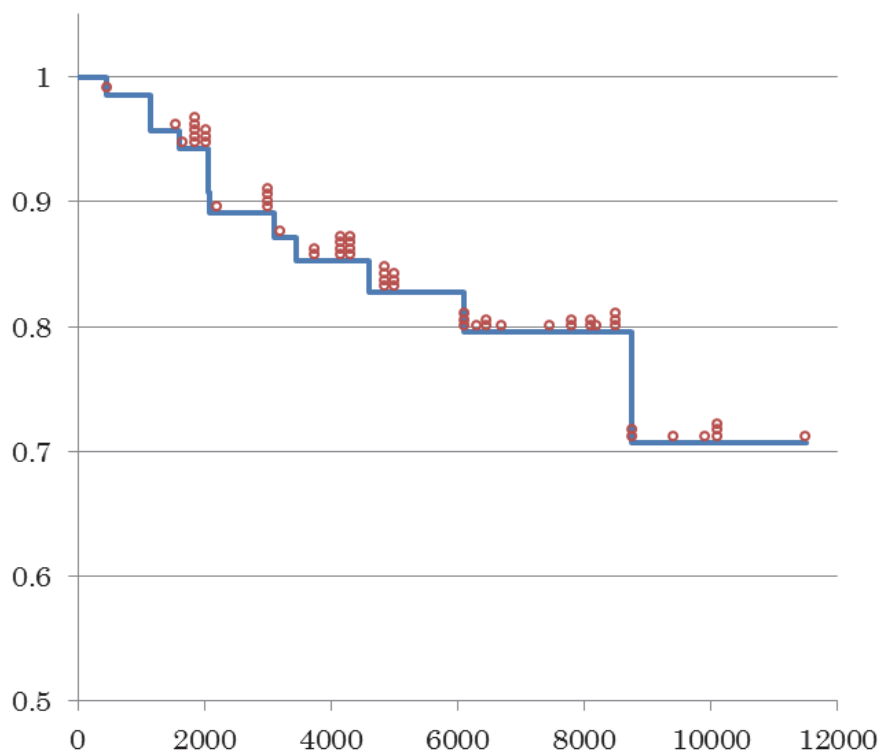


図 3.7 Excel で作成した Kaplan-Meier 曲線

統計ソフト JMP の「生存時間分析」を用いて作図した故障率と生存率の結果を 図 3.8 に示す。きめ細かくアレンジすることはできないが、標準機能として備わっており手軽に作図できる。

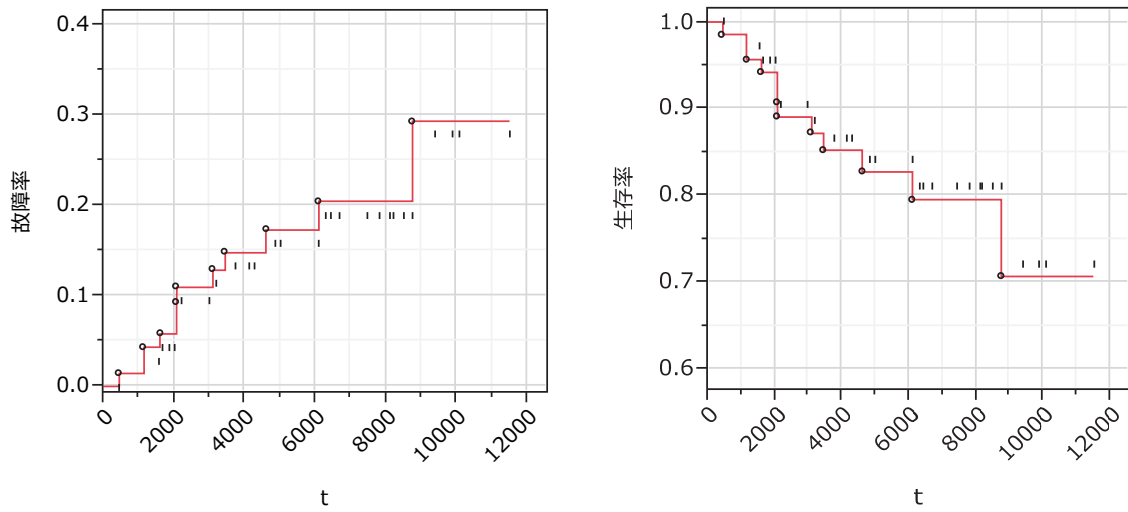


図 3.8 故障率および生存率のカプラン・マイヤー曲線 (JMP 生存時間/信頼性分析)

3.5 打ち切りデータを考慮した最尤法

各時点での故障データ t_i がワイブル分布の確率密度関数 $f(t)$ に従うとして、それを尤度として取り扱ってきた。打ち切りデータについては、その時点まで稼働しているが、その後故障がいつ起きるかはわからない。打切られた時間までは、「生存」していることは明らかであり、ワイブル分布の上側確率、生存関数 $[1-F(t)]$ を尤度とすることが望ましい。

分布の確率密度関数と累積分布関数の両方で尤度を定義してそれらの同時あてはめは、寿命データの統計解析の特徴である。ただし、それぞれのデータは、(故障, 打ち切り) のどちらかであり、それを識別するために、故障した場合を $\delta_i=1$, 打ち切りの場合を $\delta_i=0$ として、同時尤度関数を設定する。(故障, 打ち切り) を考慮した尤度関数は、

$$\begin{aligned}
 L &= \prod_{i=1}^n f(t_i)^{\delta_i} [1-F(t_i)]^{1-\delta_i} \\
 &= \prod_{i=1}^n \left\{ \frac{\beta}{\alpha^\beta} t_i^{\beta-1} \exp \left[-\left(\frac{t_i}{\alpha} \right)^\beta \right] \right\}^{\delta_i} \cdot \left\{ \exp \left[-\left(\frac{t_i}{\alpha} \right)^\beta \right] \right\}^{1-\delta_i}
 \end{aligned} \tag{3.10}$$

となる。故障データの場合は、 $\delta_i=1$ なので、 $f(t)^1=f(t_i)$ および $\{1-F(t_i)\}^{1-1}=1$ から確率密度関数 $f(t_i) \times 1=f(t_i)$ となり、打ち切りデータの場合は、 $\delta_i=0$ なので、 $[f(t_i)]^0=1$ および $[1-F(t_i)]^{(1-0)}=1-F(t_i)$ なので、 $1 \times [1-F(t_i)]=1-F(t_i)$ と生存関数となる。

式 (3.10) から、Weibull α , Weibull β の推定に際し、故障データは確率密度関数用い、打ち切りデータについては生存関数と両方を同時に考慮したパラメータの推定を行うこととなる。実際にファン 12 件の故障データ、58 件の打ち切りデータに対し、Excel のワイブル分布の確率密度関数 $f(t)$, 生存関数 $S(t)=1-F(t)$ を用いた Excel の計算シートを表 3.3 に

示す.

パラメータの初期値として $\hat{\alpha}=20,000$, $\hat{\beta}=1.0$ を与え, 対数尤度 $\ln L$ が最大値になるように Excel のソルバーで $\hat{\alpha}$ と $\hat{\beta}$ を変化させた結果が示されている. 故障データに対しては $\delta_i=0$, 打ち切りデータに対しては $\delta_i=1$ が与えられている.

表 3.3 Excel の Weibull 関数を用いた打ち切りデータを含む最尤法

	Weibull $\alpha =$	26296.9				$\ln L =$	-135.15
	Weibull $\beta =$	1.0584				$-2 \ln L =$	270.31
No	$t + \delta$	$f(t)$	$1-F(t)$	$F(t)$	L_i	$\ln L_i$	
1	450	1	0.000031	0.987	0.013	0.0000313	-10.37
3	1150	1	0.000032	0.964	0.036	0.0000323	-10.34
4	1150	1	0.000032	0.964	0.036	0.0000323	-10.34
6	1600	1	0.000032	0.950	0.050	0.0000325	-10.34
:							
45	6100	1	0.000030	0.808	0.192	0.0000299	-10.42
62	8750	1	0.000028	0.732	0.268	0.0000276	-10.50
2	460 + 0		0.000031	0.986	0.014	0.9863	-0.01
5	1560 + 0		0.000032	0.951	0.049	0.9509	-0.05
:							
69	10100 + 0		0.000026	0.695	0.305	0.6955	-0.36
70	11500 + 0		0.000025	0.659	0.341	0.6592	-0.42
	密度関数 =WEIBULL (C77, \$E\$5, \$E\$4, FALSE)						
	生存関数 =1-WEIBULL (C77, \$E\$5, \$E\$4, TRUE)						
					y_{70}, β, α		

最尤推定量は, $\hat{\alpha}=26,296$, $\hat{\beta}=1.058$ となり, 密度 $f(t_i)$, 生存確率 $1-F(t_i)$, 故障率 $F(t_i)$ が示されている. それぞれのデータに対する尤度 L_i と対数尤度 $\ln L_i$ が示され, $\ln L_i$ のデータ全体に対する和は, $\ln L = -135.15$ となることが示されている.

最初の 450 時間の故障データの確率密度 $f(t_1)=0.000031$ であり, その対数尤度は, $\ln L_1 = -10.37$ で, 故障率は $F(t_1)=0.013$ である. 第 2 番目は 460 時間で打ち切られたデータで, 生存率 $S(t_2)=1-F(t_2)=0.9863$ で, 対数尤度は $\ln L_2 = -0.01$ である. 最後の故障は 62 番目の 8750 時間であり, 故障率は $F(t_{62})=0.268$, 生存率は $1-F(t_{62})=0.732$, と推定されている. なお, 表 3.2 のノンパラメトリックな方法で推定された 62 件目の最後の故障ファンが起きた直後の生存率は 0.707, 故障率は 0.293 であり, ほぼ同様な推定結果となっている.

表 3.3 の最尤推定で求められた $\hat{F}(t_i)$ を $\ln\{\ln[1/(1-\hat{F}(t_i))]\}$ と 2 重対数変換して Y 軸とし X 軸は, 時間 t を対数目盛としてプロットし, それらのプロット点を結ぶと図 3.9 に示すように直線となる. 表 3.2 で示した故障率をメディアン・ランク法で, 調整故障率を

$$\hat{F}'(t_i) = \text{調整故障率}_i = \frac{\text{故障率}_i \times n - 0.3}{n + 0.4}$$

として計算し、 $\hat{F}(t_i)$ と同様に $\ln\{\ln[1/(1-\hat{F}(t_i))]\}$ と 2 重対数変換してプロットした結果である。故障率の代わりに調整した故障率の使用は、すべてが故障した場合に、最後の故障率が 100% となり、確率紙にプロットできなくなるための工夫である。

ワイブル確率紙は、コンピュータによる最尤法によるパラメータ推定が困難である場合に、調整故障率をプロットし、目の子で直線を引くことによりワイブル分布のパラメータを推定する簡便法として使われている。図 3.9(左) には信頼区間の表示があるが、図 3.9(右) には示されていない。Excel による 95% 信頼区間の表示は、図 6.4 を参照のこと。

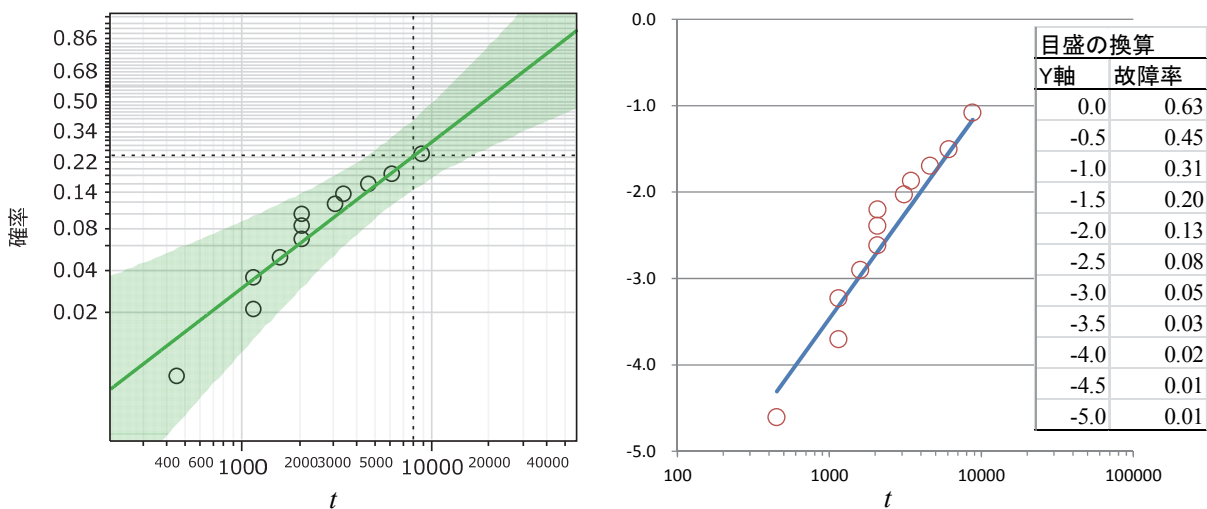


図 3.9 ワイブル確率紙へのプロット (JMP 左), (Excel 右)

最尤法で直接 $\hat{F}(t_i)$ を推定する場合にも、調整故障率が、ワイブル分布に従っているかを視覚的に判断することができるので、統計ソフトの標準的に出力されるようになっている。

表 3.3 からファンの最後の故障は 8,750 時間で、推定された故障率は $\hat{F}(t_{62}) = 0.263$ となる。最後まで観察された打ち切りデータは 11,500 時間での故障率は $\hat{F}(t_{70}) = 0.341$ と推定される。表 3.2 のノンパラメトリックな方法では、最後の故障以後すべて故障率は 0.293 と推測されていた。

推定された Weibull $\hat{\alpha} = 26,296$, Weibull $\hat{\beta} = 1.058$ を用いて 11,500 時間以後のワイブル分布の確率密度関数および累積分布関数 (生存率・故障率) の推定曲線を 図 3.10 に示す。故障した 12 件については、確率密度関数の推定値 $\hat{f}(t_i)$ を○印で上書きし、58 件の打ち切りデータの生存率 $1 - \hat{F}(t_i)$ を○印で上書きしてある。

このように、打ち切りがある場合に最尤法によるワイブル分布のパラメータ推定には、全データの対数尤度の和が最大になるように推定された結果となっている。ワイブル型累積ハザード紙を用いた推定法の場合は、打ち切りデータが累積故障率の算出に反映されるだけであるのに対し、最尤法の場合には、すべての打ち切りデータが、その時間までは故障していないとの生存率が対数尤度としてワイブル分布のパラメータ推定に寄与している。

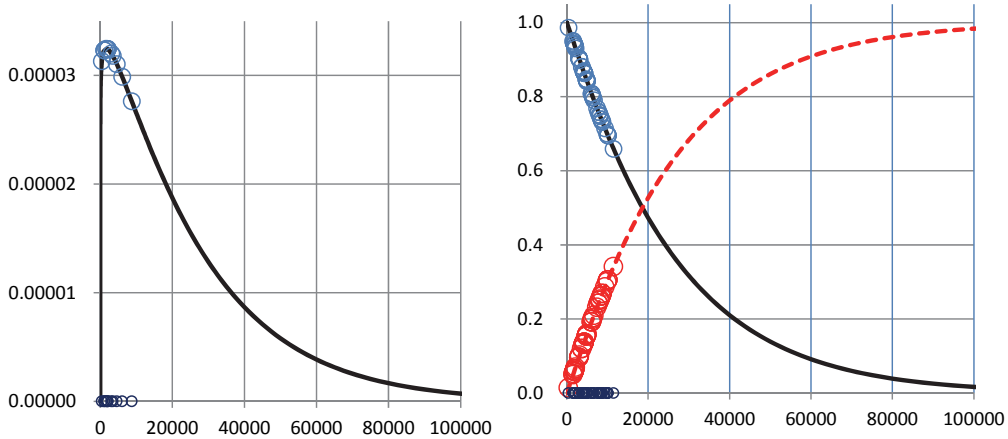


図 3.10 ワイブル分布の確率密度関数と累積分布関数・生存関数の同時あてはめ

ワイブル分布の累積分布関数から、時間を指定した時の故障率は容易に推定できる。逆に、ある故障率となる時間の推定も行いたい。そのためには、ワイブル分布のパーセント点を求める式が必要となる、式 (3.1) を

$$P = 1 - \exp \left\{ - \left(\frac{t}{\hat{\alpha}} \right)^{\hat{\beta}} \right\} \quad (3.11)$$

とにおいて、 t について解いて、次式を得る。

$$t_p = \hat{\alpha} \{ -\ln(1-P) \}^{1/\hat{\beta}} \quad (3.12)$$

このデータから、これまでの保障期間は 8,000 時間での故障率と信頼区間、50%および 90% のファンが故障するまでの時間と信頼区間の算出が要求であった。8,000 時間での故障率は、

$$F(t=8000) = 1 - \exp \left\{ - \left(\frac{t}{\hat{\alpha}} \right)^{\hat{\beta}} \right\} = 1 - \exp \left\{ - \left(\frac{8,000}{26,296} \right)^{1.058} \right\} = 0.247$$

と推定される。50%および 90% のファンが故障する時間は、

$$t_{P=0.5} = \hat{\alpha} \{ -\ln(1-P) \}^{1/\hat{\beta}} = 26,296 \{ -\ln(1-0.5) \}^{1/1.058} = 18,600 \text{ 時間}$$

$$t_{P=0.9} = \hat{\alpha} \{ -\ln(1-P) \}^{1/\hat{\beta}} = 26,296 \{ -\ln(1-0.9) \}^{1/1.058} = 57,825 \text{ 時間}$$

と推定される。さて、これらの推定値の信頼区間は、どのように推定したらよいのであろうか。推定のためには、Weibull $\hat{\alpha}$ 、および Weibull $\hat{\beta}$ の分散と共分散が必要となるが、これまで示してきた手順の中には含まれていない。

表 3.4 時間ごとの故障率と故障までの時間の推定

P	t_P	t	$F(t)=P$
0.100	3137	8,000	0.2471
0.500	18600	80,000	0.9611
0.900	57825	=WEIBULL(\$L32,\$F\$5,\$F\$4,TRUE	
0.950	74147	t, β, α	
0.975	90260		
=\$F\$4 * (-LN(1-\$O32)) ^ (1/\$F\$5)			
α	t_P	β	

3.6 JMPによる解析

これまで、Excelによる解析結果と対比してJMPの「寿命の一変量」および「生存時間／信頼性分析」での図表を示してきた。ここでは、JMPによって出力される結果の使い方を主体にする。

ヘルプメニューからサンプルデータを選択し、信頼性／生存時間を選択すると表 3.5 に示すようにJMPデータリストが表示される。ディーゼル発電機のファンデータは、Fun を選択することにより得られる。

表 3.5 JMPのサンプル・ライブラリ

信頼性分析／生存時間分析	
Appliance	生存時間分析
Comptime	モデルのあてはめ(生存時間(パラメトリック))、Weibull分布および対数正規分布
Devalt	二変量の関係(二変量)、生存時間分析、モデルのあてはめ(生存時間(パラメトリック))
Engine Valve Seat	再生モデルによる分析
Fan	非線形回帰(指数損失、Weibull損失、極値損失)
ICdevice02	生存時間分析、モデルのあてはめ(生存時間分析(パラメトリック))
Locomotive	寿命の一変量
Microprocessor Data	寿命の一変量
Motorette	モデルのあてはめ(生存時間(パラメトリック))
Tobit2	生存時間(パラメトリック)
Blenders	生存時間分析
Rats	生存時間分析、比例ハザード
Bladder Cancer	再生モデルによる分析
VA Lung Cancer	生存時間(パラメトリック)
Small Production	信頼性予測(ネバダ形式)
Small Production part 1	信頼性予測(日付形式)
NewEngineOperation	信頼性成長(Crow-AMSAA)
TurbineEngineDesign1	Reliability Growth (Piecewise Weibull NHPP)

Fan データを 表 3.6 に示す。ノートの記載から出典が Nelson 1982 と記載されており、奥野訳本の原著であることがわかる。▼ 寿命の一変量(Weibull 分布) を選択しスクリプトを実行すると、自動的に解析結果を得ることができる。

表 3.6 ディーゼル発電機のファンデータ

Fan - JMP [2]							
ファイル(F) 編集(E) テーブル(T) 行(R) 列(C) 実験計画(DOE)(D) 分析(A) グラフ(G) ツール(O) 表示(V) ウィンドウ(W) ヘルプ(H)							
Fan							
ロックされたファイル C:\%Progra							
ノート	損失関数のある生存時間分	時間	打ち切りの有無	指数損失関数	Weibull 損失関数	失関数	極値損失関数
	1	450	0	10.2804462	450	1.96405344	
非線形回帰 (指数モデル)	2	460	1	0.01602601	460	0.16922454	
非線形回帰 (Weibullモデル)	3	1150	0	10.3048336	1150	1.28329941	
非線形回帰 (極値モデル)	4	1150	0	10.3048336	1150	1.28329941	
寿命の一変量(指数分布)	5	1560	1	0.05434909	1560	0.57389193	
寿命の一変量(Weibull分布)	6	1600	0	10.3205112	1600	1.11860348	
:							
すべての行	70	66	9900	1	0.34490768	9900	3.64200647
選択されている行	0	67	10100	1	0.35187551	10100	3.71558236
除外されている行	0	68	10100	1	0.35187551	10100	3.71558236
表示しない行	0	69	10100	1	0.35187551	10100	3.71558236
ラベルのついた行	0	70	11500	1	0.40065033	11500	4.23061357

「寿命の一変量」では、図 3.11 に示すような「イベントプロット」が作成できる。Y 軸のラベルは、ディーゼル発電機のファンデータの 1~70 の行番号となっている。故障データは✕と表示され、打ち切りデータは▷で表示されている。このデータでは、右側打ち切りしか示していないが、左側打ち切り、区間打ち切りも表示できる。

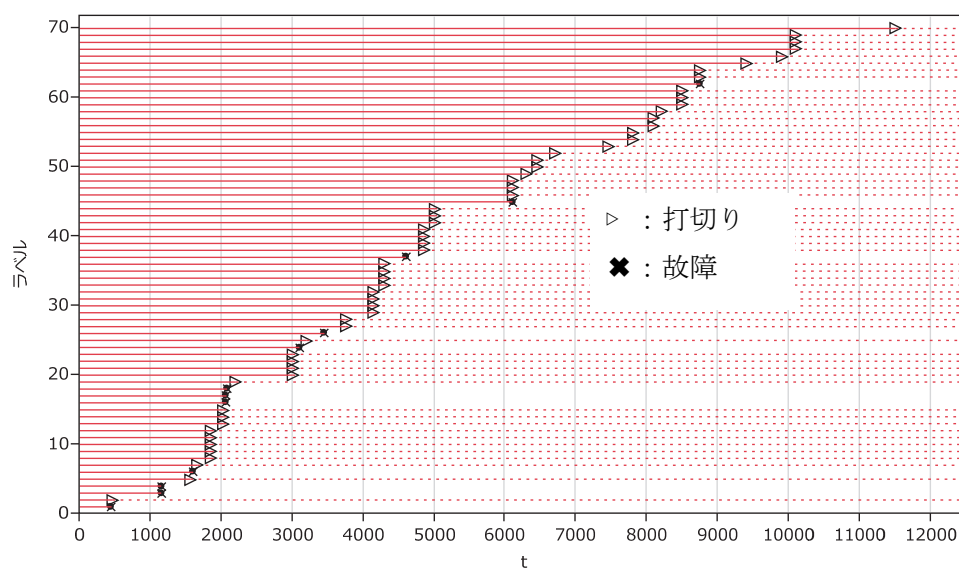


図 3.11 イベントプロット (表 3.2 の 70 件分のデータ)

図 3.12 には、あてはめられたワイブル分布の累積確率分布に 95%信頼区間が上書きされて、確率密度関数も図 3.10 と同様に表示されている。ただし、累積確率分布の散布図上の点は、故障データのみとなっている。ワイブル累積分布には、95%信頼区間も表示され、実際の故障データがない 10,000 時間以後は信頼区間幅の広がりが大きくなっている。

図 3.12 も含めて、標準的な実行結果を JMP のグラフィカル・ユーザ・インターフェイス を用いて見栄えが良くなるように加工した結果となっている。詳細は示さないのが、試行錯誤によって好みの図表に仕上げてもらいたい。

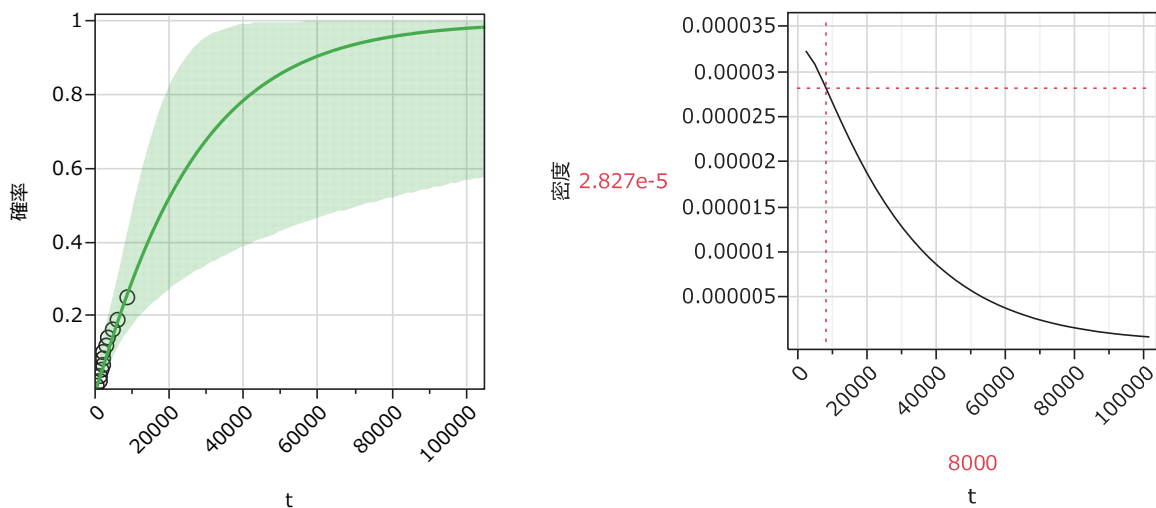


図 3.12 ワイブル累積分布 (左) および確率密度曲線 (右)

図 3.13 には、時間 t が対数目盛の X 軸で、Y 軸は 2 重対数変換後のワイブル確率目盛のワイブル確率プロットが示されている。「寿命の一変量」では、ワイブル分布以外に多くの分

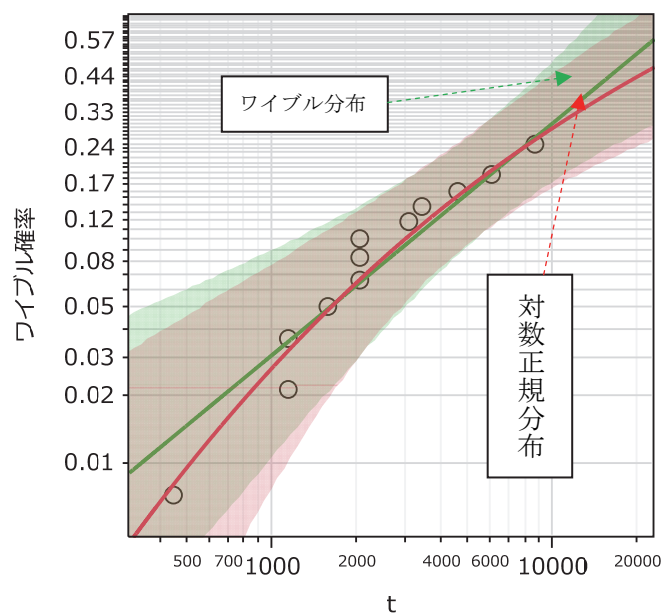


図 3.13 ワイブル確率プロット

布を自由に選択し表示できるようになっている。この図では、対数正規分布のあてはめを上書きしている。Y 軸の目盛を対数正規確率目盛にすると、対数正規分布のあてはめは直線となり、逆にワイブル分布のあてはめは曲線となる。なお、図 3.12 は、目盛を「ノンパラメトリック」を選択した場合である。

表 3.7 はワイブル分布のパラメータの推定値である。表の 1 行目：位置=10.177, 2 行目：尺度=0.945, 3 行目に、Weibull α = 26296.45, 4 行目に Weibull β = 1.058 と Excel で計算した表 3.3 の結果が表示されている。1 行目と 2 行目は、時間 t を対数変換して最小極値分布をあてはめた場合のパラメータの推定値となっている。

表 3.7 ワイブル・パラメータの推定 (Wald 法)

パラメータ	推定値	標準誤差	下側95%	上側95%	規準	
位置	10.1772	0.4659	9.2641	11.0903	(-2)*対数尤度	270.3054
尺度	0.9448	0.2394	0.4755	1.4141	AICc	274.4845
Weibull α	26296.8452	12251.4283	10552.0697	65534.4483	BIC	278.8024
Weibull β	1.0584	0.2683	0.7072	2.1031		

ワイブル分布は、図 3.1 に示したように Weibull β の大きさによって分布の形状の変化が大きいため、各種の推測を行うことが容易ではない。時間 t を対数変換することにより、一山型の左に裾を引く正規分布に近い釣鐘型の形状となり、統計的な取り扱いが容易となる。

データが対数正規分布に従う場合に、対数変換後に正規分布のパラメータ推定を行うことと同様であり、位置=10.1772 の指数をとると $\exp(10.1772) = 26296$ となり、一致する、対数正規分の「平均値」にたいする「幾何平均」との関係と同様である。尺度=0.945 は、標準偏差 σ に対応していて、尺度 $\sigma = 1/\beta$ の関係となっている。

最小極値分布については、第 6 章で導入し、詳しく述べるが、ここでは、最小極値分布の 95%信頼区間とワイブル分布のパラメータの 95%信頼区間の関係を簡単に説明する。位置=10.177 の 95%信頼区間は、(9.264, 11.090) であり指数をとると (10552, 65534) と Weibull α の 95%信頼区間が計算される。尺度=0.945 の 95%信頼区間 (0.475, 1.414) の逆数を取り上下を入れ替えると (0.707, 2.103) Weibull β の 95%信頼区間が計算される。

これまで、ワイブル分布をあてはめた場合にパラメータの標準誤差を Excel による計算をあえて示してこなかった。これは、最尤法によってパラメータを推定した場合に、それらの標準誤差の計算のためには対数尤度の式に対して 2 階の偏微分の計算が必要となり、煩雑な説明になるために避けてきた。

表 3.8 に示すように、「寿命の一変量」でパラメータの分散共分散行列を得ることができる。この行列の対角要素の平方根が、それぞれのパラメータの標準誤差となっている。位置の標準誤差は、 $\sqrt{0.217053} = 0.4645$ と計算できる。位置と尺度の欄の 0.090442 は共分散であり、図 3.12 および 図 3.13 で示されているワイブル分布 95%信頼区間の計算に用いられる。詳細は、第 6 章で取り上げる。

表 3.8 パラメータの分散共分散行列

パラメータ	位置	尺度
位置	0.217053	0.090442
尺度	0.090442	0.057333

「寿命の一変量」では、寿命時間 t を指定した場合の 95%信頼区間を計算することができる。図 3.14 (左) は、時間 $t = 8000$ とした場合に下側確率 0.247、95%信頼区間が (0.145, 0.399) が Y 軸に表示されている。半分が故障する時間は、図 3.14 (右) で 18600 時間、95%信頼区間が (8524, 40584) であることが示されている。

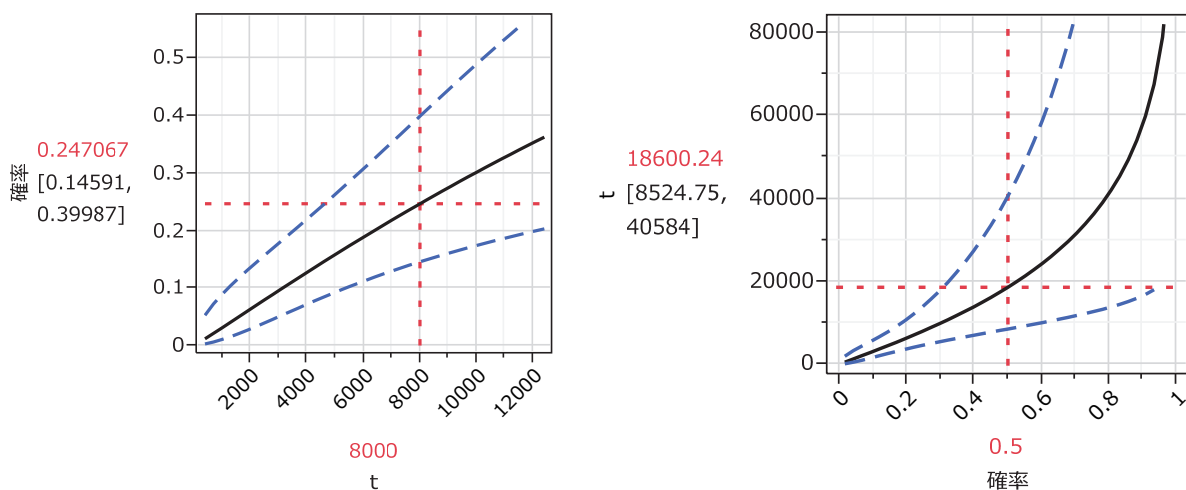


図 3.14 95%信頼区間の推定

表 3.9 には、複数の指定した時間対して、推定値とその 95%信頼区間が表示されている。信頼区間の計算法が 2 通り示されている。(Wald) とあるのは、分散共分散から計算された近

表 3.9 時間 t に関する推定値と 95%信頼区間

t	故障確率	故障確率 下側95%	故障確率 上側95%	故障確率		生存確率
		(Wald)	(Wald)	下側95% (尤度)	上側95% (尤度)	
8000	0.2471	0.1459	0.3999	0.1386	0.3859	0.7529
80000	0.9611	0.5221	1.0000	0.5646	1.0000	0.0389

似計算結果であり、(尤度) とあるのは、正確な信頼区間となっている。(尤度) に基づいた信頼区間は、第 11 章で詳細に示すが、JMP 以外の統計ソフトでは対応しているのを見たことはないので、報告書などでは Wald の近似信頼区間を使用することを勧める。

表 3.10 は、故障率を指定して、故障時間(寿命分位点)と Wald 法とプロファイル尤度法による 95%信頼区間の計算結果である。

表 3.10 故障確率 p に関する推定値と 95%信頼区間

故障確率	寿命分位点	寿命分位点 下側95% (Wald)	寿命分位点 上側95% (Wald)	寿命分位点 下側95% (尤度)	寿命分位点 上側95% (尤度)
0.100	3137	1686	5837	1420	5662
0.500	18600	8525	40584	10506	60768
0.900	57825	16541	202156	23652	392499
0.950	74147	18949	290138	27977	597301
0.975	90260	21074	386591	31913	833880

「生存時間／信頼性分析」では、いわゆる Kaplan-Meier 曲線の作成に優れている。Kaplan-Meier 曲線は、図 3.8 で示しているが、図 3.15 で打ち切りは \circ 印とし、95% 信頼区間を上書きし、さらに内側を塗っている。

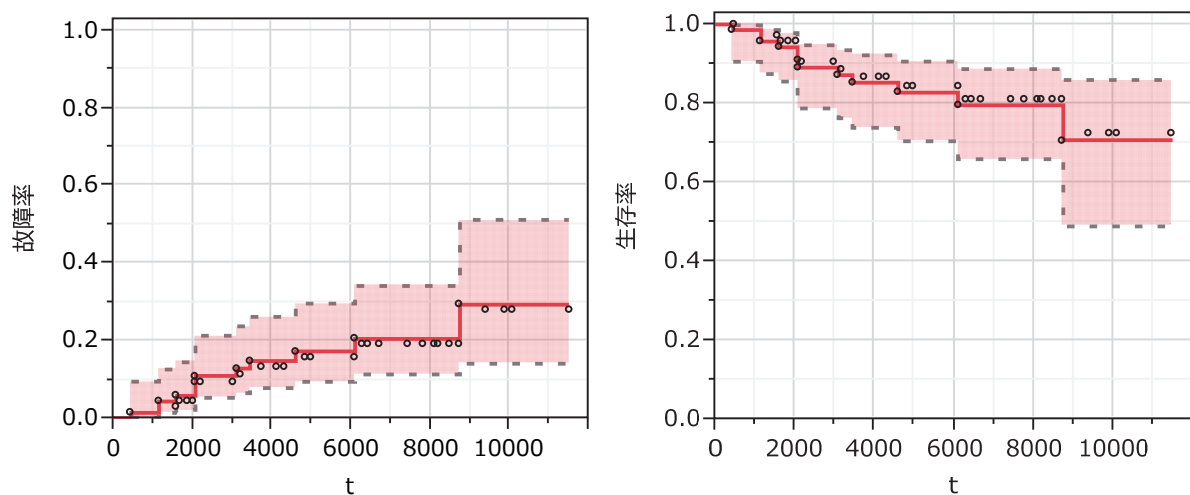


図 3.15 Kaplan-Meier 曲線に対する 95%信頼区間

「寿命の一変量」が新たに追加される前は、この「生存時間／信頼性分析」でワイブル分布のあてはめを行ってきた。「寿命の一変量」の豊富な解析機能に比べると見劣りするが、ワイブル分布のパラメータの表記が、JMP 内でも異なっていることを例示する。

パラメータの推定値は、「寿命の一変量」と一致するが、95%信頼区間の結果は、明らかに異なる。Weibull α の場合では、(10552, 65534) に対して (13631, 106087) となっている。

これは、信頼区間の求め方の違いであり、「生存時間／信頼性分析」では、プロファイル尤度を用いた信頼区間を標準としているのに対し、「寿命の一変量」では、信頼区間の計算にパラメータの標準誤差を用いた近似計算（Wald 法）が使われているためである。「寿命の一変量」で、信頼区間の計算方法が（Wald, 尤度）選択できるようになっている。「尤度」を選択すれば 95%信頼区間は一致する。

表 3.11 最小極値およびワイブル・パラメータの推定（尤度法）

	パラメータ	推定値	下側95%	上側95%	故障数
位置 μ	λ	10.177204	9.5201194	11.572014	12
尺度 σ	δ	0.9447814	0.6031578	1.6502353	12
$\alpha = \exp(\lambda)$, $\beta = 1/\delta$ のときの極値に等しい					
	パラメータ	推定値	下側95%	上側95%	故障数
	α	26296.845	13631.239	106086.92	12
	β	1.0584459	0.6059742	1.6579409	12

信頼区間は尤度を使った方法，第 11 章を参照のこと。

「生存時間／信頼性分析」でも故障確率 p に関する推定および時間 t に関する推定が行えるようになっている。故障確率 p に関する推定は、「寿命の一変量」の Wald での結果と一致するが、時間 t に関する推定では、8000 時間の場合に少数点以下 3 桁で異なるが、原因は特定できない。

各種の推定値の 95%信頼区間の計算方法は、書物によっても異なる。筆者は、奥野監訳本に基づき Excel による数値計算をしてきた。しかしながら、信頼区間の計算が異なり、JMP とは Weibull β の場合に一致しない。ワイブル分布の最尤法によるあてはめは、第 8 章「多重打ち切りデータの最尤法によるあてはめ」の第 8.3 節に述べられている。そこで、「図 8.3.2 ファンの故障データに関する統計ソフト WEIB の出力結果:p310」の結果を表 3.12 に引用し、JMP の結果と対比した結果、不一致の場合の原因を示した。

表 3.12 奥野監訳本のワイブル・パラメータ

	推定値	下側 95%	上側 95%	信頼区間
Weibull α	26297	10552	65535	Wald で一致
Weibull β	1.0584	0.6448	1.7394	不一致
50%点	18600	8524.8	40584	Wald で一致
$t=8000$ 生存率	0.7529	0.60817	0.85681	奥野監訳本の数値
故障率	0.2471	0.14319	0.39183	少数点 3 ケタで不一致

Weibull β の不一致は、Weibull α と同様に対数の計算式を用いたためである。8000 時間の不一致は、ワイブル分布の分散共分散を極値分布の分散共分散としてから JMP と同様の計算を行ったための計算精度が落ちたためであろう。

4. 最尤法による加速試験データの統計解析

4.1 加速試験データの統計解析の歴史的背景

これまでは、1 群の寿命データの場合を扱ってきた。データが正規分布に従い、打ち切りがない完全データだけの場合であれば、統計解析の基礎の問題と思われるかもしれない。データから推定された平均 $\hat{\mu}$ と不偏分散 $\hat{\sigma}^2$ から、平均値の分散を $\hat{\sigma}^2/n$ として平均値の 95% 信頼区間を算出することは容易である。

データがワイブル分布に従うと仮定すると、Weibull パラメータの推定すら容易ではない。さらに、打ち切りデータを含めて Weibull パラメータをどのように推定したらよいのだろうか。これらの課題に対して、これまで、理論的な背景を示しつつ Excel での計算結果を示し、段階的な学習の支援することを念頭に述べてきた。

筆者も必要に迫られ、がんの臨床試験の生存時間データの解析を行ってきたが、頼りとなる参考書も手元にない時代から、統計ソフトを暗中模索しつつ使ってその場をしのいで経験を積み重ねてきたのであるが、しょせん 2 群の比較程度の問題であった。

製品寿命が長く、限られた時間内では実際の試験データが得られない場合に、過酷な環境下での試験結果から日常的な環境下での様々な製品寿命に関連する推定とその 95% 信頼区間を含む予測をどのように行ったらよいのであろうか。

統計ソフト JMP の最新のバージョンで、寿命データの統計解析の機能が拡充されたことを知った。しかしながら、操作マニュアルを紐解いても、茫洋とし全体像が把できにくい。そこで、理論的な背景を精査し、本質の理解のための Excel による「手計算」を行い、JMP による解析結果と対比しつつ、加速試験データの統計解析について、詳細な解説書を作成することにした。

表 2.2 で示したデバイス A の高温下での寿命試験では、この製品が使われる温度が 10°C での 5000 時間では、30 件のデバイス A はすべて生存しており、これからでは寿命に関する推定は、単に「5000 時間で故障するデバイス A は 30 件中 0 件であった」と表現とするだろうか。製品寿命の予測に高温下での試験結果をどのように反映したらよいのであろうか。

奥野忠一監訳 (1988)、「寿命データの解析」の第 13.10 節は、加速試験についての文献の紹介が約 1 ページ分あり、1980 年代は専門家の課題であったことが推測される。筆者も臨床試験の生存時間データの解析を始めた際に、参考書としてこの著書を購入したが、読み応え

のある本であり，途中で挫折してしまった．以下に引用する．

13.10 加速寿命試験

製品や材料の加速寿命試験は，その寿命分布に関する情報を速く得るために用いられる．そのような試験では，試験ユニットは標準よりも厳しい条件が課される．これにより，標準状態のもとで観測されるよりも寿命が短くなる．加速試験条件は，典型的には，高水準の温度，電圧，圧力，振動，サイクル率，荷重などやこれらの組合せにおいて，ユニットを試験することで得られる．ある定まった加速変数や負荷変数を用いることは，多くの製品や材料に対して，技術的に十分確立されている．他の分野では，寿命と寿命に影響する変数の関係を推定する問題が起こる．

厳しい条件や加速条件において得られたデータは，適切なモデルを用いて標準の条件へ外挿し，標準状態のもとでの寿命分布の推定値を求める．そのような試験によって，標準状態における試験に比べ，時間と費用が節減できる．実際，多くの製品や材料については，標準状態における寿命が長すぎて，このような状態で試験することは問題外である．

Nelson (1974a), General Electric (1975, 11 節), Little and Jebe (1975), Yurkowsky ほか(1967) は，ユニットに高負荷を課した加速試験の計画と解析の方法を概観している．とくに，すべての試験ユニットが故障する前にそのようなデータを解析する方法に関する文献が与えられている．この重要な進歩により，(1) すべてのユニットが故障する前に試験を終了することができ，時間と費用の節約になり，(2) 低負荷において試験することができ，外挿を減らすことができる．また，文献には，種々の故障モードが混在するデータの適切な解析法も示されている．以前は，故障モードの混在するデータを用いて，設計条件における寿命分布を推定する方法は知られていなかった．Meeker (1979) は，統計的方法と応用を含め，加速試験に関する文献を与えている．

種々の統計手法と専門的な問題は，Nelson(1971, 1973, 1974a, 1974b, 1975), Nelson and Hahn (1972), Hahn and Nelson (1971) にみられる．<<文献は省略>>

統計ソフト JMP の日本語訳された 440 ページにもなるマニュアル「品質管理および信頼性 / 生存期間」で加速試験データの解析で Mekker ら (1998) の事例が用いられており，そこには，古い時代の加速試験データの統計解析に対する多くの要望が解決されていて，誰でも手軽にデータ解析を行えるようになっていたことにショックを受けた．それとは対照的に，日本での加速試験データの統計解析の状況を WEB で検索して調べた結果，ワイブル確率紙をベースにした古い時代のままであると認識した．

JMP のマニュアルは，操作法が主体であり，加速試験データの統計解析の理論的背景などは，読み取ることができない．そこで，Excel を用いて JMP が行っている加速試験データの統計解析を追試しつつ，理論的な背景を浮き彫りにすることにした．

加速試験データには，表 2.2 で示したように，通常の 10℃ の使用下でのデバイス A が，

あらかじめ設定された 5,000 時間になっても、故障が全くないとの結果であった。このようなデータであっても、その時間まで故障はしていないとの情報を持っている。どのようにこの情報を活用するかが、加速試験データの統計解析の課題である。

4.2 多群の寿命データの統計解析

奥野監訳本の第 12 章は、最尤法による比較（多重打ち切りデータ、その他のデータ）であり、2 群および多群のデータの群間比較が論じられている。ここで例示されているのは、群ごとに寿命分布をあてはめてそれぞれの群のパラメータを比較する方法に留まっており、加速試験の目的である通常状態での寿命の予測の問題には触れられていない。

表 2.2 のデバイス A の寿命データについて、ノンパラメトリックな方法での故障率の推定結果を表 4.1 に示す。群 1 の 10°C では 30 件中故障したものはないので、故障率は 0.0 である。群 2 の 40°C では、最初の故障は 1,298 時間であり、100 件中 1 件なので瞬間故障率は $1/100=0.01$ であり、2 番目の瞬間故障率は、1,390 時間であり 99 件中 1 件で $1/99=0.0101$ となり、生存率は、 $(1-1/100)\times(1-1/99)=0.980$ で、故障率は $1-0.980=0.020$ となる。途中打ち切りがないので、100 件中 2 件の故障とした結果と一致する。観察終了の 5,000 時間での故障率は、0.10 である。同様に第 3 群では、0.450、第 4 群では 0.933 となっている。

表 4.1 デバイス A のノンパラメトリックな故障率の推定

群	温度	t	件数	逆順	打切	δ	故障	瞬間 故障率	生存率	故障率
1	10	5000	30	30	1	0	0 / 30	0.0000	1.000	0.000
2	40	1298	1	100	0	1	1 / 100	0.0100	0.990	0.010
2	40	1390	1	99	0	1	1 / 99	0.0101	0.980	0.020
:										
2	40	4982	1	91	0	1	1 / 91	0.0110	0.900	0.100
2	40	5000	90	90	1	0	0 / 90	0.0000	0.900	0.100
3	60	581	1	20	0	1	1 / 20	0.0500	0.950	0.050
3	60	925	1	19	0	1	1 / 19	0.0526	0.900	0.100
:										
3	60	4674	1	12	0	1	1 / 12	0.0833	0.550	0.450
3	60	5000	11	11	1	0	0 / 11	0.0000	0.550	0.450
4	80	283	1	15	0	1	1 / 15	0.0667	0.933	0.067
4	80	361	1	14	0	1	1 / 14	0.0714	0.867	0.133
:										
4	80	2884	1	2	0	1	1 / 2	0.5000	0.067	0.933
4	80	5000	1	1	1	0	0 / 1	0.0000	0.067	0.933

観察打ち切りを無視して、温度を X 軸とし、Y 軸を対数故障時間とした散布図を 図 4.1 に示す。知りたいのは、10°Cにおいて、故障が 10%となる時間の推定である。散布図中の直線は、表 4.1 の故障率を元に、各温度でおよそ故障率が 0.10 となる点を通る直線をフリーハンドで引いたものである。この直線が 10°Cを通る時間は、およそ 40,000 時間である。

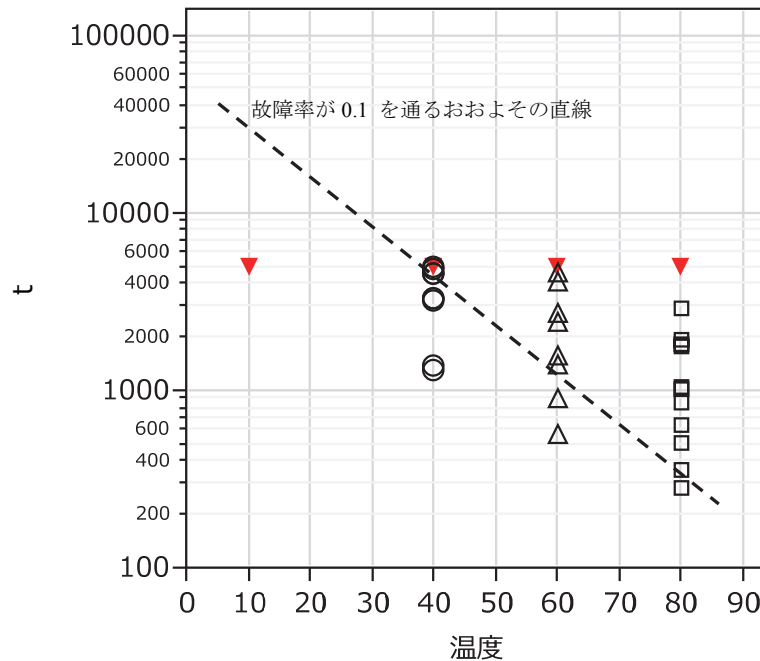


図 4.1 故障率が 10%と想定されるおよその直線 (JMP 二変量の関係)

4.3 群ごとに異なるワイブル分布のあてはめ

Excel を用いて、40 度、60 度、80 度の 3 群についてそれぞれの群ごとに Weibull パラメータ α と Weibull パラメータ β を推定する。ただし、別々のシートではなく 1 枚のシートで同時に推定してみる、これは、形状パラメータを 3 群で共通とした場合にそれぞれの Weibull パラメータの推定に拡張するためである。

ワイブル分布の確率密度関数と累積分布関数は、1 群の場合には、式 (3.1) および式(3.2) に示したように

$$F(y) = 1 - \exp\left\{-\left(\frac{y}{\alpha}\right)^\beta\right\}, \quad f(y) = \frac{\beta}{\alpha^\beta} y^{\beta-1} \exp\left\{-\left(\frac{y}{\alpha}\right)^\beta\right\} \quad (4.1)$$

であった。4 群の場合のパラメータを $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$ とする。これらのパラメータを同時推定するために、新たに変数 $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ を次のように用意する。

10 度の場合 : $x_1 = 1$, それ以外 : $x_1 = 0$

40 度の場合： $x_2 = 1$ ， それ以外： $x_2 = 0$

60 度の場合： $x_3 = 1$ ， それ以外： $x_3 = 0$

80 度の場合： $x_4 = 1$ ， それ以外： $x_4 = 0$

この変数 \mathbf{x} と Weibull α との要素の積和で，

$$\alpha = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4, \quad \alpha = \mathbf{x}\alpha^T \quad (4.2)$$

とし， Weibull β との要素も積和で，

$$\beta = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4, \quad \beta = \mathbf{x}\beta^T \quad (4.3)$$

とする．表 4.2 に示すように，それぞれのデータごとに変数 \mathbf{x} を群ごとに新たに付け加え，さらに積和の結果も付け加える．その結果として，群ごとにそれぞれ異なる Weibull α と Weibull β が，データ 1 行ごとに設定される．

このような計算シートを作成することにより，ワイブル分布の確率密度関数と累積分布関数が，各行ごとに計算される．行ごとに計算された対数尤度 $\ln L_i$ を全て足し合わせた $\ln L = \ln L_1 + \ln L_2 + \dots + \ln L_n$ を計算し， $\ln L$ が最大になるように $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ と $(\beta_1, \beta_2, \beta_3, \beta_4)$ を逐次的に変化させる．もちろん，この探索には Excel のソルバーを使う．対数尤度 $\ln L_i$ は，

表 4.2 群ごとに異なる α_j と β_j を用いたあてはめ構造

					α_1	α_2	α_3	α_4	$\alpha = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4$				
					β_1	β_2	β_3	β_4	$\beta = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$				
群					1	2	3	4	$\ln(L_i) = f(t_i; \alpha, \beta)^{\delta_i} \cdot F(t_i; \alpha, \beta)^{1-\delta_i}$				
					変数 \mathbf{x}				ワイブル分布				
群	温度	t	件数	δ	x_1	x_2	x_3	x_4	α	β	$f(t; \alpha, \beta)$	$1-F(t)$	対数尤度 $\ln L_i$
1	10	5000	30	0	1	0	0	0	α_1	β_1	$f(t_1)$	$1-F(t_1)$	$\ln L_1 = 1-F(t_1)$
2	40	1298	1	1	0	1	0	0	α_2	β_2	$f(t_2)$	$1-F(t_2)$	$\ln L_2 = f(t_2)$
2	40	1390	1	1	0	1	0	0	α_2	β_2	$f(t_3)$	$1-F(t_3)$	$\ln L_3 = f(t_3)$
		:							:	:	:	:	:
2	40	4982	1	1	0	1	0	0	α_2	β_2			
2	40	5000	90	0	0	1	0	0	α_2	β_2			
3	60	581	1	1	0	0	1	0	α_3	β_3			
3	60	925	1	1	0	0	1	0	α_3	β_3			
		:							:	:			
3	60	4674	1	1	0	0	1	0	α_3	β_3			
3	60	5000	11	0	0	0	1	0	α_3	β_3			
4	80	283	1	1	0	0	0	1	α_4	β_4			
4	80	361	1	1	0	0	0	1	α_4	β_4			
		:							:	:			
4	80	2884	1	1	0	0	0	1	α_4	β_4			
4	80	5000	1	0	0	0	0	1	α_4	β_4	$f(t_n)$	$1-F(t_n)$	$\ln L_n = 1-F(t_n)$
合計=												$\ln L$	

$$\ln(L_i) = f(t_i; \alpha, \beta)^{\delta_i} \cdot F(t_i; \alpha, \beta)^{1-\delta_i} \quad (4.4)$$

ただし、 $\delta = 0$ の場合： $\ln L_i = f(t_i)$ ，
 $\delta = 1$ の場合； $\ln L_i = 1 - F(t_i)$

となる。

表 4.3 に群ごとに異なる α_j と β_j の推定のための Excel シートを示す。第 1 群は、故障データがないので、第 2 群から第 4 群について、パラメータの推定結果が示されている。第 2 群の $t_{1000} = 1.298$ の場合であれば Weibull α と Weibull β は、

$$\alpha = [1 \ 0 \ 0] \begin{bmatrix} 13.718 \\ 7.406 \\ 2.233 \end{bmatrix} = 13.717, \quad \beta = [1 \ 0 \ 0] \begin{bmatrix} 2.233 \\ 1.249 \\ 1.312 \end{bmatrix} = 2.233 \quad (4.5)$$

として計算され、確率密度関数と生存関数は、Excel の WEIBULL () 関数を用いて

$$f(1.298) = \text{WEIBULL}(1.298, 2.233, 13.717, \text{FALSE}) = 0.009$$

$$1 - F(1.298) = \text{WEIBULL}(1.298, 2.233, 13.717, \text{TRUE}) = 0.995$$

が計算されている。尤度 L_i は、 $\delta = 1$ なので、 $f(1.298) = 0.009$ となり、その対数は、 $\ln L_i = -4.727$ となる。 $\ln L$ は、 $\ln L_i$ の合計で、

$$\begin{aligned} \ln L &= \ln L_2 + \ln L_3 + \dots + \ln L_{37} \\ &= -4.727 - 6.643 - \dots - 3.994 = -94.387 \end{aligned} \quad (4.6)$$

となる。

表 4.3 群ごとに異なる α_j と β_j の推定

		$\alpha = \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4$				$\beta = \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$				α	-	13.717	7.406	1.740	$\ln L =$	-94.387
										β	-	2.233	1.249	1.312		
										群	1	2	3	4		
群	温度	t_{1000}	件数	δ	生存率	変数 x				パラメータ		ワイブル分布				
						x_1	x_2	x_3	x_4	α	β	$f(y)$	$1-F(t)$	L_i	$\ln L_i$	
1	10	5.000	30	0	1.000	1	0	0	0	-	-	-	-	-	-	
2	40	1.298	1	1	0.990	0	1	0	0	13.717	2.233	0.009	0.995	0.009	-4.727	
2	40	1.390	1	1	0.980	0	1	0	0	13.717	2.233	0.010	0.994	0.010	-4.643	
:																
2	40	4.982	1	1	0.900	0	1	0	0	13.717	2.233	0.042	0.901	0.042	-3.168	
2	40	5.000	90	0	0.900	0	1	0	0	13.717	2.233	0.042	0.900	0.900	-9.457	
3	60	0.581	1	1	0.950	0	0	1	0	7.406	1.249	0.086	0.959	0.086	-2.455	
3	60	0.925	1	1	0.900	0	0	1	0	7.406	1.249	0.093	0.928	0.093	-2.372	
:																
3	60	4.674	1	1	0.550	0	0	1	0	7.406	1.249	0.086	0.570	0.086	-2.457	
3	60	5.000	11	0	0.550	0	0	1	0	7.406	1.249	0.083	0.542	0.542	-6.735	
4	80	0.283	1	1	0.933	0	0	0	1	1.740	1.312	0.390	0.912	0.390	-0.941	
4	80	0.361	1	1	0.867	0	0	0	1	1.740	1.312	0.406	0.881	0.406	-0.900	
:																
4	80	2.884	1	1	0.067	0	0	0	1	1.740	1.312	0.127	0.144	0.127	-2.065	
4	80	5.000	1	0	0.067	0	0	0	1	1.740	1.312	0.019	0.018	0.018	-3.994	

初期値として用いた Weibull α は、図 2.6 のワイブルプロットから、それぞれの群の直線と Y 軸の下側確率 0.632 となる交点から、(14, 7, 2) とし、Weibull β は、故障率曲線から、遅発型の故障と推測されるので、(2, 2, 2) とした。対数尤度は $\ln L = -99.045$ である。

α	-	14.000	7.000	2.000	$\ln L =$	-99.045
β	-	2.000	2.000	2.000		
群	1	2	3	4		

この対数尤度 $\ln L = -99.045$ を最大にするように、(14, 7, 2), (2, 2, 2) を Excel のソルバーで変化させると、次のように

α	-	13.717	7.406	1.740	$\ln L =$	-94.387
β	-	2.233	1.249	1.312		
群	1	2	3	4		

Weibull α と Weibull β が最尤推定量として得られる。もちろん対数尤度も大きくなっている。

Excel で求めたパラメータを用いてワイブル確率紙に、メディアン・ランクで調整した生存率をプロットし、その上にワイブル分布のあてはめ直線を引いた結果を 図 4.2 (左) に示す。更に、温度ごとの散布図にワイブル分布の確率密度関数を上書きした結果 (右) を示す。

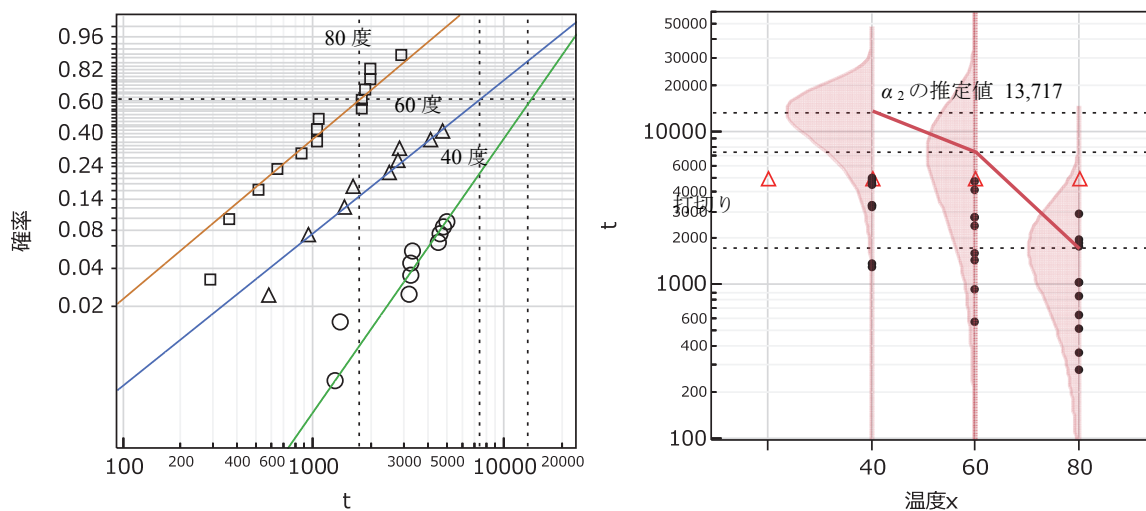


図 4.2 群ごとに α_j と β_j のあてはめ (JMP 寿命の二変量)

推定された Weibull α は ($\hat{\alpha}_2 = 13.717$, $\hat{\alpha}_3 = 7.406$, $\hat{\alpha}_4 = 1.740$) であり 1000 倍すると (13,717, 7,406, 1,740) となる。これらは、ワイブル分布の 63.2 パーセント点なので、ワイブル確率プロットの確率が 0.632 に水平線を引いたときに、最尤法で求めた各温度の直線との交点がそれぞれ 13,717, 7,406, 1,740 となる。

温度を X 軸にした散布図上に、40°C ($\hat{\alpha}_2 = 13,717$, $\hat{\beta}_2 = 2.233$), 60°C ($\hat{\alpha}_3 = 7,406$, $\hat{\beta}_3 = 1.249$), 80°C ($\hat{\alpha}_4 = 1,740$, $\hat{\beta}_4 = 1.312$) のワイブル分布の密度曲線が示されている。それぞれのワイブル分布 0.632 パーセント点を $\hat{\alpha}_2$, $\hat{\alpha}_3$, $\hat{\alpha}_4$ を直線で結んだ折れ線が示されている。

80 度では、分布の 63.2%点は $\hat{\alpha}_4 = 1,740$ 時間 と観測された故障データ内であるが、60 度の場合は $\hat{\alpha}_3 = 7,406$ 時間 と打ち切りの 5000 時間以上となり、40 度の場合は $\hat{\alpha}_2 = 13,717$ 時間 と打ち切りの 5000 時間以上をはるかに超えていることが確認される。形状 $\hat{\beta}_j$ は、それぞれ、1.312, 1.249, 2.233 であり、遅発型の故障となっている。

4.5 Weibull β を共通とするワイブル分布のあてはめ

それぞれの温度の形状を共通 $\beta = \beta_2 = \beta_3 = \beta_4$ とし、温度ごとに α_i が異なるとした場合には、Excel 上に共通パラメータ β を設定し、それぞれの行で共通 β 単に代入するように変更する。この方法より、各行の尤度の計算式は変更しないですむ。パラメータ α_i の初期値としては、推定された (13,717, 7,406, 1,740) をそのまま使い、形状パラメータは、おおよその平均値 1.6 とすると、 $\ln L = -98.598$ が計算される。

α		13.717	7.406	1.780	$\ln L_i =$	-98.559
β	-	1.600	共通			
群	1	2	3	4		

ソルバーで $\ln L$ を最大となるように α_j と β を Excel のソルバーで変化させると、

α		24.421	6.942	1.780	$\ln L_i =$	-95.463
β	-	1.427	共通			
群	1	2	3	4		

$\ln L = -95.463$ が得られる。これが最大値であることを確認するためには、推定された各パラメータを前後に振って $\ln L = -95.463$ より大きな尤度とならないことで確認できる。

表 4.4 には、計算シートを示す。形状が共通としたので、図 4.3 に示すように、ワイブル分布のあてはめ直線は平行となり、確率密度関数も位置は異なるが同じ形状のものが各群に上書きされている。

表 4.4 群ごとに異なる α_j と共通の β の推定

		$\alpha = \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4$				α				$\beta = \text{共通}$			$\ln L = -95.463$		
						β									
						群									
										パラメータ			ワイブル分布		
群	温度	t_{1000}	件数	δ	生存率	x_1	x_2	x_3	x_4	α	β	$f(t)$	$1-F(t)$	Li	$\ln L_i$
1	10	5.000	30	0	1.000	1	0	0	0	-	-	-	-	-	-
2	40	1.298	1	1	0.990	0	1	0	0	24.420	1.427	0.016	0.985	0.016	-4.108
2	40	1.390	1	1	0.980	0	1	0	0	24.420	1.427	0.017	0.983	0.017	-4.081
:															
2	40	4.982	1	1	0.900	0	1	0	0	24.420	1.427	0.027	0.902	0.027	-3.622
2	40	5.000	90	0	0.900	0	1	0	0	24.420	1.427	0.027	0.901	0.901	-9.359
3	60	0.581	1	1	0.950	0	0	1	0	6.942	1.427	0.069	0.971	0.069	-2.671
3	60	0.925	1	1	0.900	0	0	1	0	6.942	1.427	0.082	0.945	0.082	-2.499
:															
3	60	4.674	1	1	0.550	0	0	1	0	6.942	1.427	0.098	0.566	0.098	-2.320
3	60	5.000	11	0	0.550	0	0	1	0	6.942	1.427	0.096	0.535	0.535	-6.886
4	80	0.283	1	1	0.933	0	0	0	1	1.780	1.427	0.340	0.930	0.340	-1.079
4	80	0.361	1	1	0.867	0	0	0	1	1.780	1.427	0.366	0.903	0.366	-1.005
:															
4	80	2.884	1	1	0.067	0	0	0	1	1.780	1.427	0.135	0.137	0.135	-2.006
4	80	5.000	1	0	0.067	0	0	0	1	1.780	1.427	0.016	0.013	0.013	-4.365

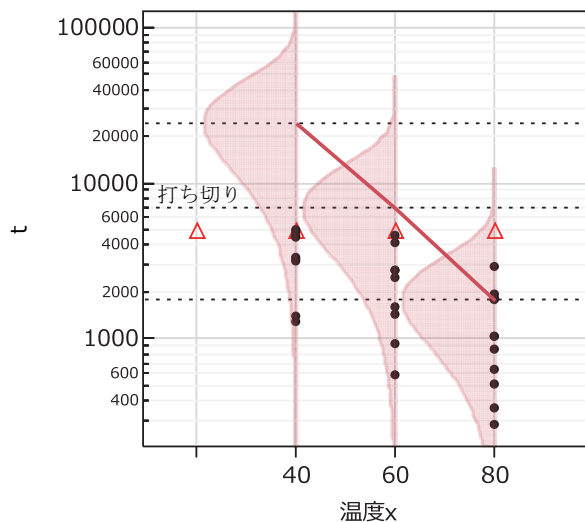
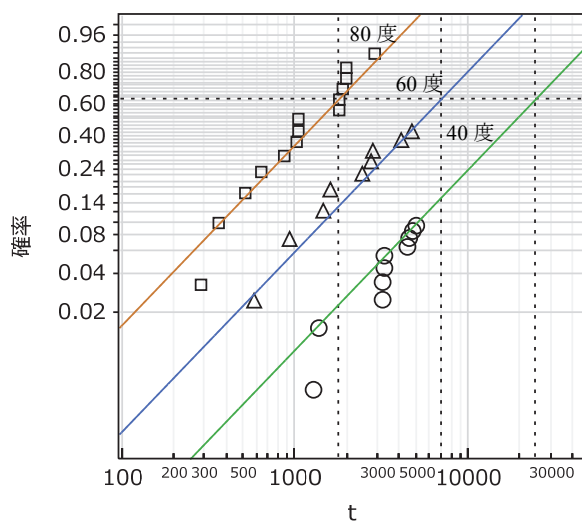


図 4.3 群ごとに α_j と共通の β によるあてはめ (JMP 寿命の二変量)

4.5 回帰パラメータを用いたワイブル分布のあてはめ

群ごとに異なるパラメータを与えた場合には、すべてのデータが打ち切りとなった10°Cでの結果がまったく反映されない。Weibull β を共通として、10°Cの場合の Weibull α を無理やり推定してみが、推定値が発散傾向 10^7 となり、異なる初期値に対して同じ解が得られない。

図 4.3 (右) の分布をみると温度変化によって対数変換した時間に対して、Weibull α が直線的に変化していることが確認される。実際、表 4.3 で推定された $\hat{\alpha}_j$ 推定値を Y 軸に温度を X 軸にプロットし、回帰式を上書きすると対数変換することにより直線状にきれいに乗ることが確認される。

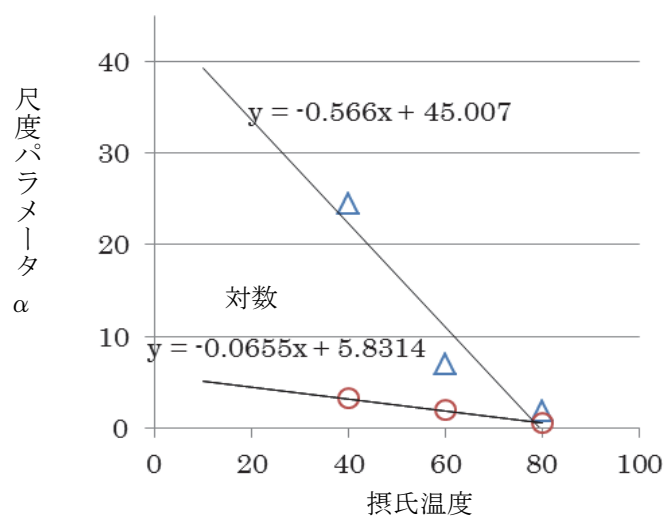


図 4.4 尺度 α_j と対数尺度 $\ln \alpha_j$ と温度の関係 (Excel)

この関係を Excel シートに落とし込んだのが、表 4.5 である。回帰式の初期値は、切片を $\hat{\alpha}_{\text{切片}} = 5.831$ 、傾きを $\hat{\alpha}_{\text{傾き}} = -0.065$ から、

$$\hat{\alpha} = \exp(5.831 - 0.065 \times \text{温度}_j)$$

とし、 $\beta = 1.43$ を初期値として、対数尤度を計算すると $\ln L = -95.700$ が得られる。

α	α 切片 =	5.831	α 傾き =	-0.065	$\ln L =$	-95.700
β	共通 $\beta =$	1.430				

Excel のソルバーで $\ln L$ を最大化した場合には、 $\hat{\alpha}_{\text{切片}} = 5.885$ 、 $\hat{\alpha}_{\text{傾き}} = -0.066$ 、 $\hat{\beta} = 1.426$ が得られる。対数尤度は $\ln L = -95.659$ と、わずかではあるが増大する。

α	α 切片 =	5.885	α 傾き =	-0.066	$\ln L =$	-95.659
β	共通 $\beta =$	1.426				

推定された 10°Cでの $\hat{\alpha}_1 = 185.324$ (千) と推定されている。また、生存関数 $1 - F(t; 10^\circ\text{C}) = 0.994$ から 5000 時間での下側確率が 0.6%であることが読み取れる。

形状を共通としたので、図 4.3 に示すように、故障データが全くない 10°Cの場合でもワイブル分布のあてはめ直線が、他の群と平行線として上書きされている。また、10°Cでのワイブル分布も示すことが可能となった。ワイブル分布の下側確率 0.632 を結ぶ直線がさらに上書きされている。

表 4.5 温度の回帰式で Weibull パラメータを設定した場合

$\alpha_i = \exp(\alpha_{\text{切片}} + \alpha_{\text{傾き}} \times \text{温度}_i)$		α		$\alpha_{\text{切片}} = 5.885$		$\alpha_{\text{傾き}} = -0.066$		$\ln L = -95.659$			
変数 x		件数 δ		パラメータ		ワイブル分布		尤度			
群	温度	t_{1000}	数	δ	生存率	α	β	$f(t)$	$1-F(t)$	L_i	$\ln L_i$
	0					359.687					
1	10	5.000	30	0	1.000	185.329	1.426	0.002	0.994	0.994	-0.173
2	40	1.298	1	1	0.990	25.351	1.426	0.016	0.986	0.016	-4.159
2	40	1.390	1	1	0.980	25.351	1.426	0.016	0.984	0.016	-4.132
:											
2	40	4.982	1	1	0.900	25.351	1.426	0.025	0.906	0.025	-3.670
2	40	5.000	90	0	0.900	25.351	1.426	0.026	0.906	0.906	-8.883
3	60	0.581	1	1	0.950	6.730	1.426	0.072	0.970	0.072	-2.626
3	60	0.925	1	1	0.900	6.730	1.426	0.086	0.943	0.086	-2.457
:											
3	60	4.674	1	1	0.550	6.730	1.426	0.100	0.552	0.100	-2.301
3	60	5.000	11	0	0.550	6.730	1.426	0.097	0.520	0.520	-7.199
4	80	0.283	1	1	0.933	1.787	1.426	0.339	0.930	0.339	-1.083
4	80	0.361	1	1	0.867	1.787	1.426	0.364	0.903	0.364	-1.009
:											
4	80	2.884	1	1	0.067	1.787	1.426	0.135	0.138	0.135	-2.001
4	80	5.000	1	0	0.067	1.787	1.426	0.016	0.013	0.013	-4.340

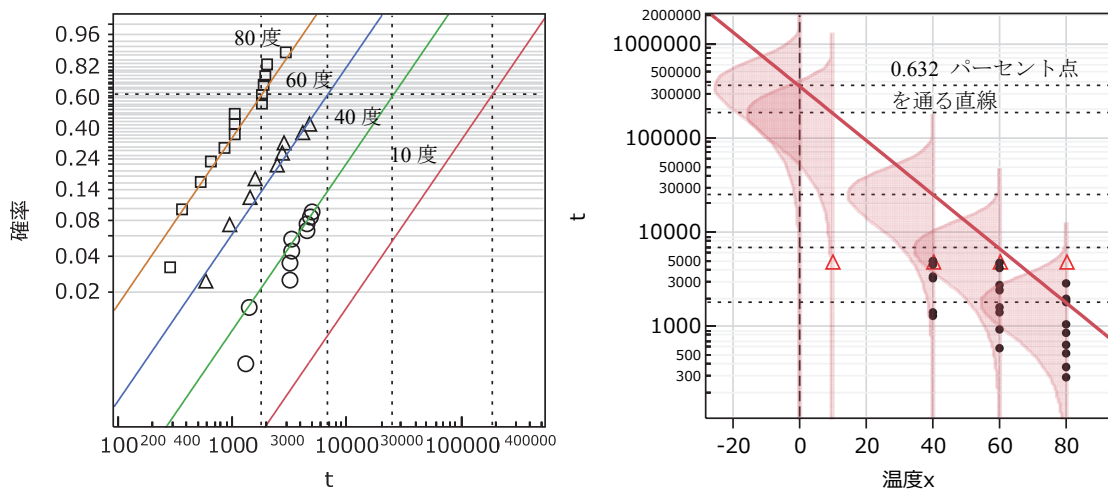


図 4.5 温度の回帰式で Weibull パラメータを設定した場合 (JMP 寿命の二変量)

4.6 アレニウス変換温度

加速因子として温度を設定した場合に、高温の場合の単位温度と低温の場合の単位温度での劣化度合いが異なることが知られている。したがって、通常の使用下での 10°C の場合の推定された Weibull $\hat{\alpha}_1 = 185.324$ (千) は過少評価となることが示唆される。

摂氏温度 x を用いた回帰式による Weibull α の推定は、

$$\alpha_i = \exp(\alpha_{\text{切片}} + \alpha_{\text{傾き}} \text{温度} x_i) \quad (4.7)$$

であった。この摂氏温度の代わりに

$$\text{アレニウス変換温度} = \frac{1}{k_B T_i} = \frac{11605}{\text{温度} x_i + 273.15} \quad (4.8)$$

を使う。ここで、 k_B はボルツマン定数: 8.6171×10^{-5} 、 T_i は絶対温度となり、 k_B の逆数は 11,605 となり、 $T_i = \text{温度} x_i + 273.15$ である。回帰式は、

$$\alpha_i = \exp\left(\alpha_{\text{切片}} + \alpha_{\text{傾き}} \frac{11605}{\text{温度} x_i + 273.15}\right) \quad (4.9)$$

となり、この式によりワイブル・パラメータの推定をおこなう。摂氏温度とアレニウス変換温度の関係を 図 4.6 に示す。摂氏温度が高くなるにつれアレニウス変換温度の変化が小さくなるのがわかる。

摂氏温度の代わりにアレニウス変換温度を用いてワイブル分布のあてはめを行う際にパラメータの初期値の設定が重要となる。闇雲に初期値を設定しても Excel では解を求めることはできにくい。アレニウス変換温度と各温度でのパラメータの対数の関係から、切片を $\alpha_{\text{切片}} = -19.701$ 、傾きを $\alpha_{\text{傾き}} = 0.6185$ が得られるのでこれらを初期値とする。

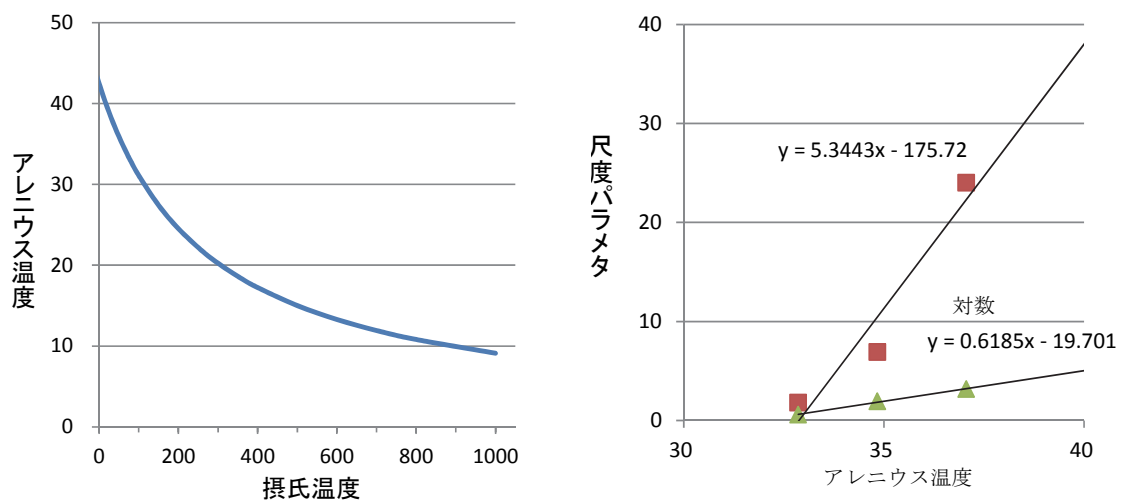


図 4.6 摂氏温度のアレニウス変換温度 (Excel)

表 4.6 で示すように Excel のソルバーで $\ln L$ を最大化した場合には、 $\hat{\alpha}_{\text{切片}} = -20.224$ 、 $\hat{\alpha}_{\text{傾き}} = 0.634$ 、 $\hat{\beta} = 1.414$ が得られる。対数尤度は $\ln L = -95.663$ となる。通常使用下での 10°C での Weibull $\hat{\alpha}_1 = 314.758$ (千) となり、摂氏温度での Weibull $\hat{\alpha}_1 = 185.324$ (千) よりも大幅な時間の推定値となる。

表 4.6 アレニウス変換温度

						α	$\alpha_{\text{切片}} = -20.224$	$\alpha_{\text{傾き}} = 0.634$	$\ln L = -95.663$			
						β	共通 $\beta = 1.414$					
						パラメータ		ワイブル分布		尤度		
群	変数 x	アレニウス温度	アレニウス温度	件数	δ	生存率	α	β	$f(t)$	$1-F(t)$	L_i	$\ln L_i$
	0	42.5					814.674					
1	10	41.0	5.000	30	0	1.000	314.758	1.414	0.001	0.997	0.997	-0.086
2	40	37.1	1.298	1	1	0.990	26.135	1.414	0.015	0.986	0.015	-4.175
2	40	37.1	1.390	1	1	0.980	26.135	1.414	0.016	0.984	0.016	-4.148
:												
2	40	37.1	4.982	1	1	0.900	26.135	1.414	0.025	0.909	0.025	-3.699
2	40	37.1	5.000	90	0	0.900	26.135	1.414	0.025	0.908	0.908	-8.675
3	60	34.8	0.581	1	1	0.950	6.380	1.414	0.079	0.967	0.079	-2.533
3	60	34.8	0.925	1	1	0.900	6.380	1.414	0.093	0.937	0.093	-2.372
:												
3	60	34.8	4.674	1	1	0.550	6.380	1.414	0.102	0.525	0.102	-2.279
3	60	34.8	5.000	11	0	0.550	6.380	1.414	0.099	0.492	0.492	-7.792
4	80	32.9	0.283	1	1	0.933	1.827	1.414	0.333	0.931	0.333	-1.101
4	80	32.9	0.361	1	1	0.867	1.827	1.414	0.357	0.904	0.357	-1.029
:												
4	80	32.9	2.884	1	1	0.067	1.827	1.414	0.139	0.149	0.139	-1.974
4	80	32.9	5.000	1	0	0.067	1.827	1.414	0.018	0.016	0.016	-4.152

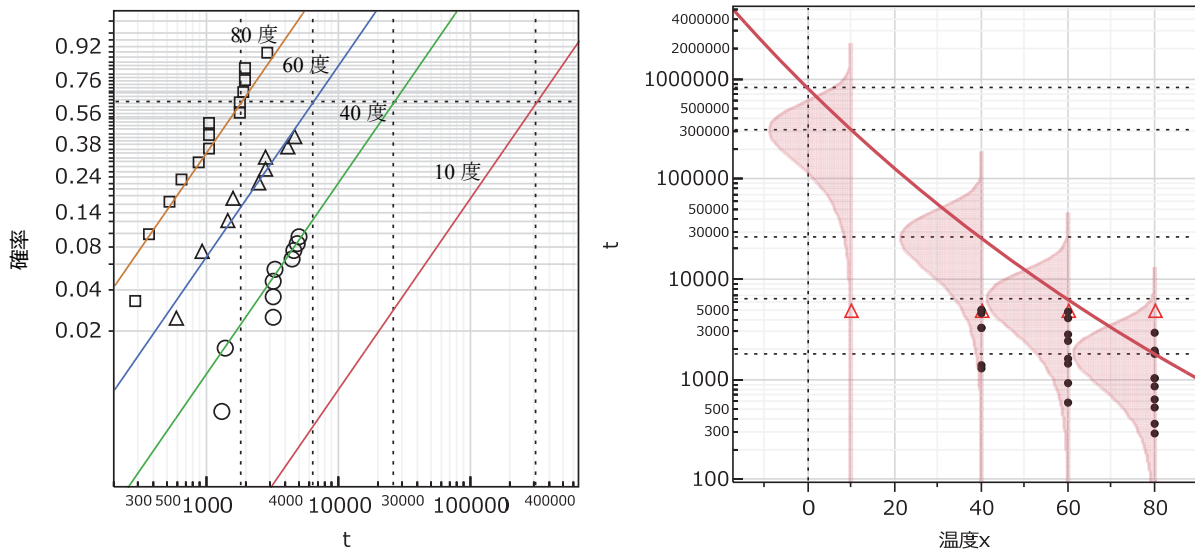


図 4.7 アレニウス変換温度を用いたあてはめ (JMP 寿命の二変量)

5. JMP の寿命の二変量を用いた加速試験の解析

5.1 寿命の二変量の基礎

第 4 章では寿命の二変量を用いた結果の表示を行ってきた。JMP になれた読者であれば、こんな結果が得られると示すだけで細かな操作手順を示さなくとも、サクサクと追試ができると想定する。しかしながら、JMP を使用したことがない読者、JMP に不慣れな読者に対しては、基本的な操作の考え方、使用上の注意点が必要と思われる。

デバイス A の故障データは、表 3.5 で示した JMP のサンプル・ライブラリの Devalt から得られる。「 寿命の二変量」を選択しスクリプトを実行すると、自動的に標準的な解析結果を得ることができる。

表 5.1 デバイス A の故障データ

	時間	ステータス	重み	温度	打ち切りの有無	x	
z	1	5000	Censored	30	10	1	40.9853435
	2	1298	Failed	1	40	0	37.0589175
	3	1390	Failed	1	40	0	37.0589175
	4	3187	Failed	1	40	0	37.0589175
	5	3241	Failed	1	40	0	37.0589175
	6	3261	Failed	1	40	0	37.0589175
:							
	32	1777	Failed	1	80	0	32.8613903
	33	1856	Failed	1	80	0	32.8613903
	34	1951	Failed	1	80	0	32.8613903
	35	1964	Failed	1	80	0	32.8613903
	36	2884	Failed	1	80	0	32.8613903
z	37	5000	Censored	1	80	1	32.8613903

以下に示すのは、元の Devalt の変数名を変更し、新たな変数を追加した JMP データセットを使用している。寿命の二変量の解析では、図 5.1 に示すように「Y.イベントまでの時間」, 「X」, 「打ち切り」, 「度数」に列の選択で表示されている変数名リストから該当する変数名を選択する。変数名「温度」は名義尺度として変更してあることに注意してもらいたい。これは、前節での解析の流れに沿って JMP の使用方法を示すためである。変数名は、時間に関しては、 t , $t/1000$, $y=\ln(t)$ の 3 変数とし、重みは件数なので n とし、「打ち切りの有無」は、

単に「打ち切り」と変更した。

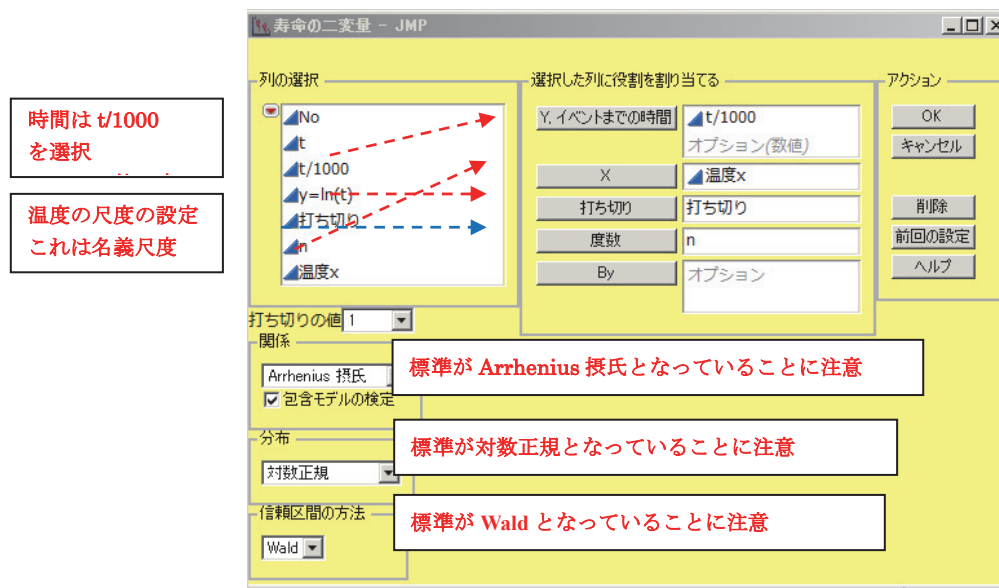


図 5.1 JMP の寿命の二変量の画面

1) 温度を名義尺度とした場合の設定画面

図 5.2 に示すように、**X** に名義尺度の「温度」を選択すると「Arrhenius 摂氏」から「位置と尺度」に自動的に変更される。分布を「対数正規」から「Weibull」に変更する。

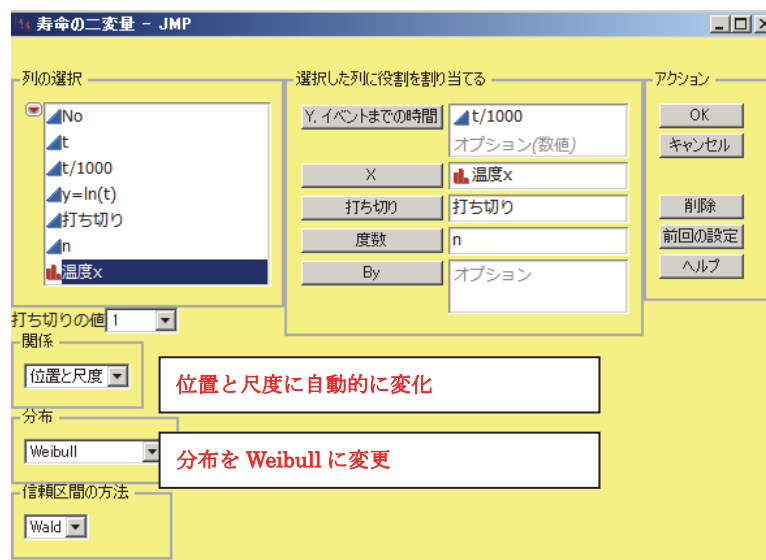


図 5.2 温度が名義尺度の場合の設定画面

2) 温度を連続尺度とした場合の設定画面

変数「温度」を連続尺度に変更した場合には、「Arrhenius 摂氏」から「線形」を選択する。

温度の尺度の設定
ここでは連続尺度

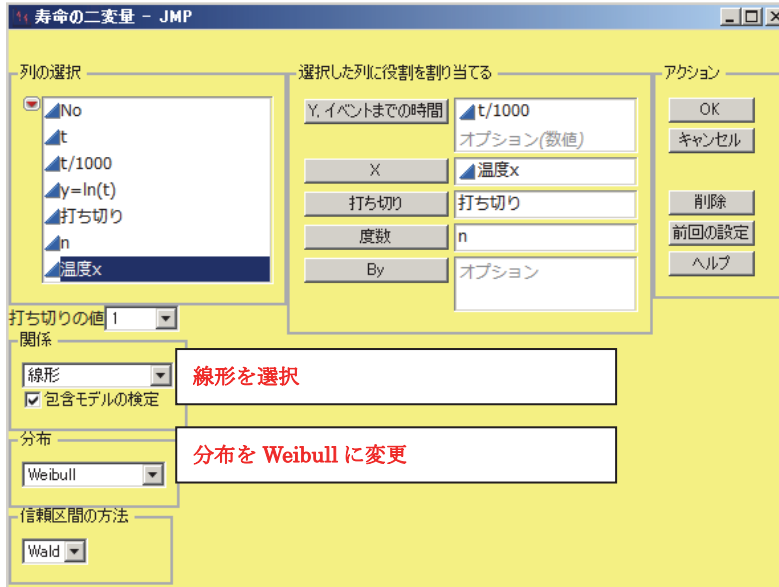


図 5.3 温度が連続尺度の場合で変換温度を使用しない設定画面

3) アレニウス摂氏温度を連続尺度とした場合の設定画面

分布を「対数正規」から「Weibull」に変更する.

温度の尺度の設定
ここでは連続尺度

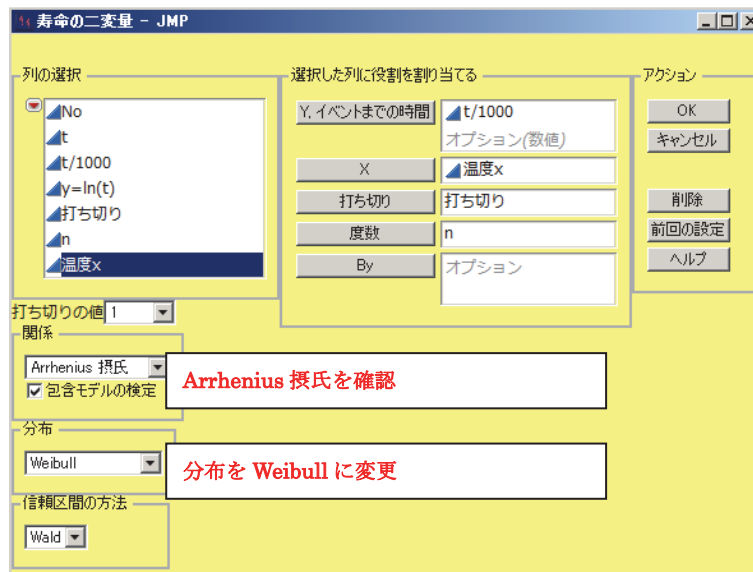


図 5.4 温度が連続尺度の場合でアレニウス摂氏温度を使う設定画面

5.2 温度を名義尺度にしたワイブル分布のあてはめ

温度を連続尺度ではなく名義尺度とした場合は多群比較となる. 図 5.1 で示した変数の選択で標準的に得られる結果と画面上で変更した図を並列して示す. 各温度に密度曲線を付加するためには, 「 散布図」から「密度曲線の追加」を選択すると, 自動的に表示される.

また、分位点 0.632 は、「分位点曲線の追加」で、数値を入力することで描ける。

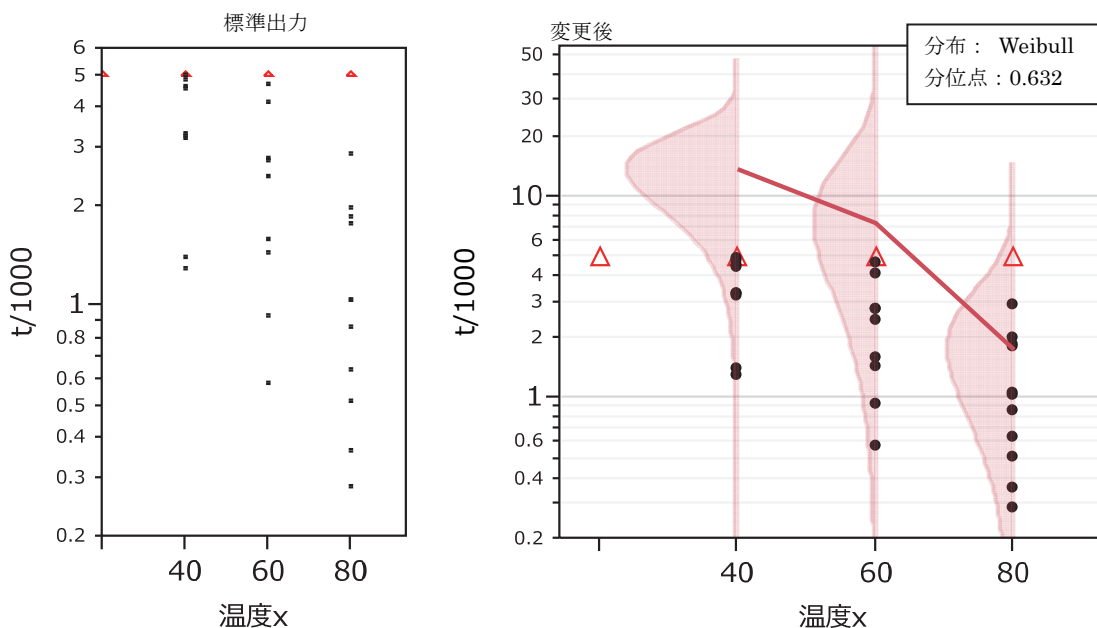


図 5.5 密度曲線・分位点の追加

ワイブル確率プロットも 図 5.6 に示すように見栄えの良い図に画面上で変更できる。

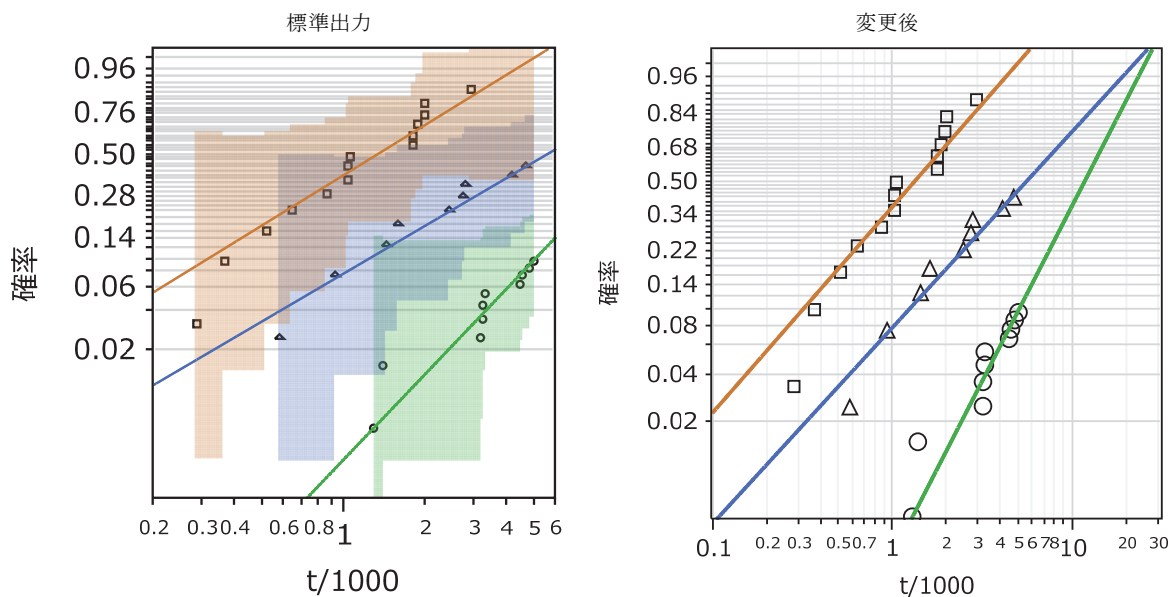


図 5.6 群ごとに異なるワイブル確率プロット

表 4.3 で Excel により計算した各温度のワイブル・パラメータについて、JMP で計算し、その結果を 表 5.2 に示す。推定値が、一致していることが確かめられる。

表 5.2 ワイブル・パラメータの推定

	パラメータ	推定値	標準誤差	下側95%	上側95%
40°C	Weibull α 0	13.717	4.781	6.927	27.161
	Weibull β 0	2.233	0.697	1.385	5.758
60°C	Weibull α 1	7.406	2.313	4.016	13.658
	Weibull β 1	1.249	0.384	0.779	3.149
80°C	Weibull α 2	1.740	0.366	1.152	2.629
	Weibull β 2	1.312	0.267	0.937	2.185

表 4.3 (再掲) 群ごとに異なる α_j と β_j の推定

$\alpha = \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4$	α	-	13.717	7.406	1.740	$\ln L_i =$	-94.387
$\beta = \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$	β	-	2.233	1.249	1.312		
	群	1	2	3	4		

5.3 温度を名義尺度だが Weibull β を共通

図 5.2 で示したように温度を名義尺度とした場合に「関係」が「位置と尺度」となっているのを「位置」に変更すれば、Weibull β を共通とするワイブル分布のあてはめが直接できる。統計的観点から「別々」と「共通」とした場合のどちらが適切化の判断を行いたい。そこで、「関係」が「位置と尺度」から、Weibull β を共通とする結果を得るため操作手順を示す。

標準出力で 図 5.7 に示すよう「位置と尺度」が異なるワイブル確率プロット (左) が結果として表示され、Weibull β を共通とするが、「位置」は異なるワイブル確率プロット (右) も同時に表示される。

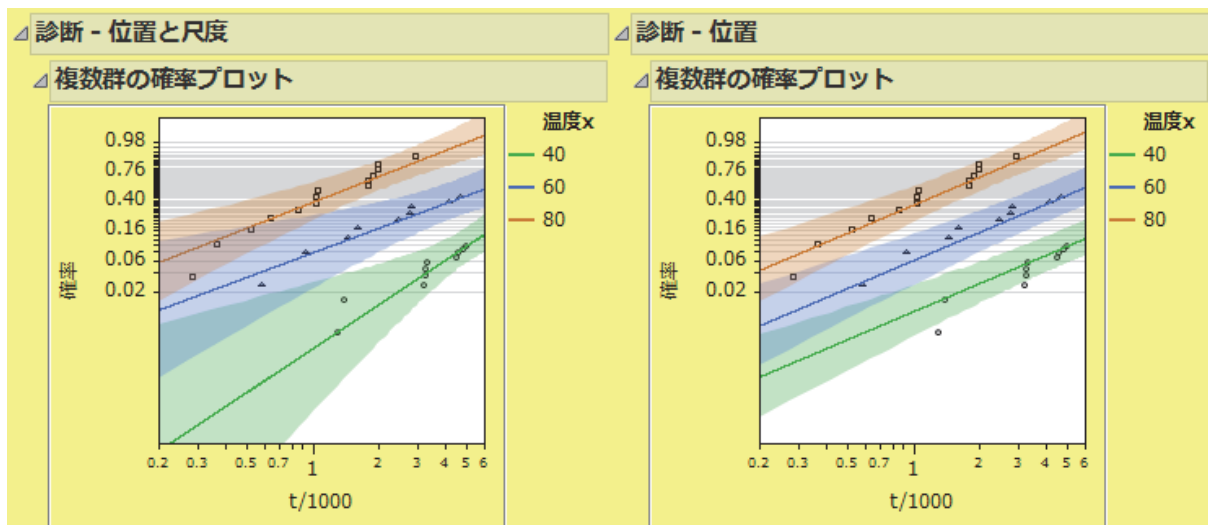


図 5.7 Weibull β を共通とするワイブル確率プロット (右)

統計的には、どちらのモデルが適切なのかに興味がある。その判断のためには、標準出力の「包括モデルの検定」を使う。

表 5.3 モデルの検定統計量

包含モデルの検定					
モデル					
診断統計量	モデル	(-2)*対数尤度	AICc	BIC	パラメータ数
<input type="checkbox"/>	効果なし	270.7126	274.7867	280.9245	2
<input checked="" type="checkbox"/>	位置	190.9255	201.3029	216.4553	5
	位置と尺度	188.775	205.6981	229.6226	8
モデル比較の検定					
説明	尤度比カイ2乗	自由度	p値(Prob>ChiSq)		
効果なし vs. 位置	79.78709	3	<.0001*		
位置 vs. 位置と尺度	2.150544	3	0.5418		

モデルの欄の「位置と尺度」は、表 4.3 の「群ごとに異なる α_j と β_j の推定」となっている。(-2)*対数尤度 の欄から、188.775 が読み取れる。これは、表 4.3 の対数尤度 のの(-2)倍の $-94.387 \times (-2) = 188.774$ である。

モデルの欄の「位置」は、表 4.4 の「群ごとに異なる α_j と共通の β の推定」となっている。(-2)*対数尤度 の欄から、190.9255 が読み取れる。これは、表 4.4 の対数尤度 のの(-2)倍の $-95.463 \times (-2) = 190.926$ である。少数点 2 桁から一致しないが、尤度を最大化する逐次計算の「終了」条件が Excel のソルバーに比べて JMP の方が高精度となっているためである。

統計的な観点では、「位置と尺度」のモデルと「位置」のモデルの間で、統計的に差があるかの判断が必要となる。このために、2つのモデルでの対数尤度の差の2倍が、カイ2乗分布に従うことを用いる。カイ2乗分布の自由度は、2つのモデルで用いられているパラメータ数の差となる。

	(-2) 対数尤度	差	自由度
位置と尺度	188.775	-	8
位置	190.9255	2.151	5

対数尤度の差は、元の尤度では比となるので、一般的には「尤度比検定」と言われている。そのために「モデル比較の検定」では、「尤度比カイ2乗」との表記となっている。「位置 vs. 位置と尺度」の行の「尤度比カイ2乗」の欄の2.151が自由度3のカイ2乗分布に従うこと

から、上側確率 0.5415 が得られる。従って、統計的には有意な差ではないので、パラメータ数の少ない「位置」のみの違いのモデルが推奨される。なお、p 値は、Excel の関数 $=1-CHISQ.DIST(2.151,3,TRUE)=0.5418$ によって確認することができる。

「位置」のモデルでの統計量の算出は、モデルの欄の 位置 を選択することにより得られる。表 5.4 に推定結果を示す。これは、表 4.4 の結果と一致している。

表 5.4 Weibull β を共通とするパラメータ

パラメータ	推定値	標準誤差	下側95%	上側95%
40℃ Weibull α 0	24.420	7.894	12.960	46.015
60℃ Weibull α 1	6.942	1.683	4.316	11.167
80℃ Weibull α 2	1.780	0.342	1.222	2.593
Weibull β	1.427	0.210	1.108	2.004

4.4 温度を連続尺度にしたワイブル回帰

温度 x を名義尺度から連続尺度に変更し、図 5.3 に示すようにモデルを設定し実行すると、故障データがなかった 10℃が表示される。これは、10℃の 5,000 時間での 30 件の打ち切りデータについて、5,000 時間でのワイブル分布の上側確率（生存率）を尤度として活用した結果となっている。

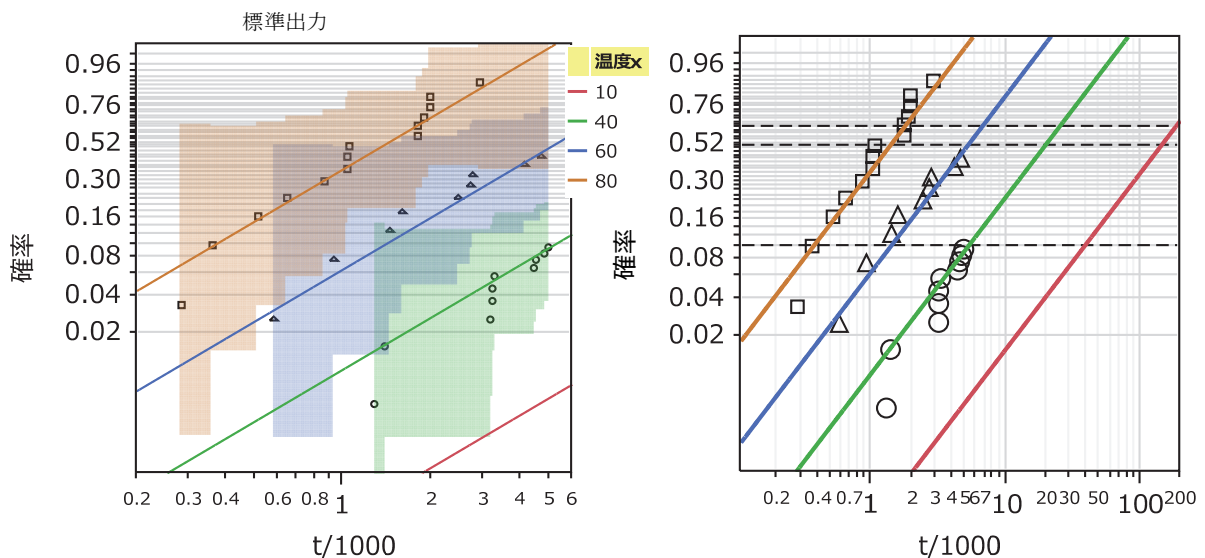


図 5.8 温度 x を連続尺度とした場合のワイブル確率プロット

表 5.5 に最尤法による回帰分析の結果を示す。得られた回帰式は、

$$\mu = \hat{\beta}_0 + \hat{\beta}_1 x = 5.882 - 0.663x$$

である。この結果は、表 4.5 に示した Excel での結果に一致するが、パラメータの表記が

異なる。表 4.5 では、Weibull α に関する α 切片, α 傾き としている。対応関係は、

$$\hat{\beta}_0 = \alpha_{\text{切片}}$$

$$\hat{\beta}_1 = \alpha_{\text{傾き}}$$

となる。

表 5.5 ワイブル分布を誤差とする回帰パラメータおよび分散共分散の推定

パラメータ	推定値	標準誤差	下側95%	上側95%	β_0	β_1	σ
β_0	5.8852	0.6849	4.5429	7.2276	0.4691	-0.0067	0.0514
β_1	-0.0663	0.0100	-0.0859	-0.0468	-0.0067	0.0001	-0.0007
σ	0.7011	0.1018	0.5016	0.9006	0.0514	-0.0007	0.0104
$\mu = 5.885247 - 0.06631011 * \text{温度}$							

切片は、 $\exp(5.885) = 359.6_{(千)}$ となる。

JMP の寿命の二変量では、時間の実時間でワイブル分布のあてはめではなく、時間の自然対数に対する最小極値分布 (第 6 章 参照) のあてはめを行っている。そのために、Weibull β ではなく σ となっている。これは、Weibull β の逆数で、

$$\sigma = 1/\beta = 1/1.426 = 0.701$$

の関係となっている。

図 5.9 の散布図には、それぞれの温度に対して最小極値分布があてはめられている。なお、実目盛とすれば、もちろんワイブル分布となるが、レンジが広くなり図としては不適當である。

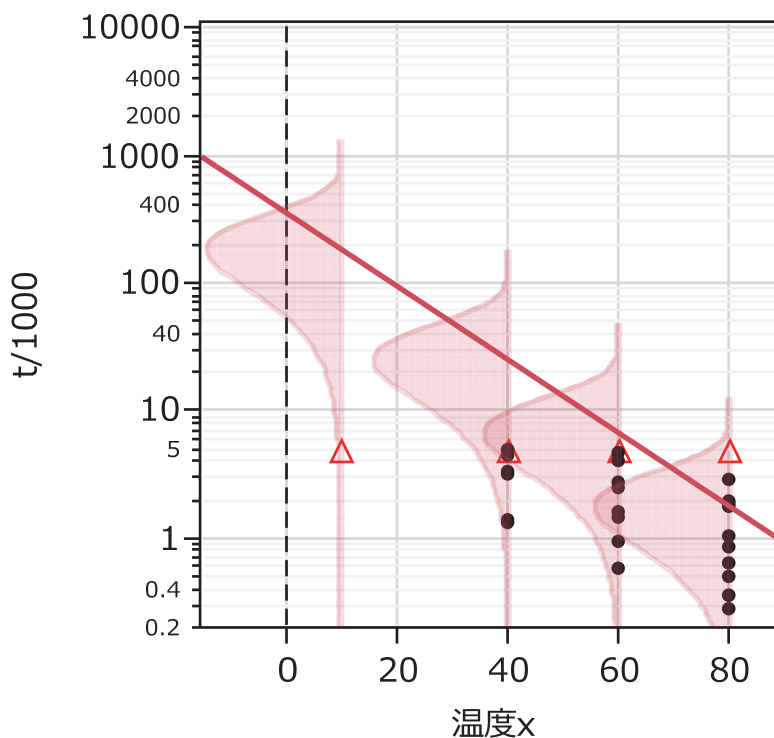


図 5.9 最小極値分布を誤差とする回帰直線

る。回帰直線は、それぞれの温度で下側確率が 0.632 となるような直線であり、どの温度でも常用対数目盛上での位置は異なるが、分布の形状は同じである。正規分布の「標準偏差 σ 」と同様の性質を持つ $\sigma = 0.701$ が使われている。

回帰直線をあてはめることの統計的な妥当性は、表 5.6 に示す包括モデルの検定結果から得られる。統計的な観点では、「回帰 vs. 別々の位置」の行の「尤度比カイ 2 乗」の欄の 0.239 が自由度 2 のカイ 2 乗分布に従うことから、上側確率 0.8219 となり、統計的には「別々の位置」のモデルではなく、「回帰」モデルが推奨される。

表 5.6 包括モデルの検定

モデル					
診断統計量	モデル	(-2)*対数尤度	AICc	BIC	パラメータ数
<input type="checkbox"/>	効果なし	270.7126	274.7867	280.9245	2
<input type="checkbox"/>	回帰	191.3177	197.4668	206.6355	3
<input checked="" type="checkbox"/>	別々の位置	190.9255	201.3029	216.4553	5
<input checked="" type="checkbox"/>	別々の位置と尺度	188.775	205.6981	229.6226	8

モデル比較の検定				
説明	尤度比カイ2乗	自由度	p値(Prob>ChiSq)	
効果なし vs. 回帰	79.395	1	<.0001*	
回帰 vs. 別々の位置	0.392	2	0.8219	
別々の位置 vs. 別々の位置と尺度	2.151	3	0.5418	

5.5 アレニウス変換温度によるワイブル回帰

JMP の寿命の二変量で、図 5.4 で示したように、連続尺度を に設定すると には、Arrhenius 摂氏が自動的に設定される。なお、Arrhenius 華氏、Arrhenius ケルビンも用意されているので、必要に応じて選択ができるようになっている。分布を Weibull に設定して実行すると、図 5.10 に示すような散布図が得られる。

この散布図では、X 軸の目盛がアレニウス摂氏形式なり、回帰直線が示されている。元の摂氏温度での表示は、「 散布図」から変換スケールのチェックを外すことにより得られる。図 5.11 に示すように、X 軸は摂氏温度目盛りとなり、下側確率 0.632 は曲線となる。

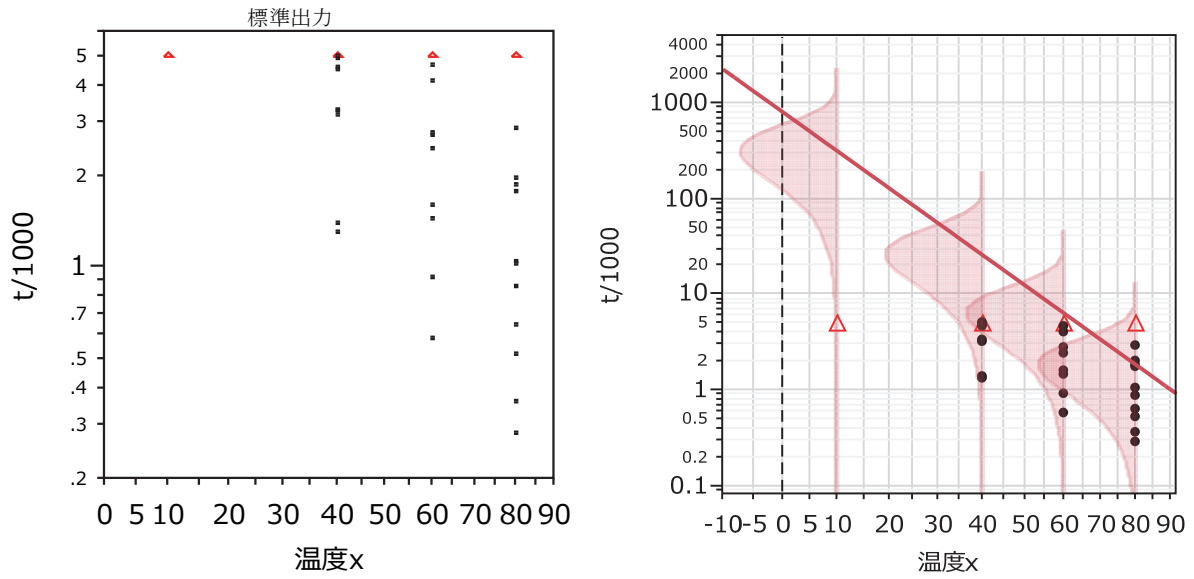


図 5.10 ワイブル分布を誤差とする回帰直線

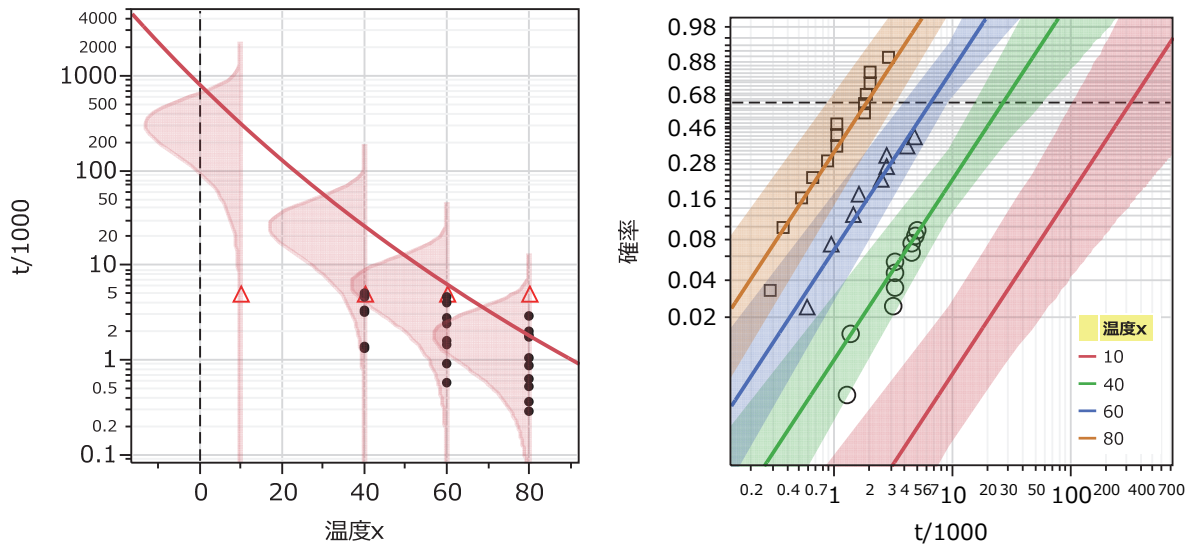


図 5.11 摂氏温度での推定曲線

推定されたパラメータを表 5.7 に示す. 切片 $\beta_0 = -20.225$ となり, 傾き $\beta_1 = 0.634$ となっている. これは, アレニウス変換温度を x とした場合の係数で, 10°C の場合は, $11605 / (10 + 273.15) = 40.9853$, 40°C の場合 37.0589 , 60°C の場合 34.8342 , 80°C の場合 32.8614 と大小関係が逆となることに起因している.

図 5.12 にアレニウス摂氏温度を用いた回帰直線を示す. 内部では, この図のような解析が行われているのであるが, グラフ表示に際して, 元の摂氏温度に換算して結果の表示が行われている.

表 5.7 アレニウス摂氏を用いたワイブル・パラメータの推定

推定値				
パラメータ	推定値	標準誤差	下側95%	上側95%
β_0	-20.225	3.313	-26.718	-13.731
β_1	0.634	0.097	0.444	0.824
σ	0.707	0.103	0.505	0.909

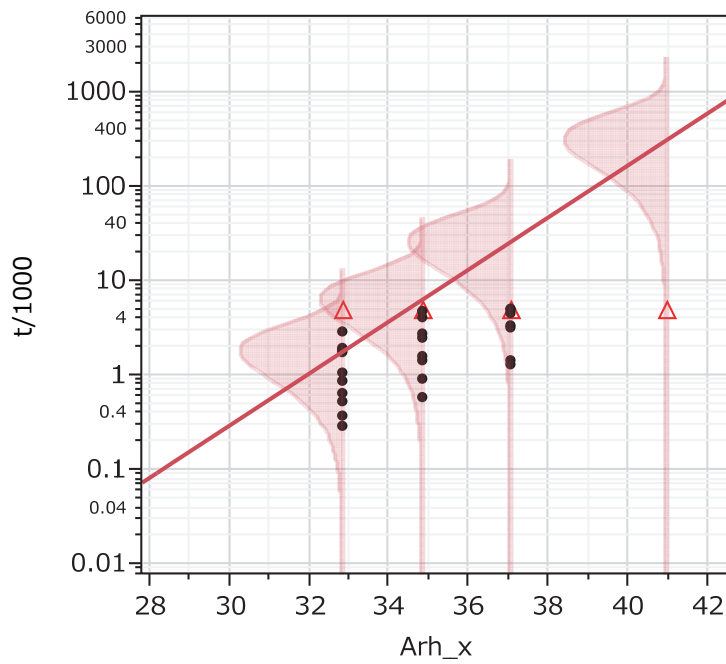
$$\mu = -20.22459 + \frac{0.6337984 * 11605}{(\text{温度}x + 273.15)}$$


図 5.12 アレニウス摂氏温度を用いた直線のあてはめ

加速係数プロファイル

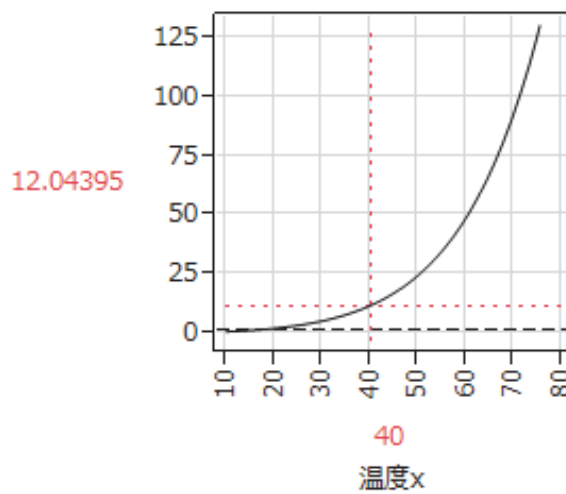


図 5.13 温度 10°Cを基準とした加速係数プロファイル

通常使用される温度，この場合は 10℃であるが，この温度を基準に 40℃での故障がどの程度加速して故障が発生しているかを係数とし算出した結果を 図 5.13 に示す． 図 5.11 に示した 10℃と 40℃の下側確率 0.632 となる Y 軸の目盛を 図 5.15 で示す方法で推定すると，それぞれ 314.70_(千)時間と 26.13_(千)時間となる． この比 314.70 / 26.13 が加速係数 12.04 となる．

5.6 各種の推定

アレニウス摂氏に変換した回帰での，各種の推定方法について示す． 図 5.14 には，温度が 10℃の場合の故障率を 0.10， 0.50， 0.90 とした場合に，故障時間が寿命分位点として表示され，その 95%信頼区間も推定されている． 故障率が 0.10 の場合は，64.1282_(千) 時間と推定され，95%信頼区間は (22.7122, 181.0668) となる．



図 5.14 故障率を指定した場合の故障時間および 95%信頼区間の計算

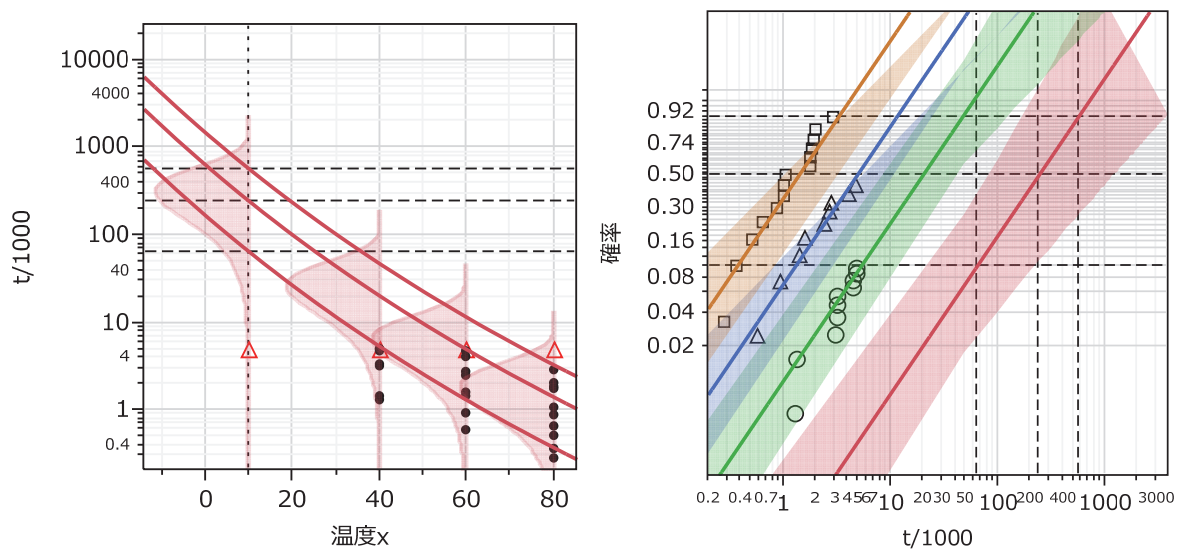


図 5.15 10℃における指定した故障率に対する各種推定

図 5.15 の散布図には故障率が (0.10, 0.50, 0.90) となる 3 本の推定曲線が上書きされている。この散布図上にさらに 10°Cでの寿命分位点 (64.12, 242.92, 567.64) を通る点線も書き加えた。この点線と 3 本の曲線が、10°C上で交叉することが確認できる。

ワイブル確率プロット上には、故障率が (0.10, 0.50, 0.90) となる点線が上書きされ、さらに寿命分位点 (64.12, 242.92, 567.64) を通る点線も書き加えた。これらの点線が、10°Cの直線上で交叉している。この交叉点を通る点線と 95%信頼区間の限界値を読めば、おおよその 95%信頼区間を読み取ることができる。

図 5.16 に、故障時間を指定した場合の故障率とその 95%信頼区間の推定結果を示す。温度が 10°Cの場合の故障時間が 30 (千) 時間を設定すると、故障率は、0.0353 と推定され、その 95%信頼区間は、(0.0093, 0.1290) となる。



図 5.16 10°Cにおける指定した故障率に対する各種推定

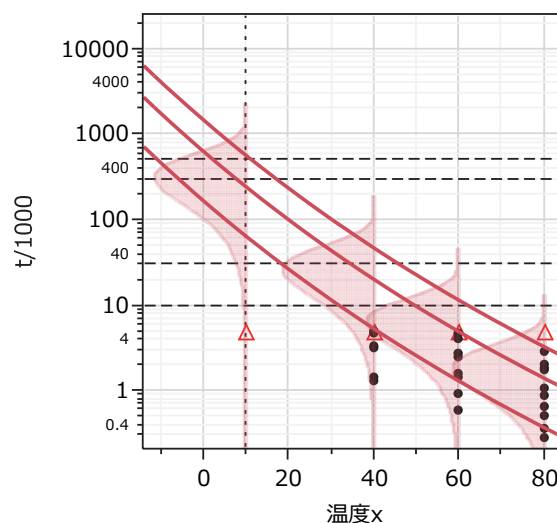


図 5.17 散布図上での 10°Cでの故障時間に対する故障率の推定

6. 寿命時間の対数変換による最小極値分布の活用

6.1 ワイブル分布を最小極値分布へ変換

これまで、時間 t を用いたワイブル分布を用いてきた。Excel にはワイブル分布の確率密度関数および累積分布関数が共にそろっていて、ソルバーを併用することで最尤法による寿命データの統計解析を手軽に行えることを示してきた。

Weibull α および Weibull β をワイブル確率紙で推定するためには、時間 t に関して対数目盛が使われている。加速試験で温度などの加速因子に対して、ワイブル分布の Weibull $\hat{\alpha}_i$ は、指数関数的に変化するので、時間について対数目盛して、回帰直線のあてはめが行われている。

ワイブル分布は、時間 t の関数であり、推定されたパラメータに対する意味付けがしやすいのであるが、加速因子を考慮した解析の場合には、各種の信頼区間の推定に必要な分散の合成が煩雑となる。

ワイブル分布の時間 t について $y=\ln(t)$ と自然対数を取り、Weibull α も自然対数 $\mu_{\text{SEV}} = \ln(\alpha)$ とし、Weibull β の逆数を $\sigma_{\text{SEV}} = \delta = 1/\beta$ とすることにより、正規分布と同様の統計的な性質を持つ最小極値分布 (SEV : Smallest Extreme Value Distribution) に置き換えることができる。これらを式 (3.1) のワイブル分布の累積分布関数 $F(t)$ に $\alpha = \exp(\mu_{\text{SEV}})$ 、 $\beta = 1/\sigma_{\text{SEV}}$ を代入する。最小極値分布のパラメータは、 μ_{SEV} 、 σ_{SEV} とするのが適切だが、文脈の中で識別できると思われる場合には、“SEV” は除くことにする。

$$\begin{aligned} F(t) &= 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right] \\ &= 1 - \exp\left[-\left(\frac{t}{\exp(\mu)}\right)^{1/\sigma}\right] \\ &= 1 - \exp\left\{-\exp\left[\ln\left(\frac{t}{\exp(\mu)}\right)^{1/\sigma}\right]\right\} \\ &= 1 - \exp\left[-\exp\left(\frac{\ln(t) - \mu}{\sigma}\right)\right] \end{aligned} \tag{6.1}$$

ここで、 $y=\ln(t)$ として、次の最小極値分布

$$F(y) = 1 - \exp\left[-\exp\left(\frac{y - \mu}{\sigma}\right)\right], \quad -\infty < y < \infty \quad (6.2)$$

を得る.

ワイブル分布と最小極値分布の関係は、正規分布と対数正規分布の関係と同様である。このような対数変換を行うことにより、ワイブル分布を正規分布と同様に、“平均値”として μ_{SEV} が、“標準偏差”として σ_{SEV} を用いて表すことができる。

最小極値分布の確率密度関数 $f(y)$ は、累積分布関数 $F(y)$ を y で微分することにより、

$$f(y) = \frac{1}{\sigma} \cdot \exp\left(\frac{y - \mu}{\sigma}\right) \cdot \exp\left\{-\exp\left(\frac{y - \mu}{\sigma}\right)\right\}, \quad -\infty < y < \infty \quad (6.3)$$

が得られる。パラメータ $\mu_{SEV} = 2.0$ としてパラメータ σ_{SEV} を (1.5, 1.0, 0.8, 0.4) と変えたときの確率密度関数および累積分布関数を 図 6.1 に示す。

ワイブル分布の確率密度関数の形状は、形状パラメータ $\beta \leq 1$ の場合に単調減少、 $\beta > 1.0$ の場合は右に裾をひく分布で一山型ではあるがその位置が一定ではなかった。時間について自然対数を取ることにより $\sigma_{SEV} = 1/\beta$ の大きさにかかわらず確率密度関数の最大値は、 $\mu_{SEV} = 2.0$ の位置となり、 $\sigma_{SEV} = 1.0$ の場合には、左に裾は引くものの正規分布の確率密度関数の形状に近い。

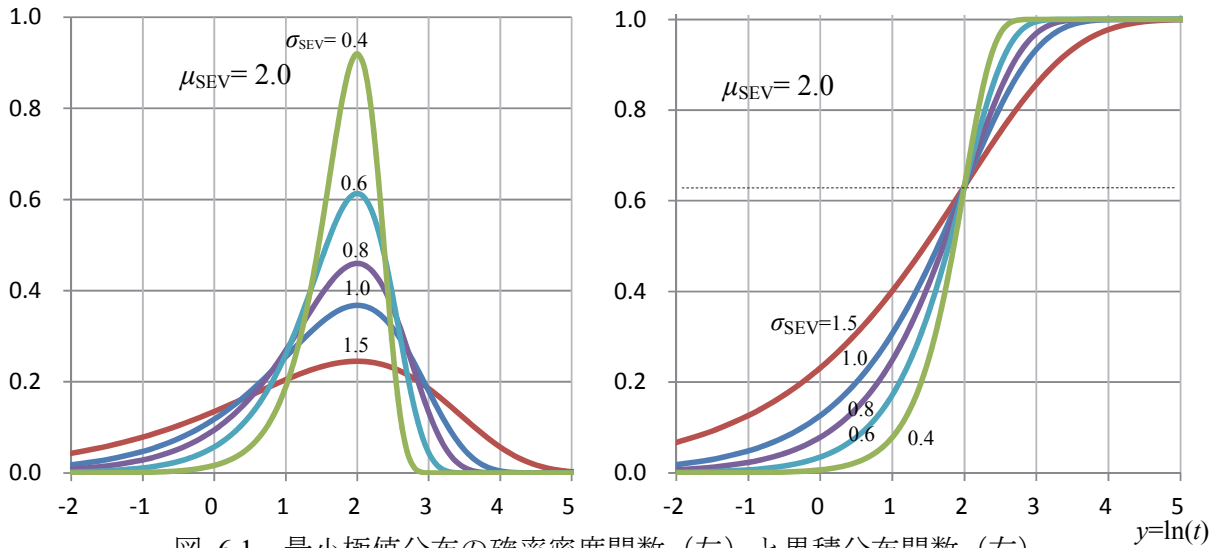


図 6.1 最小極値分布の確率密度関数 (左) と累積分布関数 (右)

累積分布関数は、ワイブル分布の場合には、図 3.1 に示したように Weibull $\beta \leq 1$ の場合に上に凸の単調増加、 $\beta > 1.0$ の場合に S 字型であったが、最小極値分布では、 $\sigma_{SEV} = 1/\beta$ の大きさにかかわらず S 字型となっている。また、 $y = \mu_{SEV} = 2.0$ における下側確率は、

$$F(y = \mu) = 1 - \exp\left[-\exp\left(\frac{\mu - \mu}{\sigma}\right)\right] = 1 - e^{-1} = 0.6321 \quad (6.4)$$

と一定であることは、ワイブル分布の $t = \alpha$ 場合に 0.632 になることと同じである。

このように最小極値分布は、パラメータ μ_{SEV} および σ_{SEV} に関して正規分布と同様の性質を持つために、多様な寿命データの解析に際して、従来の実験計画法に対する統計モデルの適用と同様な対応が可能となる。ワイブル分布の場合について様々な表記法が使われていることは、すでに述べたが、最小極値分布についても様々な表記法が混在する。JMP の中でさえも混在している。表 6.1 に表記法について示す。

表 6.1 各種のパラメータの表記法

ワイブル分布	JMP a: 生存時間分析	α	β	閾値
	JMP b: 寿命の一変量	Weibull: α	Weibull: β	閾値
	奥野(Nelson)	尺度: α	形状: β	シフト: γ
	日本の信頼性	尺度: η	形状: m	位置: γ
	Excel	β	α	
最小極値分布	JMP c: 生存時間分析	$\lambda = \ln(\alpha)$	$\delta = 1/\beta$	
	JMP d: 寿命の一変量	位置: location	尺度: scale	
	JMP e: 寿命の二変量	位置: $\mu = \ln(\alpha)$	尺度: $\sigma = 1/\beta$	
	奥野監訳本(Nelson)	位置: $\lambda = \ln(\alpha)$	尺度: $\delta = 1/\beta$	
	日本の信頼性	位置: u	尺度: $\alpha = 1/\beta$	

JMP の「品質管理および信頼性 / 生存時間」操作マニュアル p380 には、関連する書物の著者によりさまざまな表記法があることが、次のように示されている。

表 19.2 JMP の Alpha と Beta に該当する Weibull パラメータ

JMP Weibull	alpha	beta
Wayne Nelson	alpha=alpha	beta=beta
Meeker and Escobar	eta=alpha	beta=beta
Tobias and Trindade	c = alpha	m = beta
Kececioglu	eta=alpha	beta=beta
Hosmer and Lemeshow	exp(X beta)=alpha	lambda=beta
Blishke and Murthy	beta=alpha	alpha=beta
Kalbfleisch and Prentice	lambda = 1/alpha	p = beta
JMP 極値	lambda=log(alpha)	delta=1/beta
Meeker and Escobar s.e.v.	mu=log(alpha)	sigma=1/beta

JMP に新たに加えられた「寿命の一変量」, 「寿命の二変量」などは, Meeker and Escobar の著書に準拠して設計されており, 最小極値分布の場合は, μ_{SEV} および σ_{SEV} が使われている. 本書では, ワイブル分布の場合には, Weibull α , Weibull β を使い, 最小極値分布の場合は, μ_{SEV} , σ_{SEV} を用いる.

最小極値分布からワイブル分布への変換は,

$$\text{最小極値: } F(y) = 1 - \exp \left[-\exp \left(\frac{y - \mu_{\text{SEV}}}{\sigma_{\text{SEV}}} \right) \right]$$

について, $y = \ln(t)$, $\mu_{\text{SEV}} = \ln(\alpha)$, $\sigma_{\text{SEV}} = 1/\beta$ で置き換えて,

$$\begin{aligned} F(t) &= 1 - \exp \left\{ -\exp \left(\frac{\ln(t) - \ln(\alpha)}{1/\beta} \right) \right\} \\ &= 1 - \exp \left\{ -\exp \left(\ln \left(\frac{t}{\alpha} \right)^\beta \right) \right\} \\ &= 1 - \exp \left\{ -\left(\frac{t}{\alpha} \right)^\beta \right\} \end{aligned} \tag{6.5}$$

のように簡単にワイブル分布に戻すことができる.

6.2 最小極値分布を用いた最尤法

故障データに対しては, 各時点での故障率が最小極値分布の確率密度関数 $f(y)$ に従い, 打ち切りデータについては, 最小極値分布の上側確率となる生存関数 $[1 - F(y)]$ に従うとする. 故障した場合を $\delta_i = 1$, 打ち切りがある場合を $\delta_i = 0$ とした場合に尤度関数は,

$$\begin{aligned} L &= \prod_{i=1}^n f(y_i)^{\delta_i} [1 - F(y_i)]^{1-\delta_i} \\ &= \prod_{i=1}^n \left[\frac{1}{\sigma} \cdot \exp \left(\frac{y_i - \mu}{\sigma} \right) \right]^{\delta_i} \cdot \exp \left[-\exp \left(\frac{y_i - \mu}{\sigma} \right) \right] \end{aligned} \tag{6.6}$$

となる. 式 (6.6) から, 故障データは $\delta_i = 1$ なので確率密度関数に等しくなり, 打ち切りデータについては, $\delta_i = 0$ なので生存関数 $[1 - F(y)]$ に等しくなる. ディーゼル発電機ファンの 12 件の故障データ, 58 件の打ち切りデータに対し, Excel の最小極値分布の確率密度関数, 生存関数を用いた Excel の計算シートを表 6.2 に示す.

表 3.3 に示したワイブル分布のあてはめの Excel シートの関数の式を変更し, 初期値としては, $\hat{\mu}_{\text{SEV}} = \ln(\hat{\alpha}) = \ln(26296.9) = 10.177$, $\hat{\sigma}_{\text{SEV}} = 1/\hat{\beta} = 1/1.058 = 0.9452$ と換算値を使用する. 表 6.2 は, これらを初期値として, ソルバーで対数尤度 $\ln L$ を最大化した結果である.

表 6.2 ディーゼル発電機ファンに対する最小極値分布のあてはめ

		$\mu_{SEV} =$	10.177				$\ln L =$	-42.2480
		$\sigma_{SEV} =$	0.9448				$-2 \ln L =$	84.4960
No	t	$y = \ln(t)$	$(y - \mu) / \sigma + \delta$		$f(y)$	$1 - F(y)$	Li	$\ln Li$
1	450	6.109	-4.306	1	0.0141	0.9866	0.0141	-4.2624
3	1150	7.048	-3.313	1	0.0372	0.9642	0.0372	-3.2922
:								
45	6100	8.716	-1.547	1	0.1822	0.8082	0.1822	-1.7027
62	8750	9.077	-1.165	1	0.2417	0.7320	0.2417	-1.4199
2	460	6.131	-4.282	+ 0	0.0144	0.9863	0.9863	-0.0138
5	1560	7.352	-2.990	+ 0	0.0506	0.9509	0.9509	-0.0503
:								
69	10100	9.220	-1.013	+ 0	0.2673	0.6955	0.6955	-0.3632
70	11500	9.350	-0.875	+ 0	0.2907	0.6592	0.6592	-0.4167
		$f(y) = (1 / \sigma_{SEV}^5) * \text{EXP}(\sigma_{SEV}^7) * \text{EXP}(-\text{EXP}(\sigma_{SEV}^7))$						
		σ_{SEV}	$(y - \mu) / \sigma$			$(y - \mu) / \sigma$		
		$1 - F(y) = \text{EXP}(-\text{EXP}(\sigma_{SEV}^7))$			$Li = \text{IF}((G7=1), H7, I7)$			
			$(y - \mu) / \sigma$		δ	$f(y)$	$1 - F(y)$	

最初の $t_1 = 450$ 時間, $y_1 = \ln(450) = 6.109$ は, 故障データなので尤度 L_1 は, 確率密度 $f(y_1) = 0.0141$ となり, ワイブル分布の 0.000031 に比べて大きくなっている. これは, $y = \ln(t)$ としたために, 分布の取りうる範囲が狭くなったために $f(y_i)$ が相対的に大きくなるからである. 生存率は $[1 - \hat{F}(y_1)] = 0.9866$ とワイブル分布での結果と一致している.

最後の故障は 62 番目 $y_{62} = \ln(8750) = 9.077$ であり, 生存率は $[1 - \hat{F}(y_{62})] = 0.7320$ と推定されている. 最後まで観察された打ち切りデータは 11,500 時間で, 生存率は $[1 - \hat{F}(y_{70})] = 0.6592$ と推定されている. なお, JMP による最小極値分布のあてはめの解析結果は, 時間 t を用いた Weibull パラメータの推定結果に併記され, 表 3.7 では (位置, 尺度) で, 表 3.11 では (λ, δ) で表されている.

図 6.2 に位置パラメータ $\hat{\mu}_{SEV} = 10.177$, 形状パラメータ $\hat{\sigma}_{SEV} = 0.9448$ を用いて最小極値分布の確率密度曲線 $f(y)$, 生存関数 $[1 - F(y)]$ の上に各データを○印で表示する. X 軸は常用対数目盛となっているので, $\hat{\mu}_{SEV}$ について指数換算値を示してある.

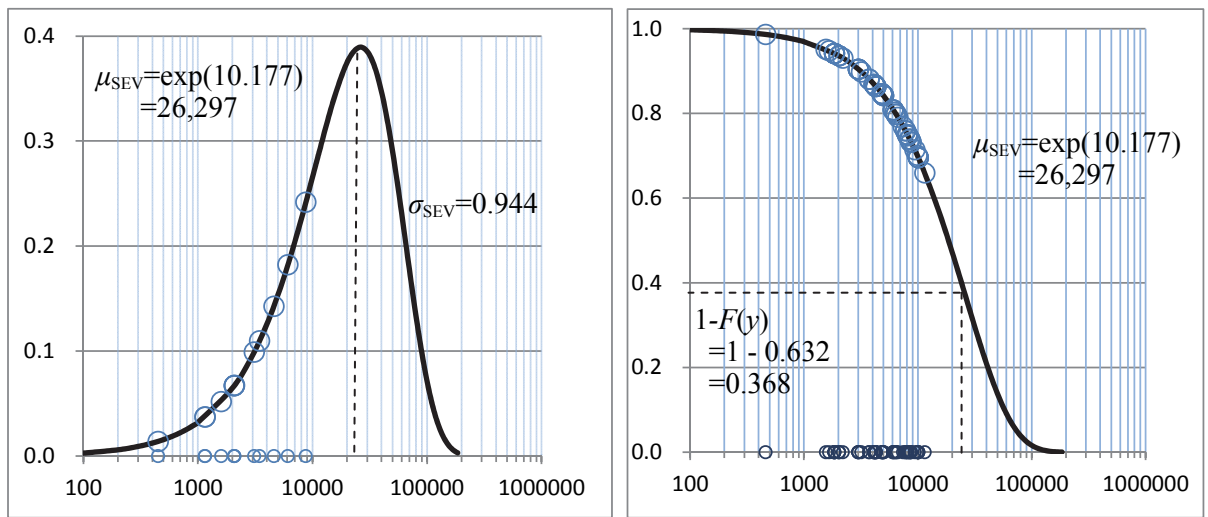


図 6.2 最小極値分布の確率密度関数(左)と生存関数(右)のあてはめ結果

このデータから、これまでの保障期間は 8,000 時間での故障率と信頼区間に対し、50%および 90%のファンが故障するまでの時間と信頼区間の算出を行う。表 6.3 に示すように、8,000 時間での故障率は、式 (6.1) から、

$$F(y) = 1 - \exp\left\{-\exp\left(\frac{\ln(8000) - 10.177}{0.9448}\right)\right\} = 0.247$$

と推定される。50%および 90%のファンが故障する時間は、式 (6.2) で $F(y) = P$ とおいて、 y について解いた式から、

$$y_P = \sigma_{SEV} \ln\{-\ln(1-P)\} + \mu_{SEV}$$

$$y_{P=0.5} = 0.9448 \cdot \ln\{-\ln(1-0.5)\} + 10.177 = 9.831$$

$$y_{P=0.9} = 0.9448 \cdot \ln\{-\ln(1-0.9)\} + 10.177 = 10.965$$

$$t_{P=0.5} = \exp(y_{P=0.5}) = 18,600 \text{ 時間}$$

$$t_{P=0.9} = \exp(y_{P=0.9}) = 57,823 \text{ 時間}$$

と推定される。

表 6.3 時間ごとの故障率と故障までの時間の推定

t	y	$F(y)=P$	P	y_P	t_P
8000	8.987	0.247	0.100	8.051	3137
80000	11.290	0.961	0.500	9.831	18600
$F(t) = 1 - \exp\left\{-\exp\left(\frac{y-\mu}{\sigma}\right)\right\}$			0.900	10.965	57823
			0.950	11.214	74144
			0.975	11.410	90255
			$y_P = \sigma \cdot \ln\{-\ln(1-P)\} + \mu$		

6.3 パラメータの分散共分散行列

これまで、Excel を用いた計算例では、推定されたパラメータの分散については、示してこなかった。JMP を用いた結果の表示には、標準誤差および 95%信頼区間が示されている。算術平均の標準誤差は、偏差平方和を自由度 $n-1$ で除した不偏分散をデータ数 n でさらに除して平方根をとって推定している。

ディーゼル発電機のファンの 12 件の故障データは、450, 1150, 1150, 1600, 2070, 2070, 2080, 3100, 3450, 4600, 6100, 8750 時間であった。最小 2 乗法による平均は $\hat{\mu} = 3047$ 、不偏分散は 5,754,693、標準偏差が 2399、標準誤差は 693 と計算される。表 3.1 に最尤法を適用した推定結果を示した。この結果に、最尤法で得られるパラメータに関する分散共分散行列を別途求めた結果も表 6.4 に合わせて示す。

最尤法によるパラメータの推定は、最小 2 乗法の場合と全く異なる。最小 2 乗法による標準偏差 $\hat{\sigma} = 2398.89$ 、平均の標準偏差 $\hat{SE} = 692.50$ は、平均値を推定してから別途計算して求めているのに対し、最尤法では、標準偏差 $\hat{\sigma} = 2296.77$ 、平均の標準偏差 $\hat{SE} = 663.02$ を同時に推定している。最小 2 乗法と最尤法の推定値の関係は、

$$\hat{\sigma} = \sqrt{n \times \hat{\sigma}^2 / (n-1)} = \sqrt{12 \times 2296.77 / 11} = 2398.89$$

$$\hat{SE} = \sqrt{n} \times \hat{SE} / \sqrt{n-1} = \sqrt{12} \times 663.02 / \sqrt{11} = 692.50$$

と換算できる。この関係は、最小 2 乗法での不偏分散は、偏差平方和を自由度 $n-1$ で除しているのに対し、最尤法での分散はデータ数 n で除した結果に一致することから得られる。

最尤法では、平均値の標準誤差も同時に推定され、さらに標準偏差の標準誤差も表 6.4 に示すように推定されている。この計算の元になったのは、分散共分散行列である。この行列の対角要素の平方根 $\sqrt{439594.6} = 663.02$ 、 $\sqrt{219797.3} = 468.83$ が標準誤差とあることが確認される。

表 6.4 時間ごとの故障率と故障までの時間の推定

最小 2 乗法		最尤法: パラメータ			最尤法: 分散共分散	
統計量	パラメータ	パラメータ	推定値	標準誤差	位置	尺度
平均 $\hat{\mu}$	3047.50	位置 平均 $\tilde{\mu}$	3047.50	663.02	439594.6	-2.05E-11
標準偏差 $\hat{\sigma}$	2398.89	尺度 標準偏差 $\tilde{\sigma}$	2296.77	468.83	-2.05E-11	219797.3
平均の標準偏差	692.50					

最尤法の計算は、JMP の「寿命の一変量」で分布を「正規」とすることで得た。

最尤法は、誤差が正規分布に従う場合、最小極値分布に従う場合、そのような分布に従うかは関係なしに、定義された対数尤度関数に対して同じ手順で解析する。最尤法の計算に、

Excel のソルバーにより対数尤度を最大化することで、最尤推定値は得られるのであるが、分散共分散行列の出力が得られないために、各種の信頼区間の推定ができない。

もちろん、JMP の出力として分散共分散行列を得ることはできる。どのような計算手順で分散共分散行列が計算されているのか、ひも解いてみよう。表 6.2 のディーゼル発電機ファンのデータに対し、最小極値分布の場合について分散共分散行列を Excel で求める。分散共分散行列 Σ は、次式で計算することができる。

$$\Sigma = \begin{bmatrix} -\frac{\partial^2 \ln L}{\partial^2 \mu} & -\frac{\partial^2 \ln L}{\partial \mu \partial \sigma} \\ -\frac{\partial^2 \ln L}{\partial \sigma \partial \mu} & -\frac{\partial^2 \ln L}{\partial^2 \sigma} \end{bmatrix}^{-1} \quad (6.7)$$

この式から、対数尤度関数 $\ln L$ についてパラメータに関して 2 階の偏微分式が必要となる。なお、この式 (6.7) は、最尤法の一般公式であり、ワイブル分布をあてはめる場合には、ワイブルの尤度関数に対して Weibull α 、および Weibull β についての 2 階の偏微分となる。これについては、第 7 章に詳しく述べている。

尤度関数 $\ln L$ は、それぞれのデータについての対数尤度の和

$$\ln L = \sum_{i=1}^n \ln L_i \quad (6.8)$$

である。それぞれの $\ln L_i$ についての対数尤度関数を整理すると、次の式が得られる。

$$\begin{aligned} \ln L_i &= \ln \left[\frac{1}{\sigma} \cdot \exp\left(\frac{y_i - \mu}{\sigma}\right) \right]^{\delta_i} + \ln \left[\exp\left\{-\exp\left(\frac{y_i - \mu}{\sigma}\right)\right\} \right] \\ &= \delta_i \left(-\ln(\sigma) + \frac{y_i - \mu}{\sigma} \right) - \exp\left(\frac{y_i - \mu}{\sigma}\right) \end{aligned} \quad (6.9)$$

この式をパラメータ μ と σ で偏微分すると、以下の結果が得られる。

$$\frac{\partial \ln L_i}{\partial \mu} = \delta_i \left(-\frac{1}{\sigma} \right) + \frac{1}{\sigma} \exp\left(\frac{y_i - \mu}{\sigma}\right) \quad (6.10)$$

$$\frac{\partial \ln L_i}{\partial \sigma} = \delta_i \left(-\frac{1}{\sigma} - \frac{y_i - \mu}{\sigma^2} \right) + \left(\frac{y_i - \mu}{\sigma^2} \right) \cdot \exp\left(\frac{y_i - \mu}{\sigma}\right) \quad (6.11)$$

さらに、2 階の偏微分も次式ように得られる。

$$\frac{\partial^2 \ln L_i}{\partial \mu^2} = -\frac{1}{\sigma^2} \exp\left(\frac{y_i - \mu}{\sigma}\right) \quad (6.12)$$

$$\frac{\partial^2 \ln L_i}{\partial \sigma^2} = \delta_i \left(\frac{1}{\sigma^2} + 2 \frac{y_i - \mu}{\sigma^3} \right) - \left\{ 2 \frac{y_i - \mu}{\sigma^3} + \left(\frac{y_i - \mu}{\sigma^2} \right)^2 \right\} \exp\left(\frac{y_i - \mu}{\sigma}\right) \quad (6.13)$$

$$\frac{\partial^2 \ln L_i}{\partial \mu \partial \sigma} = \delta_i \left(\frac{1}{\sigma^2} \right) - \left\{ \frac{1}{\sigma^2} + \frac{y_i - \mu}{\sigma^3} \right\} \exp \left(\frac{y_i - \mu}{\sigma} \right) \quad (6.14)$$

これらの式を Excel のシートに展開した結果を、表 6.5 に示す。それぞれぞれのデータ y_i についての対数尤度 $\ln L_i$ について μ と σ で偏微分した結果が同じ行に展開されている。これらの偏微分式のうち μ について展開すると

$$\frac{\partial \ln L}{\partial \mu} = \frac{\partial \left(\sum_{i=1}^n \ln L_i \right)}{\partial \mu} = \frac{\partial \ln L_1}{\partial \mu} + \frac{\partial \ln L_2}{\partial \mu} + \dots + \frac{\partial \ln L_n}{\partial \mu} \quad (6.15)$$

となる。全体で 70 件のファンの 2 階の偏微分した結果の合計を 2×2 の偏微分行列 \mathbf{Z} に代入する。さらに、 $\Sigma = (-\mathbf{Z})^{-1}$ により分散共分散行列が計算されている。

表 6.5 最小極値分布の分散共分散行列の推定

							\mathbf{Z}		$\Sigma = (-\mathbf{Z})^{-1}$		SE		
							$\hat{\mu} =$	10.177	-13.4437	21.2070	0.2171	0.0904	0.4659
							$\hat{\sigma} =$	0.9448	21.2070	-50.8951	0.0904	0.0573	0.2394
No	t	y	u	u/σ	u/σ^2	δ	$\partial \mu$	$\partial \sigma$	$\partial \mu \partial \mu$	$\partial \mu \partial \sigma$	$\partial \sigma \partial \sigma$	$\partial \sigma \partial \mu$	
1	450	6.109	-4.306	-4.56	-4.82	1	-1.0442	3.4374	-0.0151	1.1703	-8.6772	1.1703	
2	1150	7.048	-3.313	-3.51	-3.71	1	-1.0199	2.3201	-0.0408	1.2147	-6.4794	1.2147	
:													
11	6100	8.716	-1.547	-1.64	-1.73	1	-0.8330	0.2299	-0.2386	1.2507	-2.1776	1.2507	
12	8750	9.077	-1.165	-1.23	-1.30	1	-0.7282	-0.2103	-0.3496	1.1779	-1.1493	1.1779	
13	460	6.131	-4.282	-4.53	-4.80	0	0.0146	-0.0626	-0.0155	0.0508	-0.1512	0.0508	
14	1560	7.352	-2.990	-3.16	-3.35	0	0.0532	-0.1592	-0.0563	0.1121	-0.1668	0.1121	
:													
69	10100	9.220	-1.013	-1.07	-1.13	0	0.3844	-0.3893	-0.4069	0.0052	0.4068	0.0052	
70	11500	9.350	-0.875	-0.93	-0.98	0	0.4410	-0.3861	-0.4668	-0.0581	0.4596	-0.0581	
計	344440	580.56	-139.55	-147.70	-156.34	12	0.0000	0.0000	-13.4437	21.2070	-50.8951	21.2070	

$$u = (y - \hat{\mu}) / \hat{\sigma}$$

パラメータ $\hat{\mu}_{SEV}$ と $\hat{\sigma}_{SEV}$ の分散共分散行列は、Excel の Minverse () 関数を用いて計算することができる。表 6.6 に表 6.5 から抜粋した分散共分散行列、JMP ですでに計算した表 3.8 の結果を合わせて示す。

表 6.6 分散共分散行列

$\Sigma = (-\mathbf{Z})^{-1}$		パラメータ	位置	尺度
0.2171	0.0904	位置	0.217053	0.090442
0.0904	0.0573	尺度	0.090442	0.057333

表 6.5 (抜粋) Excel の関数による結果

表 3.8 (再掲)

この分散共分散行列から、パラメータ $\hat{\mu}_{\text{SEV}}$ と $\hat{\sigma}_{\text{SEV}}$ の分散と共分散は、

$$\text{Var}(\hat{\mu}_{\text{SEV}}) = 0.2171, \quad \text{Var}(\hat{\sigma}_{\text{SEV}}) = 0.0573, \quad \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) = 0.0904$$

となる。

6.4 パラメータの95%信頼区間

誤差に正規分布を仮定した最小2乗法の場合には、パラメータ $\hat{\mu}$ についての95%信頼区間の計算は、分散共分散行列の対角要素を分散 $\text{Var}(\hat{\mu})$ とし、データ数 n から推定に用いたパラメータ数 p を差し引いた自由度の t 分布 α パーセント点を $t_{(1-\alpha/2)}$ としたときに、

$$\hat{\mu} \pm t_{(1-\alpha/2)} \sqrt{\text{Var}(\hat{\mu})} \quad (6.16)$$

で計算される。最尤法の場合は、自由度を無視して正規分布のパーセント点 $z_{(1-\alpha/2)}$ を用いる。推定された $\hat{\mu}_{\text{SEV}}$ に対して、

$$\hat{\mu}_{\text{SEV}} \pm z_{(1-\alpha/2)} \sqrt{\text{Var}(\hat{\mu}_{\text{SEV}})} = 10.1772 \pm 1.96 \sqrt{0.2171} = (9.2641, 11.0903)$$

が得られる。最尤法で指定された $\hat{\sigma}_{\text{SEV}}$ についても同様に、

$$\hat{\sigma}_{\text{SEV}} \pm z_{(1-\alpha/2)} \sqrt{\text{Var}(\hat{\sigma}_{\text{SEV}})} = 0.9448 \pm 1.96 \sqrt{0.0573} = (0.4755, 1.4141)$$

となる。これは、表 3.7 の JMP での結果に一致する。

表 3.7 (再掲) ワイブル・パラメータの推定

パラメータ	推定値	標準誤差	下側95%	上側95%	規準	
位置	10.1772	0.4659	9.2641	11.0903	(-2)*対数尤度	270.3054
尺度	0.9448	0.2394	0.4755	1.4141	AICc	274.4845
Weibull α	26296.8452	12251.4283	10552.0697	65534.4483	BIC	278.8024
Weibull β	1.0584	0.2683	0.7072	2.1031		

表 3.7 で Weibull $\hat{\alpha}$ の推定値は、26,296 時間、95%信頼区間は (10,552, 65,534) と推定されているが、これは、 $\hat{\alpha} = \exp(\hat{\mu}_{\text{SEV}}) = \exp(10.1772)$ で計算されたもので、95%信頼区間も ($\exp(9.2641)$, $\exp(11.0903)$) で計算した結果となっている。標準誤差は 12,251 となっているが、これは、最小極値分布の分散を用いて、合成分散の一般式から、 $\hat{\alpha} = \exp(\hat{\mu}_{\text{SEV}})$ を $\hat{\mu}_{\text{SEV}}$ で微分すると $\exp(\hat{\mu}_{\text{SEV}})$ が得られるので、

$$\sqrt{\text{Var}(\hat{\alpha})} = \sqrt{[\exp(\hat{\mu}_{\text{SEV}})]^2 \text{Var}(\hat{\mu}_{\text{SEV}})} = \sqrt{26296^2 \times 0.4659^2} = 12251$$

となる。合成分散の一般式については、第 10 章を参照のこと。

Weibull $\hat{\beta}$ は、故障のメカニズムを想定するために使われていて、その 95%信頼区間の推定は、統計的な判断をするために必要である。最小極値分布の $\hat{\sigma}_{SEV}$ から、

$$\text{Weibull } \hat{\beta} = 1 / \hat{\sigma}_{SEV} = 1 / 0.9448 = 1.0584$$

が得られる。Weibull $\hat{\beta}$ の 95%信頼区間は、 $\hat{\sigma}_{SEV}$ の 95%信頼区間の逆数によって

$$(1 / \tilde{\sigma}_{SEV}, 1 / \underline{\sigma}_{SEV}) = (1 / 1.4141, 1 / 0.4755) = (0.7072, 2.1031)$$

が得られる。この結果は、95%信頼区間が 1.0 を含んでいるので、統計的には、偶発故障型と判定される。なお、 $(\underline{\sigma}_{SEV}, \tilde{\sigma}_{SEV})$ は、95%信頼区間の下限と上限を表している。Weibull $\hat{\beta}$ の標準誤差は 0.2683 となっている。これは、 $\hat{\beta} = 1 / \hat{\sigma}_{SEV}$ の関係なので、 $\hat{\sigma}_{SEV}$ で微分すると $-1 / \hat{\sigma}_{SEV}^2$ となり、合成分散の一般式から、

$$\sqrt{\text{Var}(\hat{\beta})} = \sqrt{\frac{\text{Var}(\hat{\sigma}_{SEV})}{(\hat{\sigma}_{SEV}^2)^2}} = \sqrt{\frac{0.2394^2}{0.9447^4}} = 0.2683$$

が計算されている。

6.5 時間 t における故障確率の 95%信頼区間

時間 t における故障確率は、式 (6.1) から $y = \ln(t)$ として、

$$\hat{F}(y) = 1 - \exp\left\{-\exp\left(\frac{y - \hat{\mu}_{SEV}}{\hat{\sigma}_{SEV}}\right)\right\} \quad (6.17)$$

となる。ここで、

$$u = \frac{y - \hat{\mu}_{SEV}}{\hat{\sigma}_{SEV}}$$

とおき、 u の分散 $\text{Var}(u)$ を $\hat{\mu}_{SEV}$ および $\hat{\sigma}_{SEV}$ の分散から合成する必要がある。合成分散の一般式から、合成分散 $\text{Var}(u)$ は、 u を $\hat{\mu}_{SEV}$ および $\hat{\sigma}_{SEV}$ で偏微分し

$$d_1 = \frac{\partial u}{\partial \hat{\mu}_{SEV}} = \frac{-1}{\hat{\sigma}_{SEV}}, \quad d_2 = \frac{\partial u}{\partial \hat{\sigma}_{SEV}} = \frac{y - \hat{\mu}_{SEV}}{\hat{\sigma}_{SEV}^2} = \frac{u}{\hat{\sigma}_{SEV}} \quad (6.18)$$

その結果を、 $\mathbf{d} = [d_1, d_2]^T$ と縦ベクトルとし、分散共分散行列を Σ としたときに、次の 2 次形式で計算することができる。

$$\begin{aligned} \text{Var}(\hat{u}) &= \mathbf{d}^T \Sigma \mathbf{d} \\ &= \begin{bmatrix} -1 / \hat{\sigma}_{SEV} & u / \hat{\sigma}_{SEV} \end{bmatrix} \begin{bmatrix} \text{Var}(\hat{\mu}_{SEV}) & \text{Cov}(\hat{\mu}_{SEV}, \hat{\sigma}_{SEV}) \\ \text{Cov}(\hat{\mu}_{SEV}, \hat{\sigma}_{SEV}) & \text{Var}(\hat{\sigma}_{SEV}) \end{bmatrix} \begin{bmatrix} -1 / \hat{\sigma}_{SEV} \\ u / \hat{\sigma}_{SEV} \end{bmatrix} \\ &= \left(\frac{1}{\hat{\sigma}_{SEV}}\right)^2 \left[\text{Var}(\hat{\mu}_{SEV}) + 2u \text{Cov}(\hat{\mu}_{SEV}, \hat{\sigma}_{SEV}) + u^2 \text{Var}(\hat{\sigma}_{SEV}) \right] \end{aligned} \quad (6.19)$$

となるので、 u の 95%信頼区間は、

$$\underline{u} = u - 1.96\sqrt{Var(u)}, \quad \tilde{u} = u + 1.96\sqrt{Var(u)}$$

となる。この下限と上限から、故障率は

$$F(\underline{y}) = 1 - \exp\{-\exp(\underline{u})\}, \quad F(\tilde{y}) = 1 - \exp\{-\exp(\tilde{u})\} \quad (6.20)$$

で求める。

時間 $t=8,000$ の場合には,

$$u = \frac{y - \hat{\mu}_{SEV}}{\hat{\sigma}_{SEV}} = \frac{\ln(8000) - 10.1772}{0.9448} = -1.2596$$

$$Var(\hat{u}) = \left(\frac{1}{0.9448}\right)^2 \{0.2171 + 2u \times 0.0904 + u^2 \times 0.0573\} = 0.0898$$

$$\underline{u} = -1.2596 - 1.96\sqrt{0.0898} = -1.8470$$

$$\tilde{u} = -1.2596 + 1.96\sqrt{0.0898} = -0.6721$$

$$F(\underline{y}) = 1 - \exp\{-\exp(-1.8470)\} = 0.1459$$

$$F(\tilde{y}) = 1 - \exp\{-\exp(-0.6721)\} = 0.3999$$

と計算できる。Excel で故障率の 95%信頼区間を計算した結果を 表 6.7 に示す。

表 6.7 時間 t での故障率の 95%信頼区間の計算

	$\hat{\mu}_{SEV} =$	10.1772		$\Sigma = (-Z)^{-1} =$	0.2171	0.0904		
	$\hat{\sigma}_{SEV} =$	0.9448			0.0904	0.0573		
t	$y = \ln t$	u	$F(y)$	$Var(u)$	L95%(u)	U95%(u)	L95%(F)	U95%(F)
10	2.3026	-8.3349	0.0002	3.0163	-11.7389	-4.9308	0.0000	0.0072
50	3.9120	-6.6314	0.0013	1.7239	-9.2048	-4.0579	0.0001	0.0171
100	4.6052	-5.8977	0.0027	1.2822	-8.1171	-3.6783	0.0003	0.0249
1,000	6.9078	-3.4605	0.0309	0.3111	-4.5537	-2.3673	0.0105	0.0895
2,000	7.6009	-2.7269	0.0633	0.1682	-3.5307	-1.9231	0.0289	0.1360
4,000	8.2940	-1.9932	0.1274	0.0944	-2.5955	-1.3909	0.0719	0.2203
8,000	8.9872	-1.2596	0.2471	0.0898	-1.8470	-0.6721	0.1459	0.3999
15,000	9.6158	-0.5942	0.4242	0.1454	-1.3417	0.1532	0.2300	0.6883
26,297	10.1772	0.0000	0.6321	0.2432	-0.9665	0.9665	0.3164	0.9278
50,000	10.8198	0.6801	0.8611	0.4107	-0.5760	1.9362	0.4300	0.9990
80,000	11.2898	1.1776	0.9611	0.5709	-0.3033	2.6585	0.5221	1.0000
100,000	11.5129	1.4138	0.9836	0.6580	-0.1762	3.0037	0.5676	1.0000

時間 $t=8,000$ と $t=80,000$ の場合に JMP で計算した結果は、表 3.9 (再掲) に示した結果と一致する。

表 3.9 (再掲) 故障率の 95%信頼区間 (JMP 寿命の一変量)

t	故障確率	故障確率 下側95% (Wald)	故障確率 上側95% (Wald)
8000	0.2471	0.1459	0.3999
80000	0.9611	0.5221	1.0000

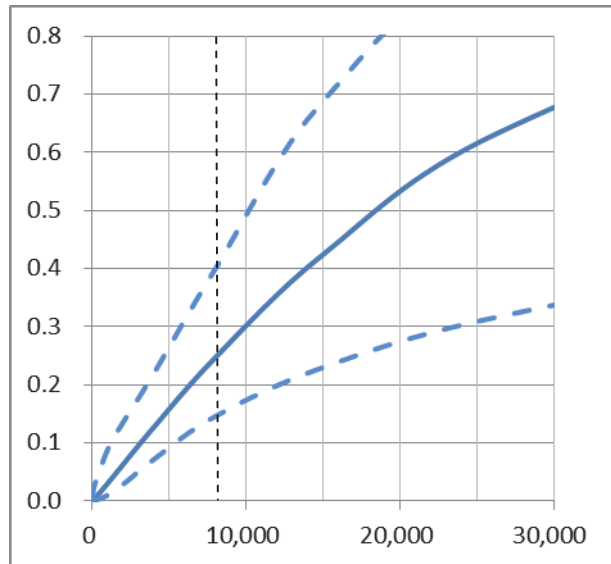


図 6.3 故障率の 95%信頼区間 (図 3.12 の JMP の結果に対応)

表 6.7 の計算結果を Excel でグラフ化した結果を Excel の平滑化の機能を用いて 図 6.3 に示す. 信頼区間の幅が裾広がりになっているのは, 故障データが 8,750 時間までしかないため外装となったためである.

ワイブル確率プロット上での故障率の 95%信頼区間は, 表 6.7 の計算結果を, 故障率については, Y 軸を

$$W = \ln \left\{ \ln \left[\frac{1}{1-F(y)} \right] \right\}, \quad \tilde{W} = \ln \left\{ \ln \left[\frac{1}{1-F(\tilde{y})} \right] \right\}, \quad \tilde{W} = \ln \left\{ \ln \left[\frac{1}{1-F(\tilde{y})} \right] \right\} \quad (6.21)$$

で換算し, X 軸は時間 t について対数目盛とすることにより得られる. 8,000 時間について計算すると,

$$W = \ln \left\{ \ln \left[\frac{1}{1-0.2471} \right] \right\} = -1.2596$$

$$\tilde{W} = \ln \left\{ \ln \left[\frac{1}{1-0.1459} \right] \right\} = -1.8470$$

$$\tilde{W} = \ln \left\{ \ln \left[\frac{1}{1-0.3999} \right] \right\} = -0.6721$$

となる. 表 6.7 のデータに対し, 2重対数変換してワイブル確率目盛とした結果を 図 6.4 に示す.

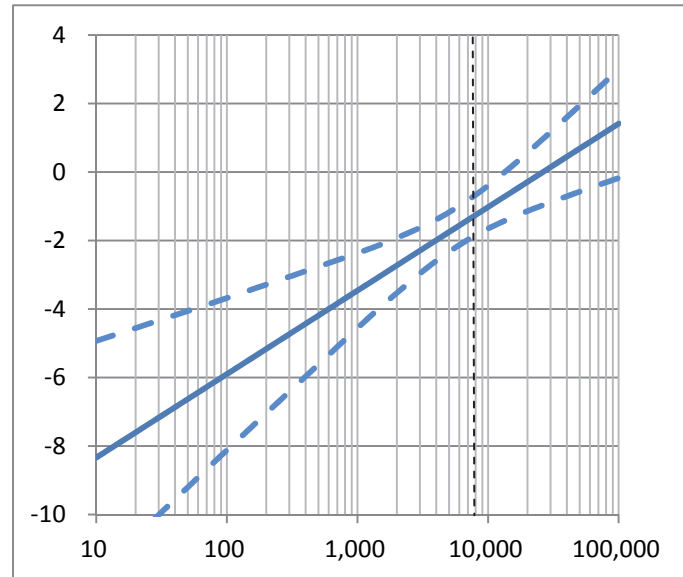


図 6.4 ワイブル確率プロット上での 95%信頼区間 (図 3.13 の JMP の結果に対応)

6.6 故障率に対する時間 t の 95%信頼区間

最小極値分布のパーセント点は、式 (6.2) から $y_p = \hat{\mu}_{SEV} + \ln[-\ln(1-P)]\hat{\sigma}_{SEV}$ として求められる。この y_p を $\hat{\mu}_{SEV}$ および $\hat{\sigma}_{SEV}$ で偏微分すると、

$$\frac{\partial y_p}{\partial \hat{\mu}_{SEV}} = 1, \quad \frac{\partial y_p}{\partial \hat{\sigma}_{SEV}} = \ln[-\ln(1-P)] \quad (6.22)$$

が得られる、 $u_p = \ln\{-\ln(1-P)\}$ とおくと合成分散の一般式から、

$$Var(y_p) = Var(\hat{\mu}) + 2u_p Cov(\hat{\mu}, \hat{\sigma}) + u_p^2 Var(\hat{\sigma})$$

となるので、 y_p の 95%信頼区間の下限と上限は、

$$\underline{y}_p = y_p - 1.96\sqrt{Var(y_p)}, \quad \tilde{y}_p = y_p + 1.96\sqrt{Var(y_p)} \quad (6.23)$$

で求めることができる。これは、時間 t の対数に対するものなので、指数をとって元のスケールのパーセント点は、

$$\underline{t}_p = \exp(\underline{y}_p), \quad \tilde{t}_p = \exp(\tilde{y}_p) \quad (6.24)$$

となる。50%のファンが故障する時間は、

$$\begin{aligned} y_{P=0.5} &= \hat{\mu}_{SEV} + \ln[-\ln(1-0.5)]\hat{\sigma}_{SEV} \\ &= 10.177 + 1.0972 \times (-0.3665) = 9.8309 \end{aligned}$$

$$t_{P=0.5} = \exp(y_{P=0.5}) = \exp(9.8309) = 18,600$$

と推定され、分散は

$$Var(y_{P=0.5}) = 0.2171 + 2 \times (-0.3665) \times 0.0904 + (-0.3665)^2 \times 0.0573 = 0.4845$$

で、95%信頼区間は、

$$\begin{aligned} \underline{y}_{P=0.5} &= y_{P=0.5} - 1.96\sqrt{Var(y_{P=0.5})} = 9.8309 - 1.96 \times \sqrt{0.1585} = 9.0507, \\ \tilde{y}_{P=0.5} &= y_{P=0.5} + 1.96\sqrt{Var(y_{P=0.5})} = 10.6111 \end{aligned}$$

表 6.8 最小極値分布のパーセント点

	Z		$\Sigma=(-Z)^{-1}$		SE			
$\mu =$	10.177	-13.4437	21.2070	0.2171	0.0904	0.4659		
$\sigma =$	0.9448	21.2070	-50.8951	0.0904	0.0573	0.2394		
P	u_P	y_P	$Var(y_P)$	$L95\%$	$U95\%$	t_P	$L95\%$	$U95\%$
0.01	-4.6001	5.8311	0.5982	4.3151	7.3470	341	75	1,552
0.05	-2.9702	7.3710	0.1856	6.5266	8.2154	1,589	683	3,697
0.1	-2.2504	8.0511	0.1003	7.4302	8.6720	3,137	1,686	5,837
0.2	-1.4999	8.7601	0.0747	8.2243	9.2959	6,375	3,730	10,893
0.5	-0.3665	9.8309	0.1585	9.0507	10.6111	18,600	8,525	40,585
0.7	0.1856	10.3526	0.2526	9.3675	11.3377	31,338	11,702	83,925
0.9	0.8340	10.9652	0.4078	9.7135	12.2168	57,825	16,540	202,161
0.95	1.0972	11.2138	0.4845	9.8495	12.5781	74,147	18,948	290,145
0.975	1.3053	11.4104	0.5509	9.9557	12.8652	90,260	21,073	386,602

で、指数を取り、

$$t_{p=0.5} = \exp(9.0507) = 8,525 \text{ 時間}, \quad \tilde{t}_{p=0.5} = \exp(10.6111) = 40,585 \text{ 時間}$$

が得られる。この計算を Excel シートで行った結果を 表 6.8 に示す。

この計算結果に基づいて、対数目盛と実目盛とした場合の 95%信頼区間について 図 6.5 に示す。

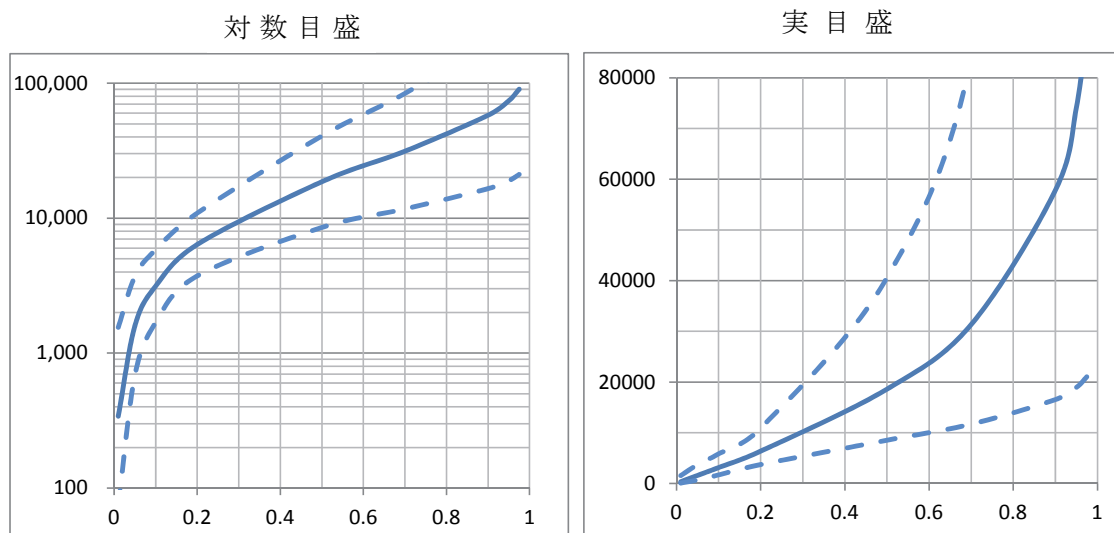


図 6.5 パーセント点での 95%信頼区間 (図 3.14 の JMP の図に対応)

7. Newton-Raphson 法による尤度の最大化

7.1 Weibull パラメータの分散共分散行列の直接推定

各種の信頼区間の計算を第 6 章では、時間 t の対数を取った最小極値分布での分散共分散を Excel で計算した結果に基づき求めた。奥野監訳本では、Weibull パラメータに関する分散共分散の推定結果から、最小極値分布での分散共分散を推定している。歴史的にはワイブル分布を前提にした計算法が先行したのであるが、対数時間による方が計算の見通しがよく、JMP では最小極値分布を主体にしている。

第 6 章では、分散共分散の推定に先立って、Excel のソルバーで最尤解を求めてから、2 階の偏微分の計算を行ってきた。Excel のソルバーを使用することは、最尤法の逐次計算プロセスを棚上げにして、とりあえず結果を出すことを優先したためである。

最尤法による対数尤度を最大化するための逐次計算には、Newton-Raphson 法がよく使われている。本章では、Excel のソルバーに頼ることなく、Excel の基本機能だけを用いて、誰にでも Newton-Raphson 法の逐次計算が体得できることを目的としている。また、奥野監訳本との対比ができるように、最小極値分布ではなく、ワイブル分布を用いることにした。

第 3 章では、ワイブル分布の確率密度関数、累積分布関数には Excel の関数を用いて、それらの対数を尤度としていた。残念ながら Excel には、自動微分の機能がないために、自ら微分式を導出する必要がある。奥野監訳本には、全データを込にした対数尤度関数の偏微分した式が示されているが、Excel での計算では、それぞれのデータに対する偏微分式が必要となる。

式 (3.10) から、それぞれのデータについての対数尤度関数は次の式となる。

$$\begin{aligned} \ln L_i &= \ln f(t_i)^{\delta_i} + \ln \{1 - F(t_i)\}^{1-\delta_i} \\ &= \ln \left[\frac{\beta}{\alpha^\beta} t^{\beta-1} \exp \left\{ -\left(\frac{t_i}{\alpha} \right)^\beta \right\} \right]^{\delta_i} + \ln \left[\exp \left\{ -\left(\frac{t_i}{\alpha} \right)^\beta \right\} \right]^{1-\delta_i} \\ &= \delta_i \cdot [\ln \beta - \beta \ln \alpha + (\beta - 1) \ln(t_i)] - \left(\frac{t_i}{\alpha} \right)^\beta \end{aligned} \quad (7.1)$$

この式を Weibull α と Weibull β で偏微分すると、以下の結果が得られる。

$$\begin{aligned}\frac{\partial \ln L_i}{\partial \alpha} &= \delta_i \cdot \left(-\frac{\beta}{\alpha} \right) - \left(-\frac{t_i}{\alpha^2} \right) \beta \left(\frac{t_i}{\alpha} \right)^{\beta-1} \\ &= \delta_i \cdot \left(-\frac{\beta}{\alpha} \right) + \frac{\beta}{\alpha} \left(\frac{t_i}{\alpha} \right)^{\beta}\end{aligned}\quad (7.2)$$

$$\begin{aligned}\frac{\partial \ln L_i}{\partial \beta} &= \delta_i \cdot \left(\frac{1}{\beta} + \ln \frac{t_i}{\alpha} \right) - \exp \left(\beta \ln \left(\frac{t_i}{\alpha} \right) \right) \\ &= \delta_i \cdot \left(\frac{1}{\beta} + \ln \frac{t_i}{\alpha} \right) - \ln \left(\frac{t_i}{\alpha} \right) \cdot \left(\frac{t_i}{\alpha} \right)^{\beta}\end{aligned}\quad (7.3)$$

さらに、Weibull α と Weibull β についての 2 階の偏微分は、次のように計算できる。

$$\frac{\partial^2 \ln L_i}{\partial^2 \alpha} = \delta_i \cdot \left(\frac{\beta}{\alpha^2} \right) - \frac{\beta + \beta^2}{\alpha^2} \left(\frac{t_i}{\alpha} \right)^{\beta} \quad (7.4)$$

$$\frac{\partial^2 \ln L_i}{\partial \alpha \partial \beta} = \delta_i \cdot \left(-\frac{1}{\alpha} \right) + \left\{ \frac{1}{\alpha} + \frac{\beta}{\alpha} \log \left(\frac{t_i}{\alpha} \right) \right\} \cdot \left(\frac{t_i}{\alpha} \right)^{\beta} \quad (7.5)$$

$$\frac{\partial^2 \ln L_i}{\partial^2 \beta} = \delta_i \cdot \left(-\frac{1}{\beta^2} \right) - \left\{ \log \left(\frac{t_i}{\alpha} \right) \right\}^2 \left(\frac{t_i}{\alpha} \right)^{\beta} \quad (7.6)$$

これらを、2 階の偏微分行列 \mathbf{Z} に代入する。

$$\mathbf{Z} = \begin{bmatrix} \frac{\partial^2 \ln L}{\partial^2 \alpha} & \frac{\partial^2 \ln L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ln L}{\partial \beta \partial \alpha} & \frac{\partial^2 \ln L}{\partial^2 \beta} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \left(\sum_{i=1}^n \ln L_i \right)}{\partial^2 \alpha} & \frac{\partial^2 \left(\sum_{i=1}^n \ln L_i \right)}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \left(\sum_{i=1}^n \ln L_i \right)}{\partial \beta \partial \alpha} & \frac{\partial^2 \left(\sum_{i=1}^n \ln L_i \right)}{\partial^2 \beta} \end{bmatrix} \quad (7.7)$$

$$\boldsymbol{\Sigma} = (-\mathbf{Z})^{-1} \quad (7.8)$$

全体の対数尤度 $\ln L$ は、それぞれのデータについての対数尤度の和

$$\ln L = \sum_{i=1}^n \ln L_i \quad (7.9)$$

である。Excel のソルバーで $\ln L$ を最大化するように、Weibull $\hat{\alpha}$ 、Weibull $\hat{\beta}$ を変化させると分散共分散行列 $\boldsymbol{\Sigma}$ が計算される。これは、Newton-Raphson 法での逐次計算を実施するための準備となっている。

表 7.1 に表 3.3 に示した、打ち切りを含むワイブル分布の最尤法によるあてはめシートを元に、1 階の偏微分および 2 階の偏微分式を付け加えた結果を示す。さらに 1 行目から 70 行目までの和が、下段に計算されている。この和を、表の頭に付けてある 1 階にの偏微分ベクトル、2 階の偏微分係数行列にまとめてある。

表 7.1 ワイブル分布の分散共分散行列の推定

			元の		新たな		1階の	2階の偏微分	負の逆行列		
			パラメータ	変化量	パラメータ		偏微分	Z	(-Z) ⁻¹		
Weible α =	26.2968	0.0000	26.2968			0.0000	-0.0194	-0.7198	150.0975	-2.6645	
Weible β =	1.0584	0.0000	1.0584			0.0000	-0.7198	-40.5511	-2.6645	0.0720	
			ln L =	-52.26	-2 ln L =	104.52					
No	t ₁₀₀₀ +	δ	f(t)	1-F(t)	Li	ln Li	∂α	∂β	∂α ∂α	∂α ∂β	∂β ∂β
1	0.450	1	0.031	0.987	0.031	-3.464	-0.040	-3.068	0.0015	-0.0397	-1.1159
2	0.460	+	0	0.031	0.986	0.986	-0.014	0.001	0.056	0.0000	-0.0017
3	1.150	1	0.032	0.964	0.032	-3.432	-0.039	-2.071	0.0014	-0.0412	-1.2494
4	1.150	1	0.032	0.964	0.032	-3.432	-0.039	-2.071	0.0014	-0.0412	-1.2494
:											
69	10.100	+	0	0.026	0.695	0.695	-0.363	0.015	0.348	-0.0011	-0.0002
70	11.500	+	0	0.025	0.659	0.659	-0.417	0.017	0.345	-0.0013	0.0020
			計	-52.260	0.000	0.000	-0.019	-0.720	-40.551		

パラメータ α と β の分散共分散行列は、Excel の Minverse () 関数を用いて計算することができ、結果を次に示す。

$$Z = \begin{bmatrix} -0.0194 & -0.7198 \\ -0.7198 & -40.5510 \end{bmatrix}, \quad \Sigma = (-Z)^{-1} = \begin{bmatrix} 150.0975 & -2.6645 \\ -2.6645 & 0.0720 \end{bmatrix}$$

この結果から Weibull $\hat{\alpha}$ と $\hat{\beta}$ の分散と共分散は、

$$Var(\hat{\alpha}) = 150.927, \quad Var(\hat{\beta}) = 0.0720, \quad Cov(\hat{\alpha}, \hat{\beta}) = -2.6645$$

となる。

なお、図 2.3 で引用した奥野監訳本のワイブル確率紙へのプロット図 (309 ページ) に引き続き 310 ページにすでにコンピュータ出力として掲載されている。原著は 1982 年の出版であり、30 年以上前にワイブル分布の分散共分散行列の推定方法は、広く知られていたことと推測される。

表 7.2 奥野監訳本にみるワイブル分布の分散共分散行列

ESTIMATED COVARIANCE MATRIX OF PARAMETER ESTIMATES:		
	SCALE	SHAPE
SCALE	150.10	-2.6645
SHAPE	-2.6645	0.71958E-01

7.2 Newton-Raphson 法によるワイブル分布での最尤法

これまで、ワイブル分布の対数尤度を最大化するために、Excel のソルバーを用いてきた。また、各種の 95%信頼区間を算出するために、パラメータの分散共分散行列 Σ を求めることが必須であった。そのために、対数尤度関数の 2 階の偏微分式を導出し、Excel でそれぞれ行ごとに偏微分式の計算を行い、 n 個のデータについての合計から、分散共分散行列 Σ を計算した。

対数尤度を最大化するためにソルバーを用いることにより、煩雑な数値計算の過程を省略することができたので、最尤解を求めるために統計ソフトの内部で使われている Newton-Raphson 法については触れずにきた。逐次的にパラメータを変化させて尤度を最大化すればよいと簡単に言ってきたが、なかなか大変な作業である。このために、コンピュータ内部で対数尤度を最大化する数値計算の方法が開発されてきた。その一つが Newton-Raphson 法であり、統計ソフトのユーザにとってはブラック・ボックスとなっている。

Excel の行列関数を用いることにより、Newton-Raphson 法の計算過程を実際に見ながら逐次的に手作業が行える。もちろん、この過程にもソルバーを使い瞬時に結果を得ることもできるが、手作業での Newton-Raphson 法を体験してもらいたい。

すでにパラメータの分散共分散行列 Σ の計算のための計算シートを示したが、ここには、すでに Newton-Raphson 法で必要になる計算式がすでに埋め込まれている。

最初に、適当な Weibull パラメータ α_0 と Weibull パラメータ β_0 を設定する。対数尤度 $\ln L$ をパラメータ α と β で偏微分した計算は、表 7.1 にすでに組み込まれていて、それぞれの t_i についての偏微係数が計算されている。その合計を、次のベクトルとして Excel シート上に配置する。

$$\begin{bmatrix} \alpha \text{で微分} \\ \beta \text{で微分} \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L}{\partial \alpha} \\ \frac{\partial \ln L}{\partial \beta} \end{bmatrix} = \begin{bmatrix} \frac{\partial \left(\sum_{i=1}^n \ln L_i \right)}{\partial \alpha} \\ \frac{\partial \left(\sum_{i=1}^n \ln L_i \right)}{\partial \beta} \end{bmatrix} \quad (7.10)$$

2 階の偏微分の結果を 2×2 の行列にしたものを \mathbf{Z} とし、その負の \mathbf{Z} の逆行列を $\Sigma = (-\mathbf{Z})^{-1}$ とした。

$$\Sigma = \begin{bmatrix} -\frac{\partial^2 \ln L}{\partial \alpha^2} & -\frac{\partial^2 \ln L}{\partial \alpha \partial \beta} \\ -\frac{\partial^2 \ln L}{\partial \beta \partial \alpha} & -\frac{\partial^2 \ln L}{\partial \beta^2} \end{bmatrix}^{-1} \quad (7.11)$$

これらの積から、パラメータの変化量を求める。

$$\begin{bmatrix} \alpha \text{ の変化量} \\ \beta \text{ の変化量} \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 \ln L}{\partial^2 \alpha} & -\frac{\partial^2 \ln L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ln L}{\partial \beta \partial \alpha} & -\frac{\partial^2 \ln L}{\partial^2 \beta} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \ln L}{\partial \alpha} \\ \frac{\partial \ln L}{\partial \beta} \end{bmatrix} \quad (7.12)$$

この変化量を元のパラメータに加える。

$$\begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} \alpha \text{ の変化量} \\ \beta \text{ の変化量} \end{bmatrix} \quad (7.13)$$

これで求めた新たなパラメータを元のパラメータに代入して、計算を繰り返す。

$$\begin{bmatrix} \alpha'_0 \\ \beta'_0 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} \quad (7.14)$$

以上、手順を示すとややこしいが、Excelには自動計算の機能が備わっており、新たなパラメータのセルの値を元のパラメータのセルに入力(コピー&値のみペースト)とし直すことで、スムーズな逐次計算が行える。

表 7.3 変化量が 0 なので解が求まっている状態

	元の パラメータ	変化量	新たな パラメータ	1階の 偏微分	2階の偏微分 Z	負の逆行列 (-Z) ⁻¹
Weible α =	26.2968	0.0000	26.2968	0.0000	-0.0194 -0.7198	150.0975 -2.6645
Weible β =	1.0584	0.0000	1.0584	0.0000	-0.7198 -40.5511	-2.6645 0.0720

表 7.4 変化量が 0 ではないので解が求まっていない状態

	元の パラメータ	変化量	新たな パラメータ	1階の 偏微分	2階の偏微分 Z	負の逆行列 (-Z) ⁻¹
Weible α =	26.0000	0.5035	26.5035	0.0480	-0.0214 -0.7137	98.1725 -1.5485
Weible β =	1.0000	0.0521	1.0521	2.7172	-0.7137 -45.2442	-1.5485 0.0465



表 7.5 逐次的な変化過程

	元の パラメータ	変化量	新たな パラメータ	元の パラメータ	変化量	新たな パラメータ	元の パラメータ	変化量	新たな パラメータ
Weible α =	26.5035	-0.2061	26.2974	26.2974	-0.0006	26.2968	26.2968	0.0000	26.2968
Weible β =	1.0521	0.0063	1.0584	1.0584	0.0001	1.0584	1.0584	0.0000	1.0584

1 回目の置き換え

2 回目の置き換え

3 回目で変化量が 0

7.3 Newton-Raphson 法による正規分布での最尤法

前節では、打ち切りデータを含むデータに対していきなり Newton-Raphson 法を用いた最尤法の推定過程を示した。これは、読者にとって段階的な学習を妨げかねない。そこで、正規分布を使った場合について、Newton-Raphson 法による最尤法を例示しよう。

図 3.4 で故障したファン 12 件に対して正規分布のあてはめでソルバーを用いてパラメータ推定を行った。一般的には、データの平均を求め、偏差平方和を自由度で割って不偏分散を求め、その平方根から標準偏差 SD 、平均の標準誤差 SE は、 $SE = SD / \sqrt{n}$ で求めている。

この問題に対して最尤法を用いて推定する実務上の必要性は全くない。しかしながら、このような周知の問題に対し Newton-Raphson 法によって、最尤解を求める過程を示すことは、読者に途切れのない段階的な学習を願っているからである。

正規分布は、パラメータを μ 、形状パラメータを σ としその確率密度関数は、

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \quad (7.15)$$

として知られている。累積分布関数 $F(t)$ は、確率密度関数の積分として定義されているが、故障データのみなので、幸い累積分布関数は必要としない。前節と同様に確率密度関数 $f(t)$ をパラメータによって偏微分する。対数を取ると。

$$\ln f(t) = -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(t-\mu)^2}{2\sigma^2} \quad (7.16)$$

なので、 $\ln f(t)$ を μ 、 (σ^2) で偏微分して、次の結果を得る。

$$\frac{\partial \ln f(t)}{\partial \mu} = \frac{(t-\mu)}{(\sigma^2)} \quad (7.17)$$

$$\frac{\partial \ln f(t)}{\partial (\sigma^2)} = -\frac{1}{2(\sigma^2)} + \frac{(t-\mu)^2}{2(\sigma^2)^2} \quad (7.18)$$

さらに 2 階の偏微分を計算する。

$$\frac{\partial \ln f(t)}{\partial \mu \partial \mu} = -\frac{1}{(\sigma^2)} \quad (7.19)$$

$$\frac{\partial \ln f(t)}{\partial \mu \partial (\sigma^2)} = \frac{(t-\mu)}{(\sigma^2)^2} \quad (7.20)$$

$$\frac{\partial \ln f(t)}{\partial (\sigma^2) \partial (\sigma^2)} = \frac{1}{2(\sigma^2)^2} - \frac{(t-\mu)^2}{2(\sigma^2)^3} \quad (7.21)$$

これらの偏微分の結果を Excel のシートに落とし込む。表 7.6 には、ファンが故障した 12

件のデータのみに対して、ソルバーによって $\ln L$ を最大化する $\tilde{\mu} = 3.0475$ (千), $\tilde{\sigma}^2 = 5.2751$ が示されている. 図 3.4 では $\tilde{\sigma} = 2296.8$ 時間となっている. これは, $\sqrt{5.2751 \times 1000^2} = 2296.8$ で換算される.

表 7.6 正規分布のニュートン・ラフソン法によるパラメータ推定結果

正規分布		元の		新たな		1階の	2階の偏微分	負の逆行列			
Newton		パラメータ	変化量	パラメータ		偏微分	Z	$(-Z)^{-1}$			
-Raphson法	$\tilde{\mu} =$	3.0475	0.0000	3.0475		0.0000	-2.2748	0.0000	0.4396	0.0000	
	$\tilde{\sigma}^2 =$	5.2751	0.0000	5.2751		0.0000	0.0000	-0.2156	0.0000	4.6378	
				$\ln L =$	-27.0053	$-2 \ln L =$	54.01				
No	$t_{1000} + \delta$	$f(t)$	$S(t)$	Li	$\ln Li$	$\partial \mu$	$\partial \sigma^2$	$\partial \mu \partial \mu$	$\partial \mu \partial \sigma^2$	$\partial \sigma^2 \partial \sigma^2$	
1	0.450	1	0.0916	0.932	0.0916	-2.3900	-0.4924	0.0264	-0.1896	-0.0933	-0.0280
3	1.150	1	0.1235	0.861	0.1235	-2.0917	-0.3597	-0.0301	-0.1896	-0.0682	-0.0066
4	1.150	1	0.1235	0.861	0.1235	-2.0917	-0.3597	-0.0301	-0.1896	-0.0682	-0.0066
6	1.600	1	0.1424	0.796	0.1424	-1.9490	-0.2744	-0.0571	-0.1896	-0.0520	0.0037
16	2.070	1	0.1587	0.712	0.1587	-1.8410	-0.1853	-0.0776	-0.1896	-0.0351	0.0115
17	2.070	1	0.1587	0.712	0.1587	-1.8410	-0.1853	-0.0776	-0.1896	-0.0351	0.0115
18	2.080	1	0.1590	0.710	0.1590	-1.8392	-0.1834	-0.0780	-0.1896	-0.0348	0.0116
24	3.100	1	0.1737	0.488	0.1737	-1.7507	0.0100	-0.0947	-0.1896	0.0019	0.0179
26	3.450	1	0.1711	0.409	0.1711	-1.7658	0.0763	-0.0919	-0.1896	0.0145	0.0169
37	4.600	1	0.1382	0.187	0.1382	-1.9789	0.2943	-0.0515	-0.1896	0.0558	0.0015
46	6.100	1	0.0718	0.040	0.0718	-2.6336	0.5787	0.0726	-0.1896	0.1097	-0.0455
63	8.750	1	0.0080	0.001	0.008	-4.8327	1.0810	0.4895	-0.1896	0.2049	-0.2036
					計	0.0000	0.0000	-2.2748	0.0000	-0.2156	

パラメータの初期値を $\hat{\mu}_0 = 3.0$, $\hat{\sigma}_0^2 = 5.0$ として, 逐次計算を行ってみよう. 結果を表 7.7 に示す. 新たなパラメータには, 元のパラメータの (3, 5) に変化量が (0.0499, 0.2540) 加えられて (3.0499, 5.2540) が示されている. これをコピーして, 元のパラメータの欄に「値のみ」をペーストする.

表 7.7 変化量が 0 ではないので解が求まっていない状態

	元の		新たな		1階の	2階の偏微分	負の逆行列		
	パラメータ	変化量	パラメータ		偏微分	Z	$(-Z)^{-1}$		
$\tilde{\mu} =$	3.0000	0.0499	3.0499		0.1140	-2.4000	0.0228	0.4170	0.0357
$\tilde{\sigma}^2 =$	5.0000	0.2540	5.2540		0.0666	0.0228	-0.2666	0.0357	3.7536

結果は, 表 7.8 に示すように変化量が (-0.0024, 0.0210) となり, あら新たなパラメータは (3.0475, 5.2754) となっている. これをまたコピーして, 元のパラメータの欄に「値のみ」をペーストすると, 表 7.8 (中) ように変化量が (0.0000, 0.0002) を加えて新たなパラメータとして (3.0475, 5.2751) が計算されている. これを元のパラメータに「値のみ」をペーストすると変化量は, (0.0000, 0.0000) となり, 最尤解が求まったことになる.

表 7.8 逐次的な変化過程

	元の		新たな		元の		新たな		元の		新たな	
	パラメータ	変化量	パラメータ	変化量	パラメータ	変化量	パラメータ	変化量	パラメータ	変化量	パラメータ	変化量
$\mu \sim =$	3.0499	-0.0024	3.0475		3.0475	0.0000	3.0475		3.0475	0.0000	3.0475	
$\sigma^2 \sim =$	5.2540	0.0210	5.2750		5.2750	0.0002	5.2751		5.2751	0.0000	5.2751	
	↑				↑							
	1回目の置き換え				2回目の置き換え				3回目で変化量が0で収束			

表 7.9 に表 7.6 に対応する最尤法による正規分布のパラメータ推定を JMP で行った結果を示す。Excel のニュートン・ラフソン法で行った最尤解に一致していることが確かめられる。

表 7.9 最尤法による正規分布のパラメータ推定 (JMP 寿命の一変量)

	パラメータ	推定値	標準誤差	下側95%	上側95%	規準	
位置 $\mu \sim$	位置	3.0475	0.6630	1.7480	4.3470	(-2)*対数尤度	54.0106
尺度 $\sigma \sim$	尺度	2.2968	0.4688	1.3779	3.2156	AICc	59.3439
	尺度 $\hat{\sigma} = \sqrt{5.2751} = 2.2968$					BIC	58.9804

図 2.1 には、故障したファンの平均値は 3074 時間、標準偏差が 2398 時間となっている。最尤法による結果は、平均値 1000 倍しては同じになるが、標準偏差は $\hat{\sigma} = \sqrt{5.2454 \times 1000^2} = 2296.8$ と異なる。最尤法による標準偏差は、偏差平方和をデータ数 n で割った分散の平方根であるので、 $\hat{\sigma} = \sqrt{n\hat{\sigma}^2 / (n-1)} = \sqrt{12 \times 5.2454 \times 1000^2 / 11} = 2398$ となり、自由度 $(n-1)$ を考慮した不偏分散の平方根で求めた標準偏差に一致する。

最小 2 乗法での平均値の標準誤差 SE は、不偏分散を元に計算するのが、最尤法では、分散共分散行列の対角要素の平方根として推定されている。表 7.9 に示すように最尤法では、2 階の偏微分係数行列 Z の負の逆行列の対角要素は、推定されたパラメータの分散となっているので、それらの平方根を取れば、表 7.6 に示したように標準誤差 SE の推定値が得られる。

表 7.10 正規分布の最尤法によるパラメータの推定 (Excel)

正規分布		元の	2階の偏微分		負の逆行列			
Newton		パラメータ	Z		$(-Z)^{-1}$		SE	
-Raphson法	$\mu \sim =$	3.0475	-2.2748	0.0000	0.4396	0.0000	0.6630	
	$\sigma^2 \sim =$	5.2751	0.0000	-0.2156	0.0000	4.6378		2.1536

さて、表 7.9 に示した JMP の出力で尺度 $\hat{\sigma}$ は $\sqrt{5.2751}$ となることは、表中に示したが、その標準誤差は 0.4688 となっていて、 $\hat{\sigma}^2$ の標準誤差は 2.1536 なので、単に平方根をとってもまったく一致していない。これも合成分散の一般式から求めることになる。

$$\tilde{\sigma}^2 = 5.2751, \quad \text{Var}(\tilde{\sigma}^2) = 4.6378$$

標準偏差 $\tilde{\sigma}$ は、 $\tilde{\sigma}^2$ の平方根なので、

$$\tilde{\sigma} = (\tilde{\sigma}^2)^{1/2} \tag{7.22}$$

を $\tilde{\sigma}^2$ で偏微分すると

$$d = \frac{\partial \tilde{\sigma}}{\partial \tilde{\sigma}^2} = \frac{1}{2} (\tilde{\sigma}^2)^{-(1/2)} = \frac{1}{2\tilde{\sigma}} \tag{7.23}$$

が得られるので、

$$\text{Var}(\tilde{\sigma}) = d \cdot \text{Var}(\tilde{\sigma}^2) \cdot d = \left(\frac{1}{2\tilde{\sigma}}\right)^2 \text{Var}(\tilde{\sigma}^2) = \left(\frac{1}{2 \times 2.2968}\right)^2 \times 4.6387 = 0.2198$$

$$SE(\tilde{\sigma}) = \sqrt{0.2198} = 0.4688,$$

と表 7.9 の尺度 $\tilde{\sigma}$ の標準誤差 0.4688 に一致する。

このように合成分散の一般式を用いた換算によって $SE(\tilde{\sigma})$ を求めたのは、正規分布の対数尤度を偏微分する際に、偏微分が容易になるように、 $\tilde{\sigma}$ でなく $\tilde{\sigma}^2$ を用いたことに起因する。ちなみに、 $\tilde{\sigma}$ で偏微分した場合には、次のようになる。

$$\frac{\partial \ln f(t)}{\partial (\sigma^2)} = -\frac{1}{2(\sigma^2)} + \frac{(t-\mu)^2}{2(\sigma^2)^2}, \quad \frac{\partial \ln f(t)}{\partial \sigma} = -\frac{1}{\sigma} + \frac{(t-\mu)^2}{\sigma^3} \tag{7.24}$$

$$\frac{\partial \ln f(t)}{\partial \mu \partial (\sigma^2)} = \frac{(t-\mu)}{(\sigma^2)^2}, \quad \frac{\partial \ln f(t)}{\partial \mu \partial \sigma} = -\frac{2(t-\mu)}{\sigma^3} \tag{7.25}$$

$$\frac{\partial \ln f(t)}{\partial (\sigma^2) \partial (\sigma^2)} = \frac{1}{2(\sigma^2)^2} - \frac{(t-\mu)^2}{2(\sigma^2)^3}, \quad \frac{\partial \ln f(t)}{\partial \sigma \partial \sigma} = \frac{1}{\sigma^2} - \frac{3(t-\mu)^2}{\sigma^4} \tag{7.26}$$

検証は、JMP で対数尤度を偏微分し、その計算式を手作業で簡単化し、その式を再度 JMP の計算式で入力して、自動微分の結果と一致することで行った

8. 回帰分析への拡張

8.1 最尤法による加速試験データに対する回帰分析

最小 2 乗法による回帰分析は、目的変数 y_i に対して、説明変数 (x_1, x_2, \dots, x_p) としたとき、

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i, \quad (8.1)$$

とし、誤差 ε_i が独立に平均が 0 の正規分布に従うと仮定する。正規方程式を解いて $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)$ を推定する。次式により

$$\hat{\varepsilon}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_p x_{pi}), \quad (8.2)$$

誤差項を事後的に推定し、分散を

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{n-p-1}, \quad (8.3)$$

として推定している。

式 (8.1) の誤差項 ε_i は、正規分布と仮定するだけで、最小 2 乗法の計算過程で正規分布の確率密度関数は、全く使用していない。最尤法を適用する場合は、誤差項が正規分布に従うとした場合に、誤差項に標準正規分布の確率密度関数 ϕ_{NOR} を使って、尤度を

$$L_i = \frac{1}{\sigma_{\text{NOR}}} \phi_{\text{NOR}} \left(\frac{y_i - \hat{y}_i}{\sigma_{\text{NOR}}} \right) \quad (8.4)$$

としている。ここで、 \hat{y}_i は、

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_p x_{pi} \quad (8.5)$$

である。最尤法は、それぞれの尤度 L_i 誤差の積、

$$L = \prod_{i=1}^n L_i = \prod_{i=1}^n \left[\frac{1}{\sigma_{\text{NOR}}} \phi_{\text{NOR}} \left(\frac{y_i - \hat{y}_i}{\sigma_{\text{NOR}}} \right) \right] \quad (8.6)$$

を最大化するような $(\hat{\sigma}_{\text{NOR}}, \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)$ を求める。

故障があるデータの場合： $\delta_i = 1$ 、打ち切りデータ： $\delta_i = 0$ としたときに尤度は、

$$L_i = \frac{1}{\sigma} \phi_{\text{NOR}} \left(\frac{y_i - \hat{y}_i}{\sigma} \right)^{\delta_i} \left[1 - \Phi_{\text{NOR}} \left(\frac{y_i - \hat{y}_i}{\sigma} \right) \right]^{1-\delta_i} \quad (8.7)$$

として表すことができる。

さて、最小極値分布の場合は、累積分布関数を Φ_{NOR} に変えて、 Φ_{SEV} を用いる。

$$\Phi_{\text{SEV}} = F(y_i) = 1 - \exp\left\{-\exp\left(\frac{y_i - \hat{y}_i}{\sigma_{\text{SEV}}}\right)\right\} \quad (8.8)$$

最小極値分布の確率密度関数は、累積分布関数を y_i で微分することにより、

$$f(y_i) = \frac{1}{\sigma_{\text{SEV}}} \cdot \exp\left(\frac{y_i - \hat{y}_i}{\sigma_{\text{SEV}}}\right) \cdot \exp\left\{-\exp\left(\frac{y_i - \hat{y}_i}{\sigma_{\text{SEV}}}\right)\right\} \quad (8.9)$$

が得られる。 Φ_{SEV} を基準化した $u_i = (y_i - \hat{y}_i) / \sigma_{\text{SEV}}$ で微分すると、

$$\phi_{\text{SEV}} = \exp(u_i) \cdot \exp[-\exp(u_i)] \quad (8.10)$$

となるので、 $f(y_i)$ は、基準化した ϕ_{SEV} を用いた場合には、

$$f(y_i) = \frac{1}{\sigma_{\text{SEV}}} \phi_{\text{SEV}} \quad (8.11)$$

と σ_{SEV} で除することになる。式 (8.7) の正規分布の場合と同様に、誤差の分布に最小極値分布を仮定した場合の尤度は、

$$L_i = \frac{1}{\sigma_{\text{SEV}}} \phi_{\text{SEV}} \left(\frac{y_i - \hat{y}_i}{\sigma_{\text{SEV}}}\right)^{\delta_i} \left[1 - \Phi_{\text{SEV}}\left(\frac{y_i - \hat{y}_i}{\sigma_{\text{SEV}}}\right)\right]^{1-\delta_i} \quad (8.12)$$

となる。ここで、 y_i ではなく、 $\ln(t_i)$ とした場合には、累積分布関数を t_i で微分するので、

$$\varepsilon_i \approx \frac{1}{t_i \sigma_{\text{SEV}}} \phi_{\text{SEV}} \left(\frac{\ln(t_i) - \hat{y}_i}{\sigma_{\text{SEV}}}\right)^{\delta_i} \left[1 - \Phi_{\text{SEV}}\left(\frac{\ln(t_i) - \hat{y}_i}{\sigma_{\text{SEV}}}\right)\right]^{1-\delta_i} \quad (8.13)$$

となる。

8.2 Excel を用いた最尤法による回帰分析

表 2.2 のデバイス A の寿命データについて、第 4 章で様々な観点からワイブル分布のあてはめを行ってきた。表 4.5 では、加速条件である温度を x とし、誤差をワイブル分布とした回帰分析を Excel で行った結果を示した。また、JMP を用いた結果は、表 5.5 および 図 5.9 に示した。

時間について自然対数を取り、誤差分布を最小極値分布に従うとした場合に、表 8.1 に示すように Excel による計算を行うことができる。表 4.5 のワイブル分布を誤差と仮定した場合の Excel の表を元に最小極値分布に変更して作成した。

表 8.1 最小極値分布を用いた加速試験の解析

		$y_i = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_i$					$\beta_{\text{切片}} = 12.7929$	$\beta_{\text{傾き}} = -0.0663$	$\ln L = -77.3611$			
							$\sigma = 0.7011$					
変数 x		件					パラメータ		最小極値分布		尤度	
群	温度	t	$y = \ln(t)$	数	δ	生存率	\hat{y}_i	u^i	$f(y)$	$1-F(y)$	L_i	$\ln L_i$
	0						12.793					
1	10	5000	8.517	30	0	1.000	12.130	-5.153	0.008	0.994	0.994	-0.173
2	40	1298	7.169	1	1	0.990	10.141	-4.239	0.020	0.986	0.020	-3.899
2	40	1390	7.237	1	1	0.980	10.141	-4.142	0.022	0.984	0.022	-3.802
:												
2	40	4982	8.514	1	1	0.900	10.141	-2.321	0.127	0.906	0.127	-2.064
2	40	5000	8.517	90	0	0.900	10.141	-2.316	0.128	0.906	0.906	-8.883
3	60	581	6.365	1	1	0.950	8.814	-3.494	0.042	0.970	0.042	-3.169
3	60	925	6.830	1	1	0.900	8.814	-2.831	0.079	0.943	0.079	-2.535
:												
3	60	4674	8.450	1	1	0.550	8.814	-0.520	0.468	0.552	0.468	-0.759
3	60	5000	8.517	11	0	0.550	8.814	-0.424	0.485	0.520	0.520	-7.199
4	80	283	5.645	1	1	0.933	7.488	-2.629	0.096	0.930	0.096	-2.346
4	80	361	5.889	1	1	0.867	7.488	-2.281	0.132	0.903	0.132	-2.028
:												
4	80	2884	7.967	1	1	0.067	7.488	0.683	0.390	0.138	0.390	-0.942
4	80	5000	8.517	1	0	0.067	7.488	1.468	0.081	0.013	0.013	-4.340

表 4.5 では、温度とそれぞれの Weibull α_i との関係を

$$\alpha_i = \exp(\alpha_{\text{切片}} + \alpha_{\text{傾き}} \text{温度} x_i) \quad (8.14)$$

とした。回帰式に指数を取ったのは、温度と Weibull α_i の関係が直線的ではなく、Weibull α_i の対数を取った場合に直線のあてはまりが良いとの結果であったためである。

最小極値分布を用いる場合は、時間 t に対数を取ったので、指数を取る必要がなくなり、一般的な回帰式のパラメータを用いて

$$y_i = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_i \quad (8.15)$$

として与える。ただし、この式に一般的な回帰式のように、誤差 ε_i を単純に加えることができない。最尤法の場合には、パラメータが推定され、

$$\hat{y}_i = \hat{\beta}_{\text{切片}} + \hat{\beta}_{\text{傾き}} \text{温度} x_i \quad (8.16)$$

とした場合に、

$$\varepsilon_i = u_i = \frac{y_i - \hat{y}_i}{\sigma_{\text{SEV}}} \quad (8.17)$$

ただし、 $\varepsilon_i \sim \phi_{\text{SEV}}$, $L_i \sim \phi_{\text{SEV}} / \sigma_{\text{SEV}}$ $\delta_i = 1$: 故障データ

$\varepsilon_i \sim \Phi_{\text{SEV}}$, $L_i \sim (1 - \Phi_{\text{SEV}})$ $\delta_i = 0$: 打ち切りデータ

となり y_i とは異なる単位系 (無単位) となるので、回帰式に付け加えることは不可能である。

式を y_i について解くと

$$y_i = \hat{y}_i + \sigma_{\text{SEV}} \varepsilon_i \quad (8.18)$$

が得られる。この式には、打ち切りデータの影響は影響を受けないので、 \hat{y}_i を回帰式に展開

すれば,

$$y_i = \hat{\beta}_{\text{切片}} + \hat{\beta}_{\text{傾き}} \text{温度} x_i + \sigma_{\text{SEV}} \varepsilon_i \quad (8.19)$$

と表すことも可能ではある. しかしながら, ε_i が $\varepsilon_i = u_i = (y_i - \hat{y}_i) / \sigma_{\text{SEV}}$ であることを前提とし, 一般的な ($\varepsilon_i = y_i - \hat{y}_i$) ではないことに注意しなければならない.

表 8.1 で, 10°C の $t_1 = 5000$ の打ち切りデータの場合について式 (8.19) をあてはめてみよう.

$$y_1 = \ln(5000) = 8.517$$

$$\hat{y}_1 = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_1 = 12.7929 - 0.0663 \times 10 = 12.130$$

$$\varepsilon_1 = u_1 = \frac{y_1 - \hat{y}_1}{\sigma_{\text{SEV}}} = \frac{8.517 - 12.130}{0.7011} = -5.153$$

$$y_1 = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_1 + \sigma_{\text{SEV}} \varepsilon_1 = 12.130 + 0.7011 \times (-5.153) = 8.517$$

$$t_1 = \exp(8.517) = 5000$$

と, 元の時間 $t_1 = 5000$ が計算される.

引き続き, 40°C の $t_2 = 1298$ の故障データの場合については,

$$y_2 = \ln(1298) = 7.169$$

$$\hat{y}_2 = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_2 = 12.7929 - 0.0663 \times 40 = 10.141$$

$$\varepsilon_2 = u_2 = \frac{y_2 - \hat{y}_2}{\sigma_{\text{SEV}}} = \frac{7.169 - 10.141}{0.7011} = -4.239$$

$$y_2 = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_2 + \sigma_{\text{SEV}} \varepsilon_2 = 10.141 + 0.7011 \times (-4.239) = 7.169$$

$$t_2 = \exp(7.169) = 1298$$

と, 元の時間 $t_2 = 1298$ が計算される.

最後の, 80°C の $t_{37} = 5000$ の打ち切りデータは,

$$y_{37} = \ln(5000) = 8.517$$

$$\hat{y}_{37} = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_{37} = 12.7929 - 0.0663 \times 80 = 7.488$$

$$\varepsilon_{37} = u_{37} = \frac{y_{37} - \hat{y}_{37}}{\sigma_{\text{SEV}}} = \frac{8.517 - 7.488}{0.7011} = 1.468$$

$$y_{37} = \beta_{\text{切片}} + \beta_{\text{傾き}} \text{温度} x_{37} + \sigma_{\text{SEV}} \varepsilon_{37} = 7.488 + 0.7011 \times 1.468 = 8.517$$

$$t_{37} = \exp(8.517) = 5000$$

と, 元の時間 $t_{37} = 5000$ が計算される.

さらに, 一般化した場合には, デザイン行列を \mathbf{X} , パラメータベクトルを $\boldsymbol{\beta}$ として,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\sigma}\boldsymbol{\varepsilon} \quad (8.20)$$

がパラメトリックな生存時間(寿命データ)解析の表記法として用いられている. 誤差項に関して, $\boldsymbol{\varepsilon}$ が σ で基準化した誤差なので, σ を掛けて一般的な誤差に戻したことになる.

8.3 JMP による解析

JMP による解析結果は、時間 t についてワイブル分布を用いた解析を行った場合でも、内部では対数時間 $\ln(t)$ による計算が行なわれ、結果も対数表示となっていた。図 8.1 に示すように、寿命の二変量で、対数時間を使う場合には、分布をワイブル分布ではなく最小極値分布とすることが必須である。



図 8.1 JMP で最小極値分布を用いた解析

JMP では、どのような組み合わせでも計算が行われる。しっかりとした基礎知識なしに、JMP の「寿命の二変量」を用いると「誤用」を誘発する。

最小極値分布を指定した場合には、Y 軸の目盛が自然対数の目盛となっている。実用的には、この目盛ではのままでは使いづらい。しかしながら、時間 t を用いて、ワイブル分布を使ったした場合であっても、JMP の内部での計算の仕組みを理解するためには、最小極値分布を明示的に使う経験が必要である。

表 8.2 最小極値分布を用いた加速試験のパラメータ推定

パラメータ	推定値	標準誤差	下側95%	上側95%	β_0	β_1	σ
β_0	12.7930	0.6849	11.4506	14.1354	0.4691	-0.0067	0.0514
β_1	-0.0663	0.0100	-0.0859	-0.0468	-0.0067	0.0001	-0.0007
σ	0.7011	0.1018	0.5016	0.9006	0.0514	-0.0007	0.0104

分散共分散行列

推定された回帰式は、

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 12.7930 - 0.0663x$$

となり、図 8.2 に示すように、通常回帰分析と同様に、切片、傾きの解釈が容易である。

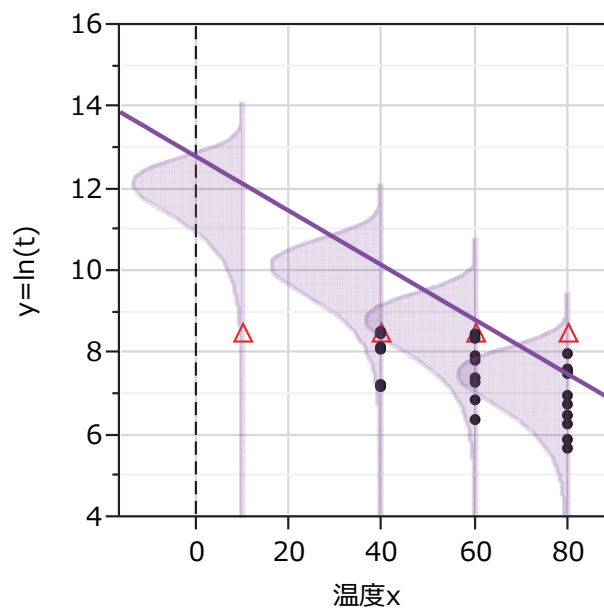


図 8.2 最小極値分布を用いた回帰直線

第5章では、 $t/1000$ についてワイブル分布のあてはめを行い、表 5.5 に結果を示したように、パラメータの推定結果は、自然対数での表示になっていることに注意してもらいたい。

表 5.5 (再掲) ワイブル分布を誤差とする回帰パラメータおよび分散共分散の推定

パラメータ	推定値	標準誤差	下側95%	上側95%	β_0	β_1	σ
β_0	12.7930	0.6849	11.4506	14.1354	0.4691	-0.0067	0.0514
β_1	-0.0663	0.0100	-0.0859	-0.0468	-0.0067	0.0001	-0.0007
σ	0.7011	0.1018	0.5016	0.9006	0.0514	-0.0007	0.0104
$\mu = 12.793 - 0.06631011 * \text{温度}$							

このように、結果の解釈を行うために、ワイブル分布に関する知識だけでは不十分で、最小極値に関する知識も必須である。

また、JMPの「寿命の一変量」によるワイブル分布のあてはめを行った際に、各種の95%信頼区間の推定を行う場合に用いる分散共分散行列も、Weibull α と Weibull β ではなく、最小極値分布を用いた結果となっている。

奥野監訳本では、表 7.2 で示したように Weibull α と Weibull β の分散共分散行列を直接推定しているが、それを最小極値分布の分散共分散行列に計算し直してから、ワイブル分布に関連する95%信頼区間の計算をしている。このようなこともあり、奥野監訳本 2.5 節 最小極値分布 (p37) に、「この節で、は最小極値分布について述べるが、それはワイプルデータに対する解析的方法 (6 章から 12 章まで) に関して必須の予備知識である。」と強調されている。

8.3 Excel による分散共分散の推定

分散共分散行列を得るために、(8.12) 式の対数をパラメータ $\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}_{SEV}$ で偏微分すると、少々煩雑であるが、以下の結果が得られる。これらの微分を行なう際には、6.3 節での $\hat{\mu}$ と $\hat{\sigma}$ についての微分式が活用できる。それは、 $\hat{\mu}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ としたときに、 $\hat{\beta}_0$ での微分は $\hat{\mu}$ での微分と同じ結果なるからで、 $\hat{\beta}_1$ での微分は、 $\hat{\beta}_0$ の微分式に x_i または x_i^2 が係数として付加される。同時間で繰り返し数 n_i がある場合には、対数尤度では、 n_i を掛けることになる

$$(6.9) \quad \ln L_i = n_i \left[\delta_i \left(-\ln(\sigma) + \frac{y_i - \hat{y}_i}{\sigma} \right) - \exp\left(\frac{y_i - \hat{y}_i}{\sigma}\right) \right] \quad (8.21)$$

$$(6.10) \quad \frac{\partial \ln L_i}{\partial \hat{\beta}_0} = n_i \left[\delta_i \left(-\frac{1}{\hat{\sigma}} \right) + \frac{1}{\hat{\sigma}} \exp\left(\frac{y_i - \hat{y}_i}{\hat{\sigma}}\right) \right] \quad (8.22)$$

$$\frac{\partial \ln L_i}{\partial \hat{\beta}_1} = x_i \cdot \frac{\partial \ln L_i}{\partial \hat{\beta}_0} \quad (8.23)$$

$$(6.11) \quad \frac{\partial \ln L_i}{\partial \hat{\sigma}} = n_i \left[\delta_i \left(-\frac{1}{\hat{\sigma}} - \frac{y_i - \hat{y}_i}{\hat{\sigma}^2} \right) + \left(\frac{y_i - \hat{y}_i}{\hat{\sigma}^2} \right) \cdot \exp\left(\frac{y_i - \hat{y}_i}{\hat{\sigma}}\right) \right] \quad (8.24)$$

$$(6.12) \quad \frac{\partial^2 \ln L_i}{\partial \hat{\beta}_0^2} = n_i \left[-\frac{1}{\hat{\sigma}^2} \exp\left(\frac{y_i - \hat{y}_i}{\hat{\sigma}}\right) \right] \quad (8.25)$$

$$\frac{\partial^2 \ln L_i}{\partial \hat{\beta}_0 \partial \hat{\beta}_1} = x_i \cdot \frac{\partial^2 \ln L_i}{\partial \hat{\beta}_0^2} \quad (8.26)$$

$$(6.14) \quad \frac{\partial^2 \ln L_i}{\partial \hat{\beta}_0 \partial \hat{\sigma}} = n_i \left[\delta_i \left(\frac{1}{\hat{\sigma}^2} \right) - \left\{ \frac{1}{\hat{\sigma}^2} + \frac{y_i - \hat{y}_i}{\hat{\sigma}^3} \right\} \exp\left(\frac{y_i - \hat{y}_i}{\hat{\sigma}}\right) \right] \quad (8.27)$$

$$\frac{\partial^2 \ln L_i}{\partial \hat{\beta}_1^2} = x_i^2 \cdot \frac{\partial^2 \ln L_i}{\partial \hat{\beta}_0^2} \quad (8.28)$$

$$\frac{\partial^2 \ln L_i}{\partial \hat{\beta}_1 \partial \hat{\sigma}} = x_i \cdot \frac{\partial^2 \ln L_i}{\partial \hat{\beta}_0 \partial \hat{\sigma}} \quad (8.29)$$

$$(6.14) \quad \frac{\partial^2 \ln L_i}{\partial \hat{\sigma}^2} = n_i \left[\delta_i \left(\frac{1}{\hat{\sigma}^2} + 2 \frac{y_i - \hat{y}_i}{\hat{\sigma}^3} \right) - \left\{ 2 \frac{y_i - \hat{y}_i}{\hat{\sigma}^3} + \left(\frac{y_i - \hat{y}_i}{\hat{\sigma}^2} \right)^2 \right\} \exp \left(\frac{y_i - \hat{y}_i}{\hat{\sigma}} \right) \right] \quad (8.30)$$

これらの偏微分式が正しいとの検証は、Excel でこれらの式を計算シート落とし込み、分散共分散行列を実際に計算し、統計ソフトの結果と一致することを確認することで行う。表 8.3 に Excel に、2 階の偏微分式を落とし込み、分散共分散行列 $(-Z)^{-1}$ の計算結果を示す。表 8.2 の結果と一致していることで確かめられる。

表 8.3 Excel による分散共分散行列の計算

		$y_i = \beta_{\text{切片}} + \beta_{\text{傾き}} \times \text{温度} x_i$								Z			$(-Z)^{-1}$				
		$\beta_{\text{切片}} =$	12.7929	$\beta_{\text{傾き}} =$	-0.0663	$\ln L =$		-77.3611		-67.1	-4191	47.0	0.469	-0.007	0.051		
		$\sigma =$	0.7011							-4191	-2.8E+05	1598	-0.007	0.000	-0.001		
										47.0	1598	-220.4	0.051	-0.001	0.010		
変数	x	件		パラメータ		最小極値分布		尤度		2階の偏微分							
群	温度	$y = \ln(t)$	数	δ	生存率	y^{\wedge}_i	u^{\wedge}_i	$f(y)$	$1-F(y)$	Li	$\ln L_i$	$\beta_0^{\wedge 2}$	$\beta_0 \beta_1$	$\beta_0 \sigma$	$\beta_1^{\wedge 2}$	$\beta_1 \sigma$	$\sigma \sigma$
	0					12.793						-67.14	-4191.36	46.961	-281187	1598	-220.37
1	10	8.517	30	0	1.000	12.13	-5.15	0.01	0.99	0.99	-0.173	-0.35	-3.5	1.47	-35	14.7	-5.73
2	40	7.169	1	1	0.990	10.14	-4.24	0.02	0.99	0.02	-3.898	-0.03	-1.2	2.13	-47	85.2	-15.49
2	40	7.237	1	1	0.980	10.14	-4.14	0.02	0.98	0.02	-3.802	-0.03	-1.3	2.14	-52	85.4	-15.11
:																	
2	40	8.514	1	1	0.900	10.14	-2.32	0.13	0.91	0.13	-2.064	-0.20	-8.0	2.30	-320	91.9	-7.56
2	40	8.517	90	0	0.900	10.14	-2.32	0.13	0.91	0.91	-8.884	-18.08	-723.0	23.78	-28920	951.2	-13.21
3	60	6.365	1	1	0.950	8.81	-3.49	0.04	0.97	0.04	-3.169	-0.06	-3.7	2.19	-222	131.3	-12.51
3	60	6.830	1	1	0.900	8.81	-2.83	0.08	0.94	0.08	-2.535	-0.12	-7.2	2.25	-432	135.3	-9.77
:																	
3	60	8.450	1	1	0.550	8.81	-0.52	0.47	0.55	0.47	-0.759	-1.21	-72.6	1.45	-4354	87.2	0.85
3	60	8.517	11	0	0.550	8.81	-0.42	0.49	0.52	0.52	-7.199	-14.65	-878.9	-8.44	-52732	-506.3	9.79
4	80	5.645	1	1	0.933	7.49	-2.63	0.10	0.93	0.10	-2.346	-0.15	-11.7	2.27	-940	181.9	-8.90
4	80	5.889	1	1	0.867	7.49	-2.28	0.13	0.90	0.13	-2.028	-0.21	-16.6	2.30	-1330	184.1	-7.38
:																	
4	80	7.967	1	1	0.067	7.49	0.68	0.39	0.14	0.39	-0.942	-4.03	-322.2	-4.74	-25777	-379.5	-2.57
4	80	8.517	1	0	0.067	7.49	1.47	0.08	0.01	0.01	-4.340	-8.83	-706.3	-21.79	-56507	-1743.1	-44.94
										計	-77.36	-67.14	-4191.36	46.96	-281187	1597.73	-220.37

注) $\hat{\beta}_{\text{切片}} = 12.7929$ であるが、JMP の表 5.5 では 12.7930 と少数点以下 4 桁が一致しない。

これは、Excel のソルバーの収束条件が JMP に比べて緩いため起きています。

Excel でのニュートン・ラフソン法では、JMP の結果と一致した。

結果は、本文中に示さないが、本章の Excel での計算シートに含まれている。

最小極値分布の分散共分散行列は、式 (6.7) で次のように示した。

$$\Sigma_i = \begin{bmatrix} -\frac{\partial^2 \ln L_i}{\partial^2 \mu} & -\frac{\partial^2 \ln L_i}{\partial \mu \partial \sigma} \\ -\frac{\partial^2 \ln L_i}{\partial \sigma \partial \mu} & -\frac{\partial^2 \ln L_i}{\partial^2 \sigma} \end{bmatrix}^{-1}$$

最小極値分布を誤差とする回帰分析の場合は、それぞれの 2 階の偏微分式を、行列で表せば、

$$\Sigma_i = \begin{bmatrix} -\frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -x_i \cdot \frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -\frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} \\ -x_i \cdot \frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -x_i^2 \cdot \frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -x_i \cdot \frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} \\ -\frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} & -x_i \cdot \frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} & -\frac{\partial^2 \ln L_i}{\partial^2 \sigma} \end{bmatrix}^{-1} \quad (8.31)$$

となる。温度だけではなく湿度も含めて加試験を行う場合、複数の加速因子を含めるような場合もある。このような場合の分散共分散行列は、容易に求めることができる。式 (8.31) は、温度 x_i について規則性が見い出される。

温度 x_i についてのパラメータ β_1 が現れない代わりに、 β_0 での偏微分式に x_i が積で付加されていること、 β_0 の 2 階の偏微分式には、 x_i^2 が積で付加されていることから、 x_i を除いた行列

$$\Sigma_i = \begin{bmatrix} -\frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -\frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -\frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} \\ -\frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -\frac{\partial^2 \ln L_i}{\partial^2 \beta_0} & -\frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} \\ -\frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} & -\frac{\partial^2 \ln L_i}{\partial \beta_0 \partial \sigma} & -\frac{\partial^2 \ln L_i}{\partial^2 \sigma} \end{bmatrix}^{-1} \quad (8.32)$$

に、次のようなベクトルの積

$$\begin{bmatrix} 1 \\ x_i \\ 1 \end{bmatrix} \begin{bmatrix} 1 & x_i & 1 \end{bmatrix} = \begin{bmatrix} 1 & x_i & 1 \\ x_i & x_i^2 & x_i \\ 1 & x_i & 1 \end{bmatrix} \quad (8.33)$$

を、行列のセル同士の積を取り、すべての i について加えればよいことがよい。湿度を加えたような場合には、温度を x_{1i} 、湿度を x_{2i} とし、

$$\begin{bmatrix} 1 \\ x_{1i} \\ x_{2i} \\ 1 \end{bmatrix} \begin{bmatrix} 1 & x_{1i} & x_{2i} & 1 \end{bmatrix} = \begin{bmatrix} 1 & x_{1i} & x_{2i} & 1 \\ x_{1i} & x_{1i}^2 & x_{1i}x_{2i} & x_{1i} \\ x_{2i} & x_{2i}x_{1i} & x_{2i}^2 & x_{2i} \\ 1 & x_{1i} & x_{2i} & 1 \end{bmatrix} \quad (8.34)$$

とればよいのだが、Excel での 4×4 の「手計算」は煩雑である。

9. 多因子による寿命試験の事例

9.1 接着剤の寿命試験

廣野(2000)は、加速因子に温度を3水準、湿度3水準として、2種類の接着剤についての加速試験の実験データの解析事例を示した。幸いデータがWEBからダウンロードできたので、多因子実験における寿命データの解析を試みる。

実験は、

G : 接着剤 (x_1), [A : $x_1=1$, B : $x_1=-1$]

T : 温度 (x_2), [30°C, 40°C, 50 度]

H : 相対湿度 (x_3), [60%, 70%, 80%]

繰り返し測定 3 回

であり、3元配置分散分析で、測定が3回繰り返されている。データを表 9.1 に示す。反応は剥がれるまでの時間であるが、途中で打ち切りとなるデータが混在している。打ち切りデータがランダムに発生しているのであれば、打ち切りデータを欠測値にして、一般的な線形モデルあるいは線形混合モデルで解いても、打ち切りデータを無視しても推定値にバイアスが入り込まない。

このデータは、ランダムな打ち切りとも思われるが、どのような打ち切りなのかが示されていないので、多因子の寿命データの解析を試みることにする。

表 9.1 接着剤の加速試験データ

相対湿度	glue = A						glue = B					
	60%		70%		80%		60%		70%		80%	
温度	Day	Cen sor	Day	Cen sor	Day	Cen sor	Day	Cen sor	Day	Cen sor	Day	Cen sor
30°C	54	0	29	0	16	1	6	1	32	0	17	0
	71	0	39	0	21	0	10	0	50	1	27	1
	113	1	62	0	34	0	24	0	13	0	7	0
40°C	38	0	21	0	11	0	4	0	22	0	12	0
	50	0	27	1	15	1	7	1	35	1	19	1
	79	0	43	0	24	0	17	0	9	0	5	0
50°C	13	0	15	0	13	0	12	0	16	0	8	0
	36	0	19	0	10	0	9	0	25	1	14	1
	57	0	31	1	17	1	15	1	6	0	3	0

Censor=1 が打ち切りデータ

廣野は、Weibull α を

$$\alpha_{ijk} = \exp\left(\alpha_{\text{切片}} + \alpha_{\text{種類}_i} x_1 + \alpha_{\text{温度}_j} \frac{11605}{x_2 + 273.15} + \alpha_{\text{湿度}_k} x_3\right) + \sigma_{\text{SEV}} \varepsilon_{ijk} \quad (9.1)$$

とし Weibull β を共通とするワイブル回帰を適用し、パラメータの推定値 $\hat{\alpha}_{\text{種類}}$ の指数をとって、 $\exp(0.2575) = 1.3$ 倍 信頼性が高いと結論した。また、アレニウス変換温度のパラメータ $\hat{\alpha}_{\text{温度}}$ が、反応論モデルでは、活性化エネルギー推定値になり

$$Ea = 0.2844, \quad 95\% \text{信頼区間} : (0.1033 \sim 0.4646)$$

となることを示した。

廣野が示した解析を、Excel によって再現してみよう。表 4.6 のデバイス A に対するアレニウス変換温度の Excel シートを元に、変数 (接着剤, 湿度) を追加し 表 9.2 の結果を得る。

表 9.2 接着剤の寿命データの解析

$\alpha = \exp\left(\alpha_{\text{切片}} + \alpha_{\text{種類}_i} x_1 + \alpha_{\text{温度}_j} \frac{11605}{x_2 + 273.15} + \alpha_{\text{湿度}_k} x_3\right)$										$\alpha_{\text{切片}}$	$\alpha_{\text{種類}}$	$\alpha_{\text{温度}}$	$\alpha_{\text{湿度}}$	$\ln L =$
										-4.8574	0.2576	0.2845	-0.0330	-165.312
										1.8596				
群	glue	± 1	温度	アレニウス温度	湿度	t	数	δ	パラメータ		ワイブル分布		尤度	
	glue								α	β	$f(t)$	$1-F(t)$	Li	$\ln L_i$
1	A	1	30	38.3	60	54	1	1	74.418	1.860	0.011	0.576	0.011	-4.516
1	A	1	30	38.3	60	71	1	1	74.418	1.860	0.010	0.400	0.010	-4.646
1	A	1	30	38.3	60	113	1	0	74.418	1.860	0.004	0.114	0.114	-2.174
2	A	1	30	38.3	70	29	1	1	53.499	1.860	0.015	0.726	0.015	-4.206
:														
8	A	1	50	35.9	70	31	1	0	27.267	1.860	0.021	0.281	0.281	-1.269
9	A	1	50	35.9	80	13	1	1	19.602	1.860	0.042	0.628	0.042	-3.174
9	A	1	50	35.9	80	10	1	1	19.602	1.860	0.040	0.751	0.040	-3.220
9	A	1	50	35.9	80	17	1	0	19.602	1.860	0.039	0.464	0.464	-0.767
10	B	-1	30	38.3	60	6	1	0	44.460	1.860	0.007	0.976	0.976	-0.024
10	B	-1	30	38.3	60	10	1	1	44.460	1.860	0.011	0.940	0.011	-4.519
10	B	-1	30	38.3	60	24	1	1	44.460	1.860	0.018	0.728	0.018	-4.022
11	B	-1	30	38.3	70	32	1	1	31.962	1.860	0.021	0.367	0.021	-3.845
:														
17	B	-1	50	35.9	70	6	1	1	16.290	1.860	0.041	0.855	0.041	-3.185
18	B	-1	50	35.9	80	8	1	1	11.711	1.860	0.070	0.611	0.070	-2.660
18	B	-1	50	35.9	80	14	1	0	11.711	1.860	0.046	0.248	0.248	-1.394
18	B	-1	50	35.9	80	3	1	1	11.711	1.860	0.045	0.924	0.045	-3.090

この Excel での結果と廣野が行った JMP の「生存時間 (パラメトリック) のあてはめ」の解析法と結果とを対比する。JMP の解析スクリプトを次に示す。

```

モデルのあてはめ(
  Y( :Day ),
  効果( :G_接着剤, :A_温度, :H_湿度 ),
  手法( Name( "生存時間(パラメトリック)" ) ),
  一変量の分布( Weibull ),
  打ち切り( :Censor ),
  実行( 尤度比検定( 1 ) )
)

```

表 9.3 接着剤の JMP データセット

		2x3x3	G_接着剤	glue_0	T_温度	A_温度	H_湿度	Censor	Day
○	1	1	A	0	30	38.28	60	0	54
○	2	1	A	0	30	38.28	60	0	71
▼	3	1	A	0	30	38.28	60	1	113
○	4	2	A	0	30	38.28	70	0	29
○	5	2	A	0	30	38.28	70	0	39
○	6	2	A	0	30	38.28	70	0	62
▼	7	3	A	0	30	38.28	80	1	16
:									
○	51	17	B	1	50	35.91	70	0	6
○	52	18	B	1	50	35.91	80	0	8
▼	53	18	B	1	50	35.91	80	1	14
○	54	18	B	1	50	35.91	80	0	3

表 9.4 に結果を示すが、シンプルな出力で、ここに示された推定値での適切な結果の解釈には、注意を要する。廣野は、G_接着剤[A] の推定値 0.2575 を用いて、この指数 $\exp(0.2575) = 1.3$ 倍としたが、名義尺度の接着剤は (1, -1) と内部として計算されており、この推定値の 2 倍の指数 1.7 倍が、接着剤 A の B に対する適切な評価となる。

表 9.4 接着剤のワイブル回帰の結果 (JMP 生存時間 (パラメトリック))

パラメータ推定値					代替パラメトリック化		
項	推定値	標準誤差	下側信頼限界	上側信頼限界	項	Exp(推定値)	$\beta=1/\delta$
切片	-4.8655	3.3982	-11.7869	2.0259	切片	0.0077	1.8596
G_接着剤[A]	0.2575	0.0927	0.0730	0.4483	G_接着剤[A]	1.2937	
A_温度	0.2847	0.0888	0.1035	0.4651	A_温度	1.3294	
H_湿度	-0.0330	0.0123	-0.0567	-0.0071	H_湿度	0.9675	
δ	0.5377	0.0698	0.4225	0.7043			

9.2 接着剤の寿命試験に対する新たな解析法

多因子の寿命データの解析に対し伝統的なワイブル回帰のような方法には、違和感を覚える。加速因子に温度と湿度を加えているので、交互作用の検討が不可欠であり、また、接着剤の種類とそれぞれの加速因子との間の交互作用についても検討が必要である。

この違和感は、3 元配置の実験データに対して、標準的な「回帰分析」を適用し、推定されたパラメータを用いて結論を出したことにある。打ち切りデータがあるので、従来の解析法が使えないので、やむなく「回帰分析」を適用したと思われる。さらに、温度にアレニウス変換を行っているため、単純な温度と反応の関係も見えにくくなっている。

この接着剤の実験データには、3 階の繰り返し測定があり、3 回とも打ち切りになったセル

がないので、Weibull β を共通にして、それぞれのセル $2 \times 3 \times 3 = 18$ のセルごとに Weibull α を推定することが可能である。そして、推定された 18 個の Weibull α について、通常の 3 元配置分散分析を適用することにより、多面的な推定結果が得られる。

不幸なことに、あるセルの 3 個の繰り返しが全て打ち切りになった場合には、湿度を固定した上で、他の温度を用いたワイブル回帰により反応を推定することもできる。

表 9.3 に示した JMP データセットの最初の列がセル番号となっているので、寿命の二変量で、セルごとの Weibull α を推定してみよう。図 9.1 に 18 個のセルに対して Weibull $\beta = 2.2808$ と共通にして、18 個の Weibull α に対する対数目盛に対するワイブル確率密度関数が図示されている。

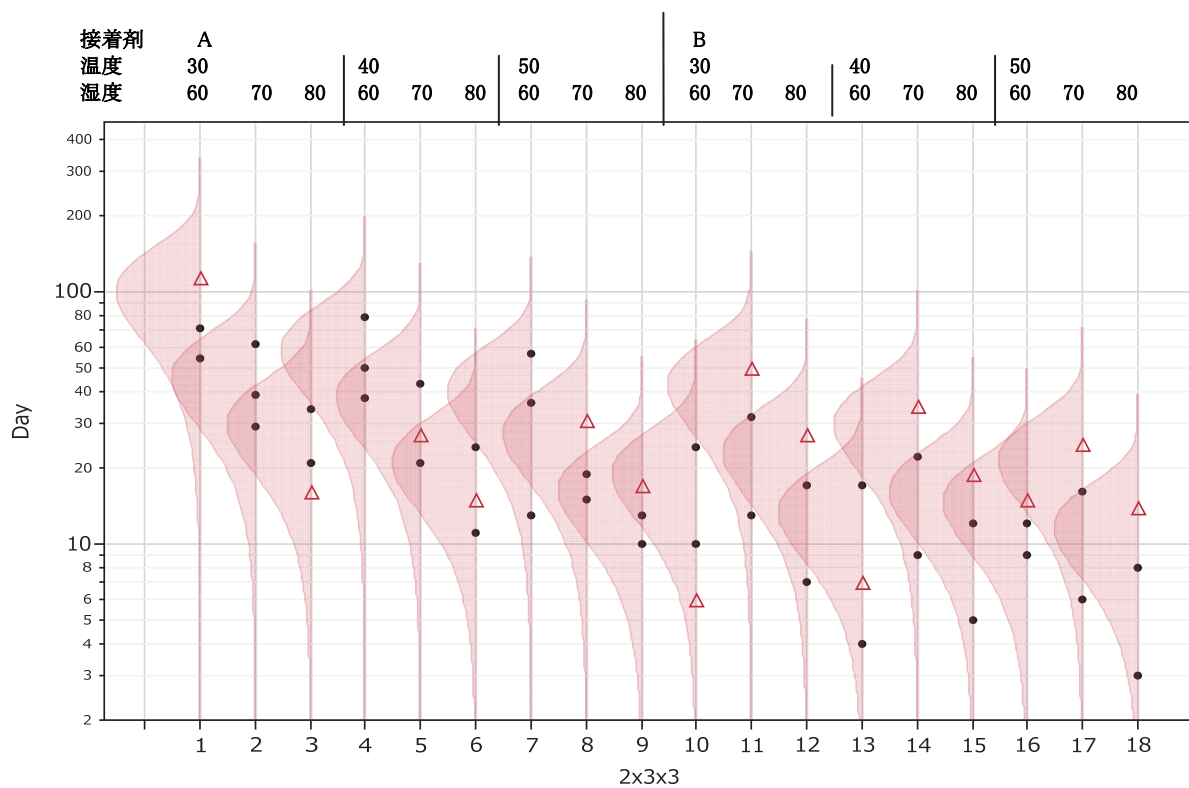


図 9.1 セルごとのワイブル回帰の適用 (JMP 寿命の二変量)

Weibull α の推定値は、それぞれの密度関数の最大値で、下側確率が 0.632 となる。温度が高くなるにつれ、剥離時間が短くなるが、湿度については、接着剤 A については、湿度が高くなるに従い剥離時間が短くなる。しかしながら、接着剤 B については、湿度が 60% の場合と 80% と同程度で、70% で剥離時間が長い結果となっている。このような製品特性を意図したものなのかは、不明である。

新製品の研究開発段階であれば、図 9.1 による評価で十分と思われるが、報告書とするためには、更なる解析が必要であろう。Weibull α について推定された統計量を表 9.5 に示す。

表 9.5 推定された Weibull α および Weibull β (JMP 寿命の二変量)

推定値				
パラメータ	推定値	標準誤差	下側95%	上側95%
Weibull α 0	100.5379	31.1860	54.7387	184.6568
Weibull α 1	46.0731	11.6797	28.0327	75.7235
Weibull α 2	30.0794	9.3298	16.3777	55.2444
Weibull α 3	58.9847	14.9507	35.8911	96.9375
Weibull α 4	38.3552	11.8980	20.8822	70.4485
Weibull α 5	21.2242	6.5829	11.5564	38.9800
Weibull α 6	40.6160	10.3306	24.6716	66.8648
:				
Weibull α 16	21.3804	6.6283	11.6447	39.2558
Weibull α 17	11.6250	3.6043	6.3311	21.3454
Weibull β	2.2809	0.3051	1.8071	3.0914

推定された Weibull α を JMP データセットに出力して, 元の 3 元表に整理した結果を 表 9.6 に示す.

表 9.6 推定された Weibull α

温度	接着剤 = A			接着剤 = B		
	相対湿度			相対湿度		
	60%	70%	80%	60%	70%	80%
30°C	100.54	46.07	30.08	19.03	42.86	23.05
40°C	58.98	38.36	21.22	13.43	29.86	16.24
50°C	40.62	27.47	16.39	14.71	21.38	11.62

この 3 元表から, 接着剤 A は, 温度と湿度の加速因子に対して常識的な結果と思われるが, 接着剤 B では, 温度に関して反応が短いレベルでほぼフラット, 湿度に関して 70% に山があることが読み取れる. そので, 形式的に (質×量×量) の 3 元元配置分散分析を行いうこととした.

温度に関する交互作用は, 検出されず主効果に有意な差が見いだされた. 湿度には, 接着剤の種類との間に交互作用が有意となった. 図 9.2 に交互作用プロットを示す. 温度と接着剤間で交互作用があるかのように見えるが, 統計的には $p=0.1557$ と有意ではない.

表 9.7 接着剤の分散分析表 (JMP モデルのあてはめ)

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	6	6719.4106	1119.90	10.3280
誤差	11	1192.7639	108.43	p値(Prob>F)
全体(修正済み)	17	7912.1745		0.0006*

パラメータ推定値				
項	推定値	標準誤差	t値	p値(Prob> t)
切片	395.1375	108.6885	3.6355	0.0039*
G_0,1	-216.374	48.7184	-4.4413	0.0010*
T_温度	-4.9607	2.6119	-1.8992	0.0841
H_湿度	-4.1639	1.5328	-2.7166	0.0201*
G_0,1*T_温度	0.9164	0.6012	1.5243	0.1557
G_0,1*H_湿度	2.2697	0.6012	3.7753	0.0031*
T_温度*H_湿度	0.0489	0.0368	1.3286	0.2109

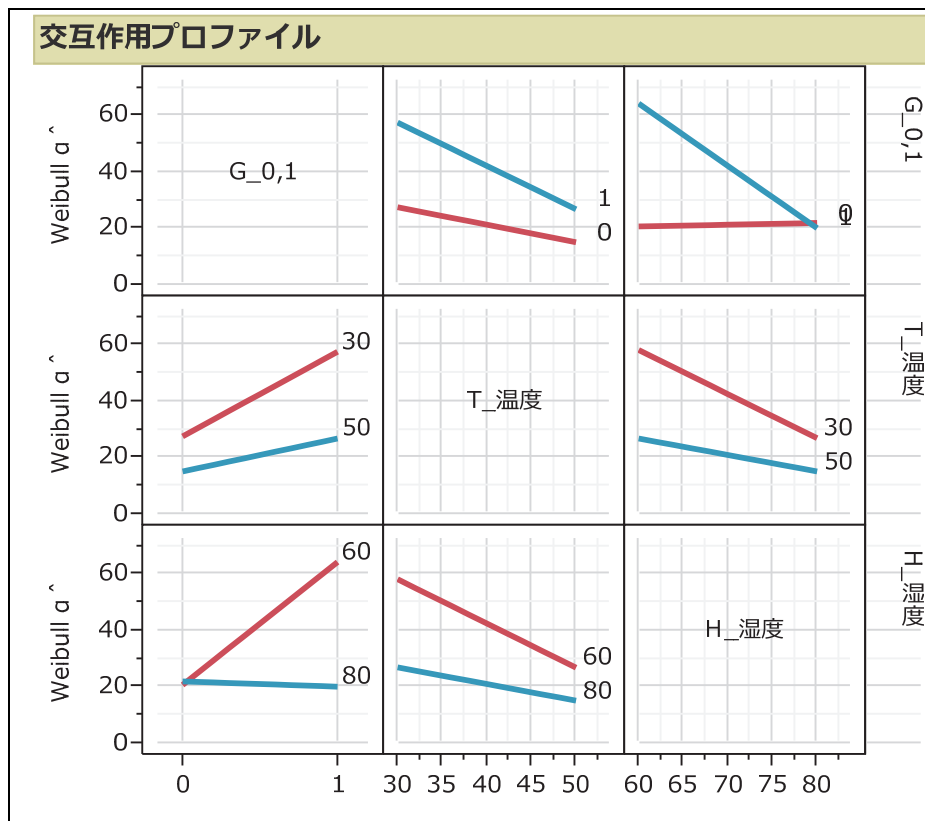


図 9.2 交互作用プロット (JMP モデルのあてはめ)

9.3 今後の課題

寿命データの解析は、試験研究段階、量産に向けた寿命の確認試験、市販後の量産品に対する品質を担保するための寿命試験を区別して対応する必要がある。その際に打ち切りデータが発生することを前提にし、本章で示したような、ワイブル回帰を用いた Weibull α などの推定値を用いた、統計解析の手順を推奨する。

試験研究段階では、加速因子のみならず他の多くの因子を含めた多因子実験が主体となり、直交表実験ならば、加速因子を外側因子として組み込むことも考えられる。あるいは、D 最適化計画で計画を立てて、応答局面解析に持ち込むこともよいだろう。

量産に向けた寿命の確認試験では、本章で示したような、加速因子を 2 次元とし、応答局面解析が主体になるであろう。

市販後の場合には、製造時期ごとに加速因子に対する故障曲線が一定であること確認することが主体になるであろう。あるいは、保障期間の半分ほどの年月が経過して時に、残りの半分の年月を保証することができるか、などの加速試験も考えられる。

10. 合成分散の一般式（デルタ法）

10.1 最小 2 乗法での場合

パラメータに関して線形の場合

最も代表的な分散の合成例は、平均値の分散の推定であろう。それぞれの x_i が互いに独立で、分散が σ^2 と共通と仮定した場合に、算術平均値

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad (10.1)$$

の分散が $Var(\bar{x}) = \sigma^2 / n$ となることは、よく知られている。なぜ、この式となるのであろうか。それぞれの x_i に $1/n$ を振り分けて

$$\bar{x} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n \quad (10.2)$$

とすると、次式のように分散に関する公式、係数 $1/n$ の 2 乗が $Var(x)$ の前に出ることを使い、

$$Var\left[\left(\frac{1}{n}\right)x_i\right] = \left(\frac{1}{n}\right)^2 Var(x_i) \quad (10.3)$$

さらに、 x_i が互いに独立であることから、分散の加法性が成り立ち、 $Var(x_i) = \sigma^2$ なので、

$$\begin{aligned} Var(\bar{x}) &= \left(\frac{1}{n}\right)^2 Var(x_1) + \left(\frac{1}{n}\right)^2 Var(x_2) + \cdots + \left(\frac{1}{n}\right)^2 Var(x_n) \\ &= \frac{n}{n^2} \sigma^2 = \frac{\sigma^2}{n} \end{aligned} \quad (10.4)$$

が求められる。

回帰分析で推定された切片 $\hat{\beta}_0$ の分散 $Var(\hat{\beta}_0)$ 、傾き $\hat{\beta}_1$ の分散 $Var(\hat{\beta}_1)$ が推定されたときに、回帰の推定値 $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ の分散 $Var(\hat{y}_i)$ は、どのようなのであろうか。切片 $\hat{\beta}_0$ と傾き $\hat{\beta}_1$ は、互いに独立ではないので、共分散 $Cov(\hat{\beta}_0, \hat{\beta}_1)$ が存在する。この場合に、次の合成分散の式が知られている。

$$Var(\hat{y}_i) = Var(\hat{\beta}_0) + 2Cov(\hat{\beta}_0, \hat{\beta}_1) + x_i^2 Var(\hat{\beta}_1) \quad (10.5)$$

非現実ではあるが、共分散が 0 と仮定すれば、分散の加法性から、

$$\begin{aligned} Var(\hat{y}_i) &= Var(\hat{\beta}_0 + \hat{\beta}_1 x_i) \\ &= Var(\hat{\beta}_0) + Var(\hat{\beta}_1 x_i) \\ &= Var(\hat{\beta}_0) + x_i^2 Var(\hat{\beta}_1) \end{aligned} \quad (10.6)$$

が得られる。

パラメータに関して非線形の場合のデルタ法

回帰分析の応用で、ある y_0 に対する \hat{x}_0 を逆推定したい。

$$y_0 = \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_0 \quad (10.7)$$

の関係から、

$$\hat{x}_0 = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1} \quad (10.8)$$

であることは、簡単な計算で確かめられる。さて、 $Var(\hat{x}_0)$ を切片 $\hat{\beta}_0$ の分散 $Var(\hat{\beta}_0)$ 、および、傾き $\hat{\beta}_1$ の分散 $Var(\hat{\beta}_1)$ から推定したい。

回帰式 $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ の分散 $Var(\hat{y}_i)$ を求める場合には、パラメータに関して線形となっていることから、 $Var(\hat{y}_i) = Var(\hat{\beta}_0) + 2Cov(\hat{\beta}_0, \hat{\beta}_1) + x_i^2 Var(\hat{\beta}_1)$ が得られた。しかしながら、式 (10.8) は、パラメータに関して非線形となっている。これは式 (10.8) を切片 $\hat{\beta}_0$ と傾き $\hat{\beta}_1$ で偏微分すると、

$$d_1 = \frac{\partial \hat{x}_0}{\partial \hat{\beta}_0} = \frac{-1}{\hat{\beta}_1}, \quad d_2 = \frac{\partial \hat{x}_0}{\partial \hat{\beta}_1} = x_0 \quad (10.9)$$

となり、切片 $\hat{\beta}_0$ で偏微分すると $\hat{\beta}_1$ が残ってしまい、線形ではなく、非線形の関係になっている。そのために、通常分散の合成式が使えない。

このような課題が、寿命データの統計解析では、しばしば起きる。このために、合成分散の一般式（デルタ法）を用いる。

逆推定値 \hat{x}_0 の分散 $Var(\hat{x}_0)$ を求めたい。回帰分析の切片 $\hat{\beta}_0$ と傾き $\hat{\beta}_1$ の分散共分散行列を

$$\Sigma = \begin{bmatrix} Var(\hat{\beta}_0) & Cov(\hat{\beta}_0, \hat{\beta}_1) \\ Cov(\hat{\beta}_0, \hat{\beta}_1) & Var(\hat{\beta}_1) \end{bmatrix} \quad (10.10)$$

としたときに、パラメータに関する偏微分式を $\mathbf{d} = [d_1, d_2]^T$ のように縦ベクトルとして、 \hat{x}_0 の分散 $Var(\hat{x}_0)$ は、 Σ に両側から \mathbf{d} を掛ける、いわゆる 2 次形式によって

$$\begin{aligned} Var(\hat{x}_0) &= \mathbf{d}^T \Sigma \mathbf{d} \\ &= \begin{bmatrix} -1/\hat{\beta}_1 & x_0 \end{bmatrix} \begin{bmatrix} Var(\hat{\beta}_0) & Cov(\hat{\beta}_0, \hat{\beta}_1) \\ Cov(\hat{\beta}_0, \hat{\beta}_1) & Var(\hat{\beta}_1) \end{bmatrix} \begin{bmatrix} -1/\hat{\beta}_1 \\ x_0 \end{bmatrix} \end{aligned} \quad (10.11)$$

求めることができる。ここでは、2次元の場合について示したが、1次元を含めて多次元にも拡張できる。合成分散の計算は、Excel の行列関数での計算を前提とするのならば、式 (10.11) のままが適しており、更なる展開をして行列計算がないようにしても、計算が煩雑で見通しが悪い。

パラメータに関して線形の場合へのデルタ法の適用

回帰式の推定値 $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ の分散 $Var(\hat{y}_i)$ について天下り式に,

$$Var(\hat{y}_i) = Var(\hat{\beta}_0) + 2x_i Cov(\hat{\beta}_0, \hat{\beta}_1) + x_i^2 Var(\hat{\beta}_1)$$

であるとしたのであるが, この場合でも, それぞれのパラメータで微分すると,

$$d_1 = \frac{\partial \hat{y}_i}{\partial \hat{\beta}_0} = 1, \quad d_2 = \frac{\partial \hat{y}_i}{\partial \hat{\beta}_1} = x_i \quad (10.12)$$

が得られる. パラメータに関する偏微分式を $\mathbf{d} = [d_1, d_2]^T$ の縦ベクトルとして, \hat{x}_0 の分散 $Var(\hat{x}_0)$ は, 次の2次形式によって

$$\begin{aligned} Var(\hat{y}_i) &= \mathbf{d}^T \Sigma \mathbf{d} \\ &= \begin{bmatrix} 1 & x_i \end{bmatrix} \begin{bmatrix} Var(\hat{\beta}_0) & Cov(\hat{\beta}_0, \hat{\beta}_1) \\ Cov(\hat{\beta}_0, \hat{\beta}_1) & Var(\hat{\beta}_1) \end{bmatrix} \begin{bmatrix} 1 \\ x_i \end{bmatrix} \\ &= Var(\hat{\beta}_0) + 2x_i Cov(\hat{\beta}_0, \hat{\beta}_1) + x_i^2 Var(\hat{\beta}_1) \end{aligned} \quad (10.13)$$

計算される. ここでは, 2次元の場合について示したが, 1次元を含めて多次元に拡張できる.

再度, 算術平均に戻ってみよう.

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n} x_1 + \frac{1}{n} x_2 + \cdots + \frac{1}{n} x_n \quad (10.14)$$

であった. それぞれの x_i の分散 σ^2 とした場合に

$$z_i = \frac{x_i}{n} \quad (10.15)$$

の分散は, z_i を x_i で偏微分すると,

$$d_i = \frac{\partial z_i}{\partial x_i} = \frac{1}{n} \quad (10.16)$$

となり, 合成分散の一般式から

$$\begin{aligned} Var(\bar{x}) &= \mathbf{d}^T \Sigma \mathbf{d} \\ &= \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} \begin{bmatrix} 1/n \\ 1/n \\ \vdots \\ 1/n \end{bmatrix} \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned} \quad (10.17)$$

となる.

10.2 最尤法における合成分散の一般式

最小 2 乗法の場合で示した合成分散の一般式は、最尤法の場合でも推定したパラメータに関する分散共分散行列 Σ を用いて、同様に計算することができる。

パラメータが 1 つの場合

第 6 章でディーゼル発電機ファンデータに対して最小極値分布をあてはめて、 $\hat{\mu}_{\text{SEV}} = 10.1722$ が推定され、その分散は $Ver(\hat{\mu}_{\text{SEV}}) = 0.4657$ であった。この $\hat{\mu}_{\text{SEV}}$ から Weibull $\hat{\alpha}$ は、 $\hat{\alpha} = \exp(\hat{\mu}_{\text{SEV}})$ で計算される。では、 $\hat{\alpha}$ の分散 $Var(\hat{\alpha})$ は、どのようにして求めたらよいのであろうか。計算式をパラメータ $\hat{\mu}_{\text{SEV}}$ で偏微分すると、

$$d = \frac{\partial \hat{\alpha}}{\partial \hat{\mu}_{\text{SEV}}} = \exp(\hat{\mu}_{\text{SEV}}) \quad (10.18)$$

が得られる。ここで、合成分散の一般式は、1 次元の場合でも適用できるので

$$Var(\hat{\alpha}) = d \cdot Ver(\hat{\mu}_{\text{SEV}}) \cdot d = [\exp(\hat{\mu}_{\text{SEV}})]^2 Ver(\hat{\mu}_{\text{SEV}}) \quad (10.19)$$

が、導き出される。

Weibull $\hat{\beta}$ の分散は、 $\hat{\beta} = 1/\hat{\sigma}_{\text{SEV}}$ の関係から、

$$\frac{\partial \hat{\beta}}{\partial \hat{\sigma}_{\text{SEV}}} = \frac{-1}{\hat{\sigma}_{\text{SEV}}^2} \quad (10.20)$$

となるので、合成分散の一般式から

$$Var(\hat{\beta}) = \frac{-1}{\hat{\sigma}_{\text{SEV}}^2} Var(\hat{\sigma}_{\text{SEV}}) \frac{-1}{\hat{\sigma}_{\text{SEV}}^2} = \frac{Var(\hat{\sigma}_{\text{SEV}})}{(\hat{\sigma}_{\text{SEV}}^2)^2} \quad (10.21)$$

計算される。

パラメータが複数の場合

時間 t における故障確率は、 $y = \ln(t)$ として、最小極値分布の累積分布関数から

$$\hat{F}(y) = 1 - \exp\left\{-\exp\left(\frac{y - \hat{\mu}_{\text{SEV}}}{\hat{\sigma}_{\text{SEV}}}\right)\right\} \quad (10.22)$$

となる。ここで、

$$u = \frac{y - \hat{\mu}_{\text{SEV}}}{\hat{\sigma}_{\text{SEV}}} \quad (10.23)$$

の分散 $Var(u)$ を求めたい。パラメータ $\hat{\mu}_{\text{SEV}}$ および $\hat{\sigma}_{\text{SEV}}$ の分散共分散行列を、

$$\Sigma = \begin{bmatrix} Var(\hat{\mu}_{\text{SEV}}) & Cov(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) \\ Cov(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) & Var(\hat{\sigma}_{\text{SEV}}) \end{bmatrix} \quad (10.24)$$

としたときに、 u を $\hat{\mu}_{\text{SEV}}$ および $\hat{\sigma}_{\text{SEV}}$ で偏微分すると

$$d_1 = \frac{\partial u}{\partial \hat{\mu}_{\text{SEV}}} = \frac{-1}{\hat{\sigma}_{\text{SEV}}}, \quad d_2 = \frac{\partial u}{\partial \hat{\sigma}_{\text{SEV}}} = \frac{-(y - \hat{\mu}_{\text{SEV}})}{\hat{\sigma}_{\text{SEV}}^2} = \frac{-u}{\hat{\sigma}_{\text{SEV}}} \quad (10.25)$$

が得られる． $\mathbf{d} = [d_1, d_2]^T$ と縦ベクトルとして，

$$\begin{aligned} \text{Var}(\hat{u}) &= \mathbf{d}^T \boldsymbol{\Sigma} \mathbf{d} \\ &= [-1/\hat{\sigma}_{\text{SEV}} \quad -u/\hat{\sigma}_{\text{SEV}}] \begin{bmatrix} \text{Var}(\hat{\mu}_{\text{SEV}}) & \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) \\ \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) & \text{Var}(\hat{\sigma}_{\text{SEV}}) \end{bmatrix} \begin{bmatrix} -1/\hat{\sigma}_{\text{SEV}} \\ -u/\hat{\sigma}_{\text{SEV}} \end{bmatrix} \\ &= \left(\frac{1}{\hat{\sigma}_{\text{SEV}}} \right)^2 \left[\text{Var}(\hat{\mu}_{\text{SEV}}) + 2u \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) + u^2 \text{Var}(\hat{\sigma}_{\text{SEV}}) \right] \end{aligned} \quad (10.26)$$

となる．

最小極値分布のパーセント点は， $y_p = \hat{\mu}_{\text{SEV}} + \ln[-\ln(1-P)]\hat{\sigma}_{\text{SEV}}$ であつた．分散 $\text{Var}(y_p)$ を求めるために， y_p を $\hat{\mu}_{\text{SEV}}$ および $\hat{\sigma}_{\text{SEV}}$ で偏微分すると，

$$d_1 = \frac{\partial y_p}{\partial \hat{\mu}_{\text{SEV}}} = 1, \quad d_2 = \frac{\partial y_p}{\partial \hat{\sigma}_{\text{SEV}}} = \ln[-\ln(1-P)] \quad (10.27)$$

が得られる， $\mathbf{d} = [d_1, d_2]^T$ と縦ベクトルとして， $u_p = \ln\{-\ln(1-P)\}$ とおくと

$$\begin{aligned} \text{Var}(y_p) &= \mathbf{d}^T \boldsymbol{\Sigma} \mathbf{d} \\ &= [1 \quad u_p] \begin{bmatrix} \text{Var}(\hat{\mu}_{\text{SEV}}) & \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) \\ \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) & \text{Var}(\hat{\sigma}_{\text{SEV}}) \end{bmatrix} \begin{bmatrix} 1 \\ u_p \end{bmatrix} \\ &= \left[\text{Var}(\hat{\mu}_{\text{SEV}}) + 2u_p \text{Cov}(\hat{\mu}_{\text{SEV}}, \hat{\sigma}_{\text{SEV}}) + u_p^2 \text{Var}(\hat{\sigma}_{\text{SEV}}) \right] \end{aligned} \quad (10.28)$$

となる．

10.3 知る人ぞ知るデルタ法

寿命データ解析では

寿命データの Excel を用いた統計解析で，95%信頼区間を求めるためには，合成分散の一般式（デルタ法）をどうしても使わなければ解決できない状況に遭遇する．先進的な統計ソフトでは，内部の計算手順に組み込められているが，解説書には出てこない．

統計ソフトのユーザに徹すれば，他者への統計的な説明をする必要がないかも知れない．ある学問領域で多くの人たちと知識を共有し，普及しようとしたときには，その統計的な背景について平易な解説が，相互理解を図るためには必要と思われる．

奥野監訳本には，デルタ法による信頼区間の計算式が天下りの的に記されているだけで，デルタ法についての解説はない．これが，「知る人ぞ知るデルタ法」の状況である．多くの書物，Web サイトにデルタ法についての解説を見出すことはできても，ほとんどが1変量に限られ

ている。

回帰直線の推定値の分散での例

回帰分析を行い、回帰式の 95%信頼区間を求めることは、どのような統計教科書でも、その計算方法が示されている。ただし、多くの教科書では、「これこれしかじかの式で与えられる」との天下りの説明が多い。

正確に説明しようとする教科書では、回帰式を

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = \bar{y} + \hat{\beta}_1 (x - \bar{x}) \quad (10.29)$$

において、 \bar{y} と $\hat{\beta}_1$ の相関が 0 となることを使って、分散の加法性によって、

$$Var(\hat{y}) = Var(\bar{y}) + (x - \bar{x})^2 Var(\hat{\beta}_1) \quad (10.30)$$

となることを解説している。

中村訳 (1968) の「2.4 節 マトリクス展開を用いた y の分散」には、式 (10.11) 式と同様な説明があるが、

$$Var(\hat{y}_i) = Var(\hat{\beta}_0) + 2x_i Cov(\hat{\beta}_0, \hat{\beta}_1) + x_i^2 Var(\hat{\beta}_1) \quad (10.31)$$

を「マトリクスで書けば」との説明となっていて、デルタ法の説明とはなっていない。

デルタ法の統計検定 1 級での扱い

幸い、Meeker ら (1998) の付録 B.2 Statistical Error Propagation – The Delta Method には、本章で説明したと同様にデルタ法についての 2 ページ分の解説がある。

なお、本章では数値例は割愛したが、第 6 章では、多くの数値例を示したので、本章と合わせて、理解を深めてもらいたい。各種の推定値に対して、標準誤差を示すことは統計解析の必須事項である。また、統計的な判断を行う際に、推定値の 95%信頼区間を示すことも必須である。

奥野監訳本で提示されていた、多くの 95%信頼区間を示した上での統計的な判断が、現代の信頼性工学の分野では無視されているかのごとくである。奥野監訳本は、QC 検定 1 級の出題範囲にも含まれていない統計が多用されている。分散の加法性が全く適用できない場合のデルタ法についても「知る人ぞ知る」状況を痛切に感じている。寿命データの解析を改めて学習したいと思われる人たちに、本書が少しでも役立ってもらいたいと願っている。

統計検定 1 級の出題範囲の項目 (学習しておくべき用語) 例 に「ワイブル分布」は含まれてはいるが、「最小極値分布」は含まれていない。非線形に関する用語も見いせないの、パラメータに関して非線形の場合の合成分散の算出も出題範囲外である。「デルタ法」は、大項目：統計的推測 (検定)、小項目：漸近的性質など、大項目：理工学分野、小項目：漸近理論、に含まれてはいる。悲しいことに、「最尤推定量」は散見するものの、対数尤度を最大化する

ためのニュートン・ラフソン法に関連する「最尤法」も見出すことができない。

「奥野監訳本で提示されていた、多くの95%信頼区間を示した上での統計的な判断が、現代の信頼性工学の分野では無視されているかのごとくである。」と述べたが、「信頼性工学」の統計を下支えする統計検定1級の出題範囲にも入っていないことは、本当に悲しいことであり、「信頼性工学」の関係者が、「寿命データの統計解析」を学習しようにも学習するための日本語の教科書が、奥野監訳本しかなく、平易な解説書すらないことが、憂うべき事柄である再認識した。

11. プロファイル尤度を用いた 95%信頼区間

11.1 プロファイル尤度による 95%信頼区間

最近追加されたの JMP の「寿命の一変量」および「寿命の二変量」では、95%信頼区間の計算法で **Wald** または **尤度** が選択できるようになっている。以前からある「生存時間分析」では、**尤度** を用いた結果のみが示されていた。

推定されたパラメータの信頼区間の計算に際して、分散共分散行列の対角要素の平方根を標準誤差とし、正規分布を仮定した 95%信頼区間は、近似 (Wald) 法として知られている。正確な 95%信頼区間の計算法として、尤度を用いた方法があるが、さらなる逐次計算が必要となるために多くの統計ソフトでは、敬遠されている。

知る人ぞ知る「プロファイル尤度」について Meeker ら (1998) に丁寧な説明がなされている。筆者も自ら計算したことがなかったので、JMP のプロファイル尤度による 95%信頼区間が、どのような方法で求められているのかを詳細に示すことにした。

ディーゼル発電機ファンについて信頼区間の計算法を「尤度」とするのは、「寿命の一変量」の設定画面で、「信頼区間の方法」で **Wald** を **尤度** に変更する。

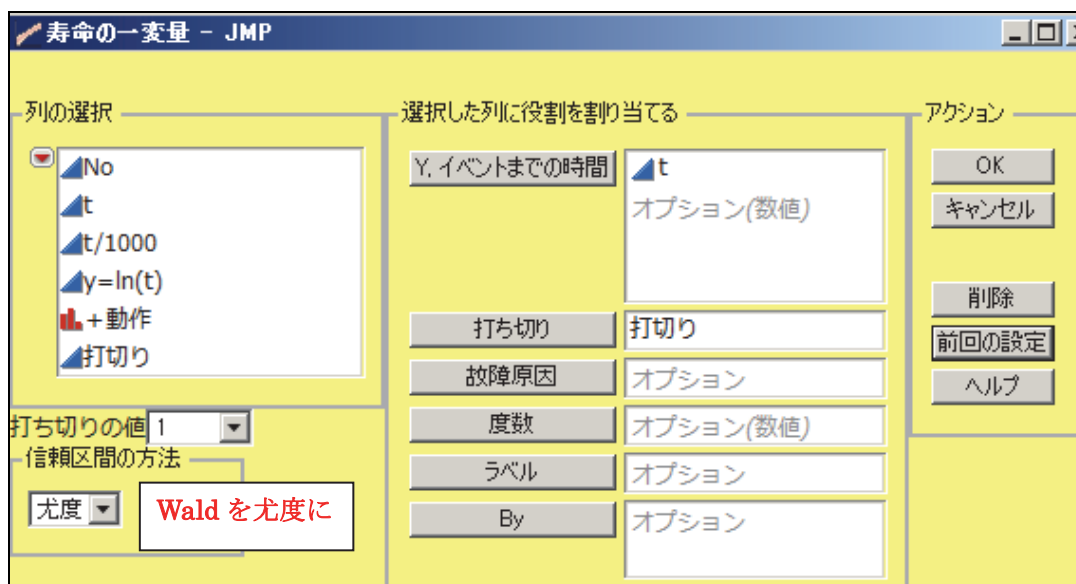


図 11.1 信頼区間の計算方法を尤度に変更

この設定で、パラメータの推定とその信頼区間を計算すると、表 11.1 の結果を得る。Wald 法での結果を表 3.7 (再掲) に示す。Wald 法での 95%信頼区間は、推定値に対して、標準誤差の 1.96 倍をプラス・マイナスしているのので、(下側 95%+下側 95%) / 2 を計算した時に、この結果が推定値と一致すれば、Wald 法と判定され、異なれば別の方法での計算と判定される。

表 11.1 ワイブル・パラメータの尤度による信頼区間の計算 (尤度法)

パラメータ	推定値	標準誤差	下側95%	上側95%	(L95+U95)/2	推定値との差
位置	10.1772	0.4659	9.5201	11.5720	10.5461	-0.3689
尺度	0.9448	0.2394	0.6032	1.6502	1.1267	-0.1819
Weibull α	26296.8452	12251.4283	13631.2390	106086.921	59859	-33562
Weibull β	1.0584	0.2683	0.6060	1.6579	1.1320	-0.0736

位置 μ の場合「推定値との差」が、-3.6389 なので、Wald 法ではないことが明らかである。尺度 σ についても -0.1819 と 0 となっていない。位置 μ の場合「推定値との差」が、-3.6389 なので、Wald 法ではないことが明らかである。

表 3.7 は Wald 法による 95%信頼区間であった。位置 μ よび尺度 σ の「推定値との差」が 0 なので、Wald 法での計算であることが確認された。

表 3.7 (再掲) ワイブル・パラメータの推定 (Wald 法)

パラメータ	推定値	標準誤差	下側95%	上側95%	(L95+U95)/2	推定値との差
位置	10.1772	0.4659	9.2641	11.0903	10.1772	0
尺度	0.9448	0.2394	0.4755	1.4141	0.9448	0
Weibull α	26296.8452	12251.4283	10552.0697	65534.4483	38043	-11746
Weibull β	1.0584	0.2683	0.7072	2.1031	1.4052	-0.3468

下限と上限の和の 1/2 と推定値との差 (0 か否かで判別)
Weibull パラメータは、(位置・尺度) から計算しているので判定は不可

ただし、Weibull α および Weibull β は、「推定値との差」が 0 ではないので Wald 法による計算がなされていない。Weibull α の 95%信頼区間は、最小極値分布の位置 μ の信頼区間から、

$$L95\% \alpha = \exp(9.2641) = 10552$$

$$U95\% \alpha = \exp(11.0903) = 65534$$

と計算されることは、3.6 節で解説した。Weibull β は、最小極値分布の尺度の逆数、

$$L95\% \beta = 1/1.4141 = 0.7072,$$

$$U95\% \beta = 1/0.4755 = 2.1031$$

で計算されている。このことは、ワイブル分布を前提にした最尤法を行ったとしても最小極値分布の標準誤差から 95%信頼区間が計算されていることになる。

ワイブル分布を仮定した場合の分散共分散行列は、表 7.1 で示したように、時間 t を千時間単位 t_{1000} として次のように

$$\text{Weibull : } \mathbf{Z} = \begin{bmatrix} -0.0194 & -0.7198 \\ -0.7198 & -40.5510 \end{bmatrix}, \quad \text{Weibull : } \boldsymbol{\Sigma} = (-\mathbf{Z})^{-1} = \begin{bmatrix} 150.0975 & -2.6645 \\ -2.6645 & 0.0720 \end{bmatrix}$$

計算されていた。最小極値分布を仮定した場合は、表 6.5 に示したように

$$\text{最小極値 : } \mathbf{Z} = \begin{bmatrix} -13.4437 & 21.2070 \\ 21.2070 & -50.8951 \end{bmatrix}, \quad \text{最小極値 : } \boldsymbol{\Sigma} = (-\mathbf{Z})^{-1} = \begin{bmatrix} 0.2171 & 0.0904 \\ 0.0904 & 0.0573 \end{bmatrix}$$

であった。ワイブル分布と最小極値分布の関係は、正規分布と対数正規分布の関係であることはすでに述べた、

ワイブル分布の $Var(\hat{\alpha}) = 150.0975$ から、最小極値分布の $Var(\hat{\mu}_{SEV}) = 0.2171$ をデルタ法で求めてみよう。 $\hat{\mu}_{SEV} = \ln(\hat{\alpha})$ の関係から

$$d = \frac{\partial \hat{\mu}_{SEV}}{\partial \hat{\alpha}} = \frac{1}{\hat{\alpha}}$$

なので、

$$\begin{aligned} Var(\hat{\mu}_{SEV}) &= d \cdot Var(\hat{\alpha}) \cdot d \\ &= \left(\frac{1}{26.2968} \right)^2 \times 150.0975 = 0.2171 \end{aligned}$$

が推定される。

同様に、ワイブル分布の $Var(\hat{\beta}) = 0.0720$ から、最小極値分布の $Var(\hat{\sigma}_{SEV}) = 0.0573$ を求めてみよう。 $\hat{\sigma}_{SEV} = 1/\hat{\beta}$ の関係から

$$d = \frac{\partial \hat{\sigma}_{SEV}}{\partial \hat{\beta}} = \frac{-1}{\hat{\beta}^2}$$

なので、

$$\begin{aligned} Var(\hat{\sigma}_{SEV}) &= d \cdot Var(\hat{\beta}) \cdot d \\ &= \left(\frac{-1}{1.0584^2} \right)^2 \times 0.0720 = 0.0573 \end{aligned}$$

が推定される。

奥野監訳本では、ワイブル分布を仮定した分散共分散行列を最尤法で求めた上で、デルタ法により最小極値分布の分散共分散行列を計算して 95%信頼区間を求め、再度ワイブル・パラメータの 95%信頼区間に戻している。

最小極値分布を仮定した Wald 法と、尤度を用いた正確な 95%信頼区間には、明らかな差がある。

Wald $\alpha = 10,552$, Wald $\tilde{\alpha} = 65,534$
 尤度 $\alpha = 13,631$, 尤度 $\tilde{\alpha} = 106,086$

Wald $\beta = 0.7072$, Wald $\tilde{\beta} = 2.1031$
 尤度 $\beta = 0.6060$, 尤度 $\tilde{\beta} = 1.6579$

表 11.1 の尤度を用いた信頼区間の推定方法については、JMP のマニュアルでは全く触れられていない。以下に示す Meeker ら(1998) の Appendix B.6.6 (p627) を参考にしつつも解いてみる。

尤度を用いた信頼区間（プロファイル尤度）は、対数尤度の - 2 倍を用いた尤度比検定の考え方に準じている。「寿命の一変量」の標準の解析結果から更に先に深堀すると、「尤度等高線の表示」が現れる。

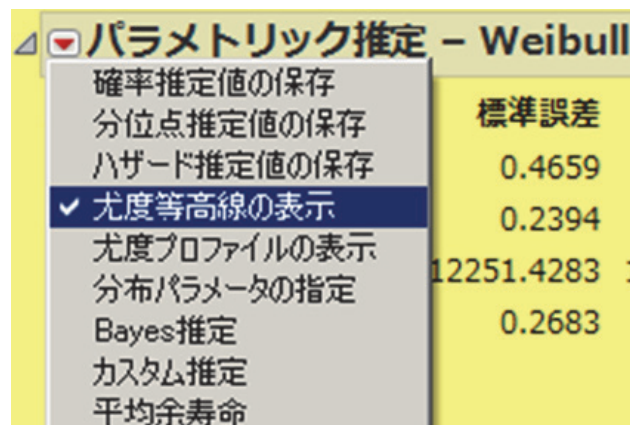


図 11.2 尤度等高線 プルダウンメニュー

図 11.3 に尤度の等高線図を示す。この結果に筆者も当惑した。等高線は「尤度」が現れるものと期待していたのだが、(0.500, 0.7500, ..., 0.990) とは何か。確率のように見える。

対数尤度の等高線

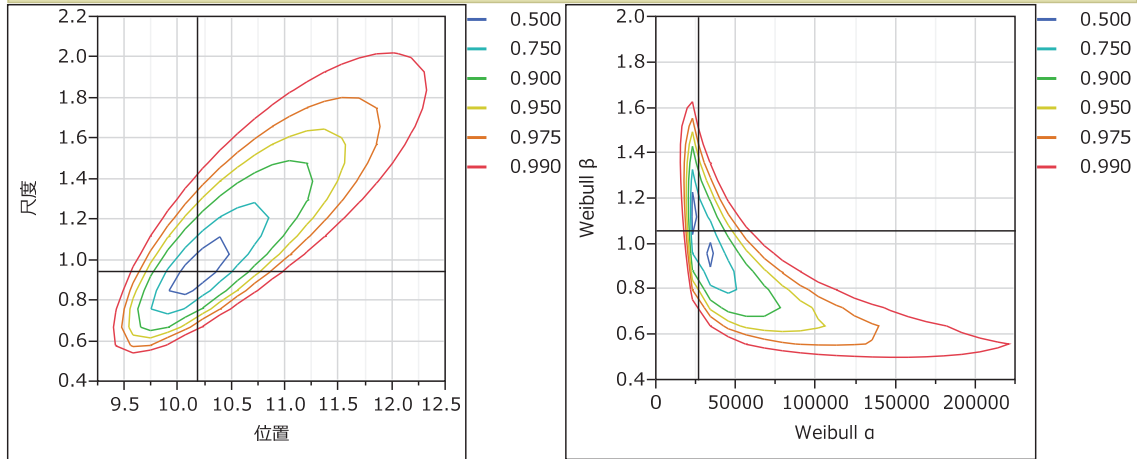


図 11.3 尤度の等高線：カイ 2 乗分布の下側確率の等高線

実際には、カイ 2 乗分布の下側確率が表示されているのだが「信頼域」との説明が統計的詳細でされている。この図からプロファイル尤度による 95%信頼区間の算出方法は、読み取るとることができない。

11.2 パラメータの格子点での対数尤度

推定されたパラメータの前後を指定した範囲と分割数を指定して、メッシュ状に対数尤度を計算する機能があることは、他の JMP プラットホームにあるので、これを活用して、さらに検討することとした

図 11.2 のプルダウンメニューの「尤度プロファイルの表示」で、推定されたパラメータについてパラメータごとのプロファイル曲線が出力されるので、最小極値分布およびワイブル分布の パラメータについての尤度プロファイルを図 11.4 に示す。

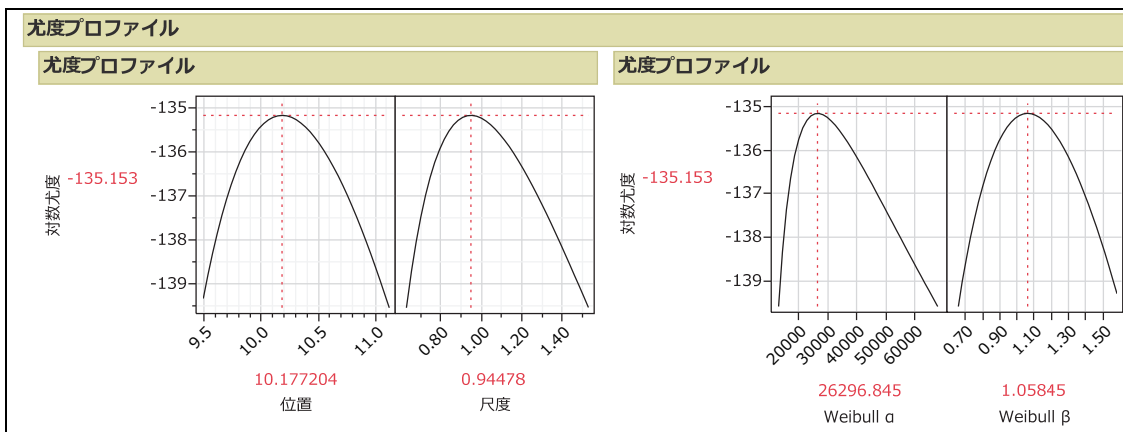


図 11.4 パラメータの尤度プロファイル

この尤度プロファイルから、尤度に基づく 95%信頼区間が求められていると思い、試行錯誤しても表 11.1 の結果が得られない。そこで、Meeker ら(1998)の次の記述を参考に、さらに検討することにした。

Confidence Interval for μ

The profile likelihood for μ is

$$R(\mu) = \max_{\sigma} \left[\frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})} \right]. \quad (8.2)$$

The interval over which $R(\mu) > \exp[-\chi^2_{(1-\alpha;1)}/2]$ is an approximate $100(1 - \alpha)\%$ confidence interval for μ . The general theory for likelihood confidence intervals is

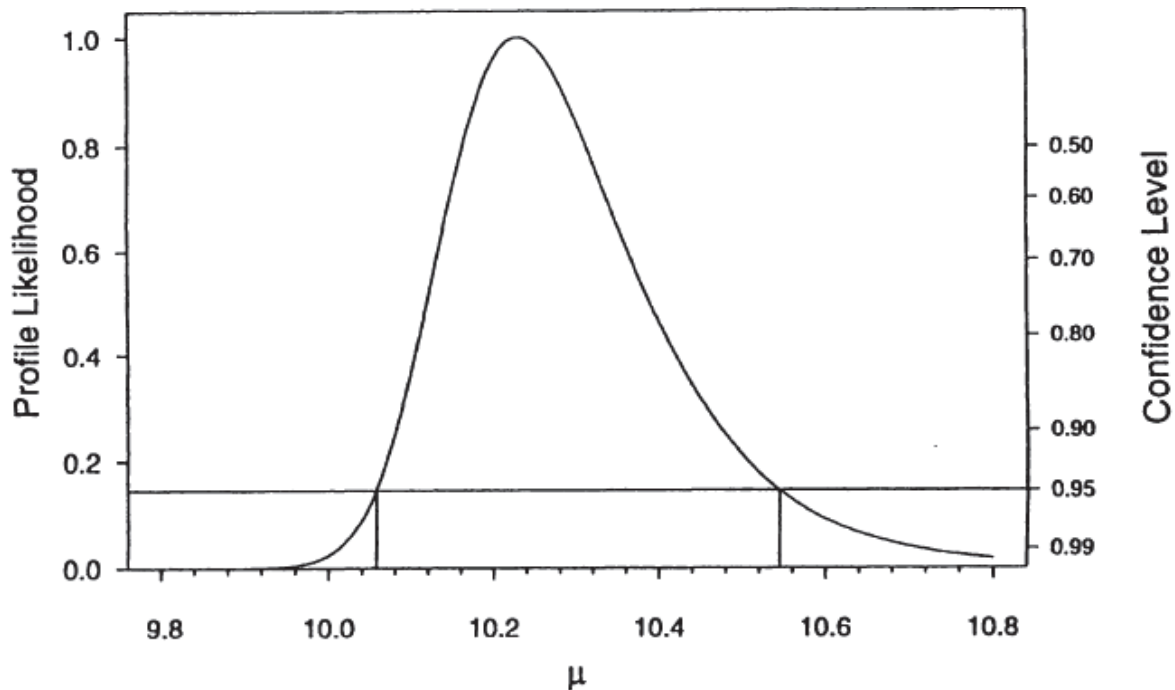


Figure 8.5. Weibull profile likelihood $R(\mu)$ for the shock absorber data.

given in Appendix Section B.6.6. Intuitively, we want to include all values of μ that have high likelihood. Using (8.2), for every fixed value of μ , find the point of *highest* relative likelihood by maximizing the relative likelihood with respect to σ . This gives the profile likelihood level for that value of μ . Values of μ with high profile likelihood are more plausible than those with lower values of profile likelihood.

Meeker and Escobar(1998),p179-180.

この記述から、極値分布の最尤推定量としての位置 $\hat{\mu}$ および尺度 $\hat{\sigma}$ を中心に、位置 μ および尺度 σ をメッシュ状に変化させて対数尤度 $\ln L(\mu, \sigma)$ を計算が必要であることが読み取れる。

そこで、推定値を中心にしたメッシュ状の格子点に対する対数尤度を JMP データセットに出力し、位置 μ および尺度 σ に関する対数尤度の等高線図を表示することにした。

図 11.5 に示すように、「尤度プロファイル」のプルダウンメニューに、「因子グリッドの設定」、「グリッドテーブルの出力」があるので、これを使うことにした。

このメニューの「因子の設定」で、 101×101 のメッシュ状の格子点を設定し、対数尤度を JMP データセット出力した。

出力された対数尤度の JMP データセットで「等高線図」を用い、対数尤度の等高線を作成した結果を図 11.6 に示す。対数尤度が最大 -135.153 となるのは、最小極値分布のあてはめでは、位置 $\hat{\mu} = 10.1772$ 、尺度 $\hat{\sigma} = 0.9448$ で、等高線図上の点線の交叉点で示されている。

尤度比検定では、 -2 倍の対数尤度の差が、自由度 1 のカイ 2 分布に従うことを用いているので、 95% 信頼区間は、対数尤度の最大値の 135.153 からカイ 2 乗分布の下側 95% 点 3.8415 の $1/2$ 、 1.921 を加えた -137.073 が 95% 信頼区間の等高線となる。

新たな課題は、この等高線図からある位置 μ_0 を固定した時に、尺度 σ に関して対数尤度が最大となるような位置 $\hat{\mu}_\sigma$ を得ることである。ある $\mu = 11.22$ を通る垂直線を引いたときに、尺度 σ を動かし場合の対数尤度のピークは、ある対数尤度の等高線の接線の位置となる。

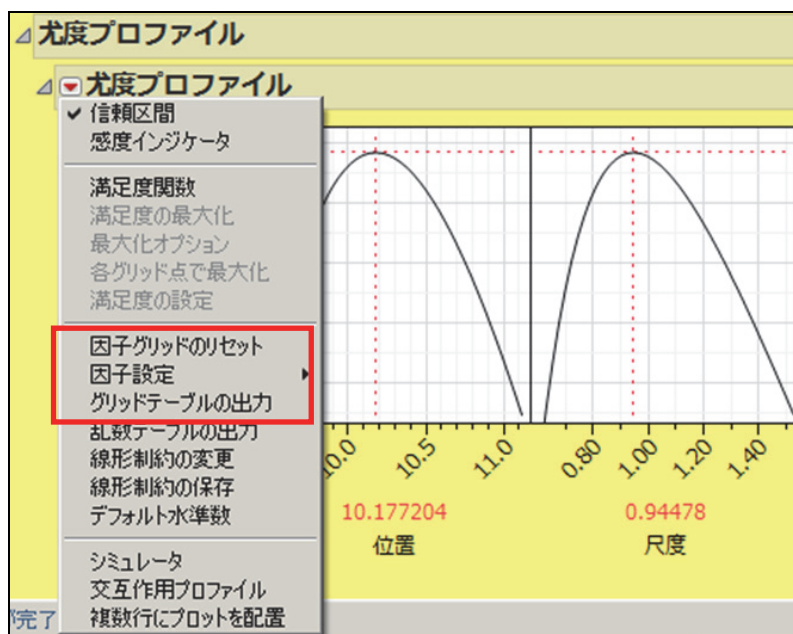


図 11.5 パラメータの尤度プロファイルのプルダウンメニュー

対数尤度の等高線図

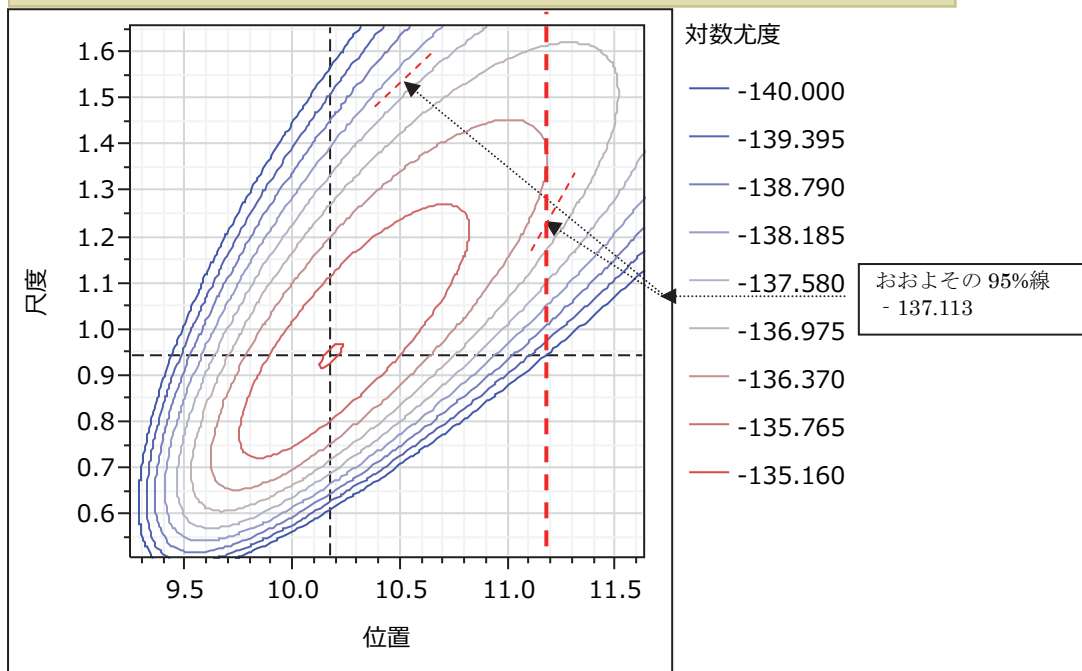


図 11.6 対数尤度の等高線図

表 11.2 に位置 $\mu = 11.22$ を固定し、尺度 σ を変化させたときの対数尤度がピークとなるのは、 $\ln L = -136430$ で、その時の尤度の最大値からの差は、1.277 となる。この 2 倍 $2 \times 1.277 = 2.554$ が自由 1 のカイ 2 乗分布に従うことから下側確率は 0.890 が求められる。

表 11.2 位置を固定し尺度を動かしたときの対数尤度のピークの同定

	位置	尺度	Weibull α	Weibull β	対数尤度	最大値か らの差	最大値か らの比	下側確率	尤度の ピーク
7544	11.220	1.328	74608	0.753	-136.477	1.324	0.266	0.896	•
7545	11.220	1.340	74608	0.746	-136.460	1.307	0.271	0.894	•
7546	11.220	1.352	74608	0.740	-136.446	1.293	0.274	0.892	•
7547	11.220	1.364	74608	0.733	-136.437	1.284	0.277	0.891	•
7548	11.220	1.376	74608	0.727	-136.431	1.278	0.278	0.890	•
7549	11.220	1.388	74608	0.720	-136.430	1.277	0.279	0.890	1
7550	11.220	1.400	74608	0.714	-136.431	1.278	0.279	0.890	•
7551	11.220	1.412	74608	0.708	-136.436	1.283	0.277	0.891	•
7552	11.220	1.424	74608	0.702	-136.445	1.292	0.275	0.892	•
7553	11.220	1.436	74608	0.696	-136.457	1.304	0.272	0.894	•
7554	11.220	1.448	74608	0.691	-136.471	1.318	0.268	0.896	•

表 11.2 で示した，位置 μ の刻みごとの対数尤度のピークとなる点のデータを抜きだして，位置 μ についての下側確率の分布図を 図 11.7 に示す．この図で，確率が 0.95 となるような位置 μ が，プロファイル尤度法における 95%信頼区間となる．表 11.1 で求められた対数尤度による 95%信頼区間 (5.5201, 11.5720) を縦線で表し，確率が 0.95 となる水平線との交点に，対数尤度のピーク曲線が通ることが確認される．

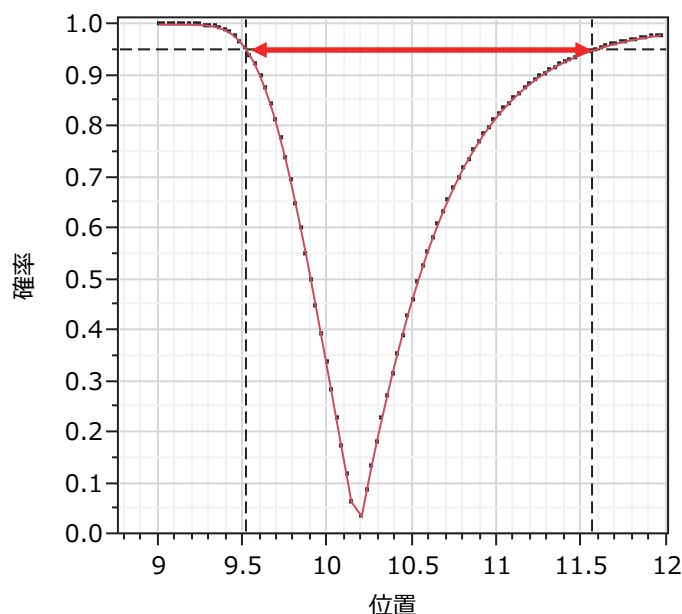


図 11.7 位置 μ に関する対数尤度のピーク点での下側確率

11.3 カイ 2 乗分布の下側確率によるプロファイル尤度

位置 μ について，プロファイル尤度での 95%信頼区間の算出法を前節で示したので，尺度 σ をカイ 2 乗分布の下側確率の等高線を用いた求め方を示す．表 11.2 を尺度・位置の順にソートすると，「尤度のピーク」は，自動的に尺度 σ についての尤度のピークとなる．

ある尺度 $\sigma=1.652$ を固定して，位置 μ を動かしたときに，ピークとなる対数尤度は，-137.080 である．対数尤度の最大値は-135.153 であるので，その差は 1.927 となる．差の 2 倍 $1.927 \times 2 = 3.854$ であり，カイ 2 乗分布の下側確率は 0.9504 となる．対数尤度では，ピークであるが，下側確率では，ボトムとなっている．

位置 μ と尺度 σ を X 軸と Y 軸として，下側確率についての等高線図を 図 11.8 示す．外側が 0.95 となるように等高線を 0.05 刻みで最小が 0.20 となるように設定した．図 11.8 上で尺度 $\sigma=1.652$ を固定し，位置 μ を左端から右方向に位置をずらして，行くとおおよそ 0.950 の等高線上で下側確率がボトムとなることが確認できる．下側確率が 0.950 となる尺度 σ が，95%

表 11.3 尺度を固定し位置を動かしたときの対数尤度のピークの同定

	位置	尺度	Weibull α	Weibull β	対数尤度	最大値か らの差	最大値か らの比	下側確率	尤度の ピーク
9772	11.250	1.652	76880	0.605	-137.099	1.946	0.143	0.9515	•
9773	11.280	1.652	79221	0.605	-137.089	1.936	0.144	0.9509	•
9774	11.310	1.652	81634	0.605	-137.082	1.929	0.145	0.9505	•
9775	11.340	1.652	84120	0.605	-137.080	1.927	0.146	0.9504	1
9776	11.370	1.652	86682	0.605	-137.082	1.929	0.145	0.9505	•
9777	11.400	1.652	89322	0.605	-137.087	1.934	0.145	0.9508	•
9778	11.430	1.652	92042	0.605	-137.097	1.944	0.143	0.9513	•

信頼区間の上限となる。

尺度 $\sigma=0.608$ に固定した場合は、対数尤度は-137.001 でピークとなり、下側確率は 0.9454 となる。この図から、尺度 σ についての 95%信頼区間は、位置 μ を水平方向に動かした場合のカイ階 2 乗分布の下側確率 0.950 との接線となる尺度 σ を読み取ることになる。

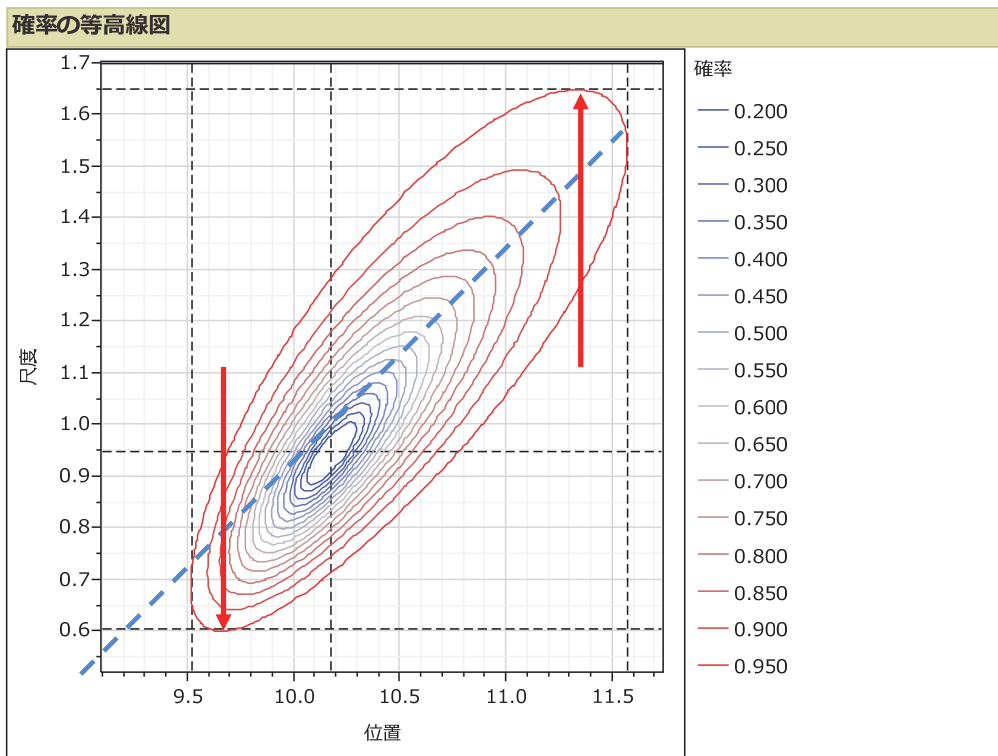


図 11.8 最小極値分布での下側確率の等高線図

図 11.8 の等高線図では、尺度 σ の正確な 95%信頼区間は読み取れないので、尤度がピークとなった行を全て抜き出して作成した対数尤度のプロファイル曲線を 図 11.9 示す。この図に、表 11.1 で得られた尤度による 95%信頼区間 (0.6032, 1.6502) を縦線で入れ、下側確率が 0.950

となる横線を入れると、プロファイル曲線上で交叉することが確認できる。

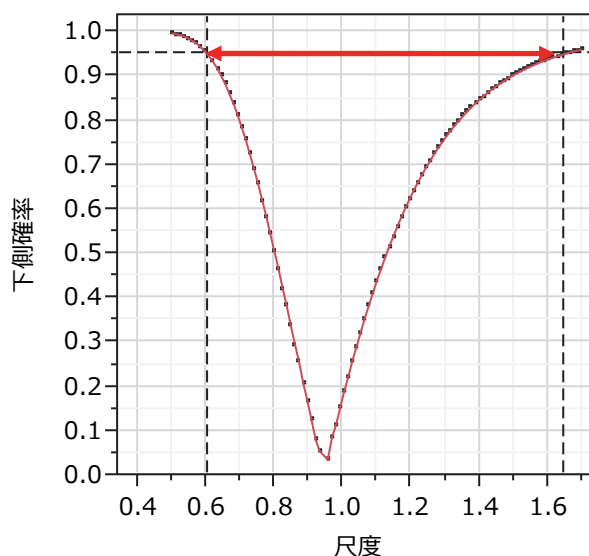


図 11.9 尺度 σ の対数尤度のピーク点での下側確率の推移

11.4 ワイブル分布での下側確率の等高線図

寿命データの解析は、ワイブル分布が標準的に使用されている。奥野監訳本でもワイブル分布を前提にした最尤法が展開されていた。最尤法は、どのような分布を仮定しようとも解析手順は同じであるが、各種の信頼区間の計算に対しては、時間について対数を取った最小極値分布の分散共分散行列を使って、ワイブル・パラメータの場合に換算してきた。

図 11.3 で示したように Weibull α と Weibull β についての対数尤度からカイ 2 乗分布の下側確率を算出した等高線図は、左下に潰れていて、最小極値分布の楕円形の等高線とは対照的である。なお、等高線が負の方向になっているのは、最小極値分布の尺度 σ の 95%信頼区間と逆数の関係からである。

表 11.2 および表 11.3 に示した Weibull α を X 軸、Weibull β を Y 軸、カイ 2 乗分布の下側確率の等高線図を図 11.10 に示す。等高線は、0.950 から 0.100 刻みに 0.050 と設定した。この図に、表 11.1 で得られた尤度による 95%信頼区間、

Weibull α 95%信頼区間 : (13,631, 106,086)

Weibull β 95%信頼区間 : (0.6060, 1.6579)

を縦線と横線で上書きした。それらの線とカイ 2 乗分布の下側確率 0.950 等高線とが接していることが確認される。

このような歪んだ対数尤度の等高線の形状は、Excel のソルバーで最尤解を求めようとしたときに、与えた初期値によっては収束せずに、発散してしまう原因となる。本書では、多くの人たちがワイブル分布に慣れ親しんでいることから、ワイブル分布を前提にした Excel での解析を主体にしてきたが、初期値の与え方によっては解がしばしば求められないことを経験した。

確率の等高線図

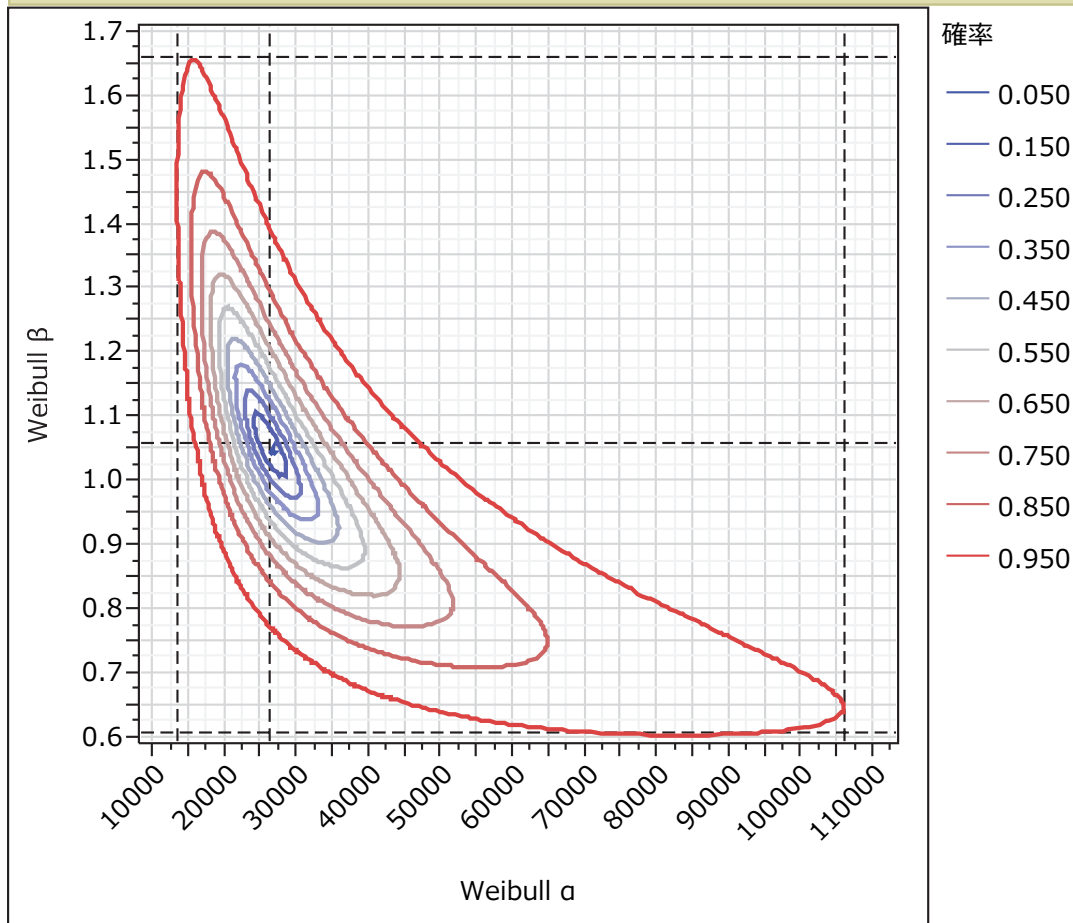


図 11.10 ワイブル・パラメータに対する下側確率の等高線図

残念ながら Web での検索では、尤度による 95%信頼区間の事例が見出せない。臨床試験の生存時間解析に関する成書には、最小極値分布に関する記述は見当たらない。唯一、奥野監訳本には、丁寧な最小極値分布の説明、多くの事例での適用例が示されているのが救いであった。

12. JMP による対数尤度関数の偏微分

12.1 微分した式

寿命データの統計解析を丁寧に行なおうとすると、典型的な事例に対して「手計算」による手順の解説が不可欠である。そのために、各種の対数尤度関数の偏微分の結果をこれまで示してきた。この結果が正しいかは、統計ソフトから得られる分散共分散行列との対比によって検証をしてきた。

その検証過程で、計算結果が一致しないとのトラブルの連続であった。特に符号の付け間違い、数式を Excel で入力する際の構文ミスなどが多発した。これらを乗り越えるために JMP の微分機能を活用して、間違いの是正を図ってきた。

最近の統計ソフトでは、非線形回帰を行う際にも、偏微分式を自ら指定する必要性はなくなっている。これは、統計ソフトの中で、自動微分の機能が組み込まれているからである。もちろん JMP でも対応し、Excel のソルバーにも組み込まれている。最尤法による寿命データの解析においても、対数尤度関数の偏微分も縁の下の方のごとく、目に見えないところで活躍している。

このことが、統計ソフトをブラック・ボックスと認識しつつも、頼りにしてしまうことになる。JMP では、自動微分の機能が画面操作でユーザに公開されているので、本書で取り上げた偏微分に即して、その操作法と活用方法について解説する。

JMP で計算式を作成する計算式エディタに 図 12.1 に示すように「微分した式」を求める機能がある。

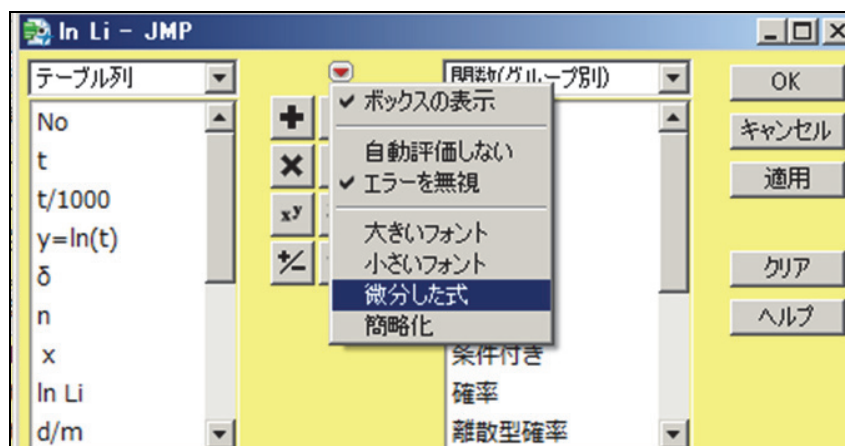


図 12.1 計算式の微分

次に示す最小極値分布の対数尤度,

$$\ln L_i = \left[\delta_i \left(-\ln(s) + \frac{y_i - (m + \beta_1 x_i)}{s} \right) - \exp \left(\frac{y_i - (m + \beta_1 x_i)}{s} \right) \right] n_i$$

のパラメータによる微分式を求めたいとする. JMP の計算機エディタで作成した対数尤度式は,

である. パラメータとした (\mathbf{m} , $\mathbf{b1}$, \mathbf{s}) で, 1 階の微分, 2 階の微分を求めたい. 微分したいパラメータを選択して, 「微分した式」を選択することにより微分式が得られる. (\mathbf{m}) について微分すると, 次のように計算式が現れる. なお, パラメータの初期値として本書では最尤推定量としているが, どのような値でも計算が可能である.



図 12.2 JMP による \mathbf{m} についての偏微分式

12.2 微分した式による数値計算

JMP の計算式では, 指定したパラメータについての表 12.1 に示すように計算結果が得られるので, Excel での計算結果と対比しつつ, 相互チェックにより検証が行える.

表 12.1 JMP による偏微分の計算結果 (抜粋)

	No	t	t/1000	y=ln(t)	δ	n	x	ln Li	d/m	d/b1	d/s	d/mm	
▼	1	1	5000	5.000	8.517	0	30	10	-0.1735	0.2474	2.4743	-1.2750	-0.3529
○	2	2	1298	1.298	7.169	1	1	40	-3.8988	-1.4058	-56.2308	4.5334	-0.0293
○	3	3	1390	1.390	7.237	1	1	40	-3.8026	-1.4037	-56.1464	4.3874	-0.0323

本書で示した対数尤度の偏微分式は、Excel 計算と JMP による計算結果が一致すること確認した。これらの計算に用いた Excel および JMP ファイルも公開するので、読者が自ら追試を行いつつ、筆者にとっても難解だった加速試験の統計解析が適切に行われるようになることを願っている。

偏微分式の相互検証

JMPの結果(抜粋)

No	t	t/1000	y=ln(t)	δ	n	x	ln Li	d/m(b0)	d/b1	d/s	d/mm	
1	1	5000	5.000	8.517	0	30	10	-0.1735	0.2474	2.4743	-1.2750	-0.3529
2	2	1298	1.298	7.169	1	1	40	-3.8988	-1.4058	-56.2308	4.5334	-0.0293
3	3	1390	1.390	7.237	1	1	40	-3.8026	-1.4037	-56.1464	4.3874	-0.0323

↙

変数x	件			
群	温度	y=ln(t)	数	δ
1	10	8.517	30	0
2	40	7.169	1	1
2	40	7.237	1	1

↕

Excelの結果(抜粋)

ln Li	β_0^{\wedge}	β_1	σ	$\beta_0^{\wedge 2}$	$\beta_0\beta_1$
-0.1734	0.2474	2.4741	-1.2749	-0.3529	-3.5292
-3.8986	-1.4059	-56.2364	4.5337	-0.0293	-1.1734
-3.8024	-1.4038	-56.1519	4.3876	-0.0323	-1.2939

完全に一致しないのは、設定したパラメータの精度の問題

2015.3.6 高橋行雄120

図 12.3 相互検証

JMP の計算式では、パラメータに対して数値を与える必要がある。この数値に、最尤法によって推定されたパラメータ、 $m=12.7929$, $s=0.7011$, $b1=-0.0663$ を与えることにより、Excel シート上で最尤解を得た時に偏微分の結果と対比することができる。これらが一致すれば、相互に検証できたことになる。

12.3 JMP の微分関数と最大化関数

JMP は、画面操作の裏で R 言語と同様な JMP スクリプト言語によって統計解析が行われている。微分を行う関数は `derivative()` であり、最尤解をもとめる関数は `maximize()` であり、「JMP スクリプト言語 (JSL)」マニュアルの第 8 章「プログラミング手法」、「その他の数値演算子」に次のような解説がある。このマニュアルは、JMP のヘルプメニュー>ドキュメンテーション>スクリプト言語 (JSL) ガイドによって参照することができる。

第2次導関数をとるには、第3引数としてもう1つ変数を指定します。第2引数と第3引数は、どちらもリストでなければなりません。JMPは、元の式、第1次導関数、そして第2次導関数の順番で結果を含んだリストを返します。

```
second=derivative(3*y*x^2,{x},{x}); show(second);
second = {3 * y * x ^ 2, 6 * x * y, 6 * y}

second=derivative(3*y*x^2,{y},{y}); show(second);
second = {3 * y * (T#1 = x ^ 2), 3 * T#1, 0}

second=derivative(3*y*x^2,{y},{x}); show(second);
second = {3 * y * (T#2 = x ^ 2), 3 * T#2, 6 * x}
```

JMP10 バージョン p237 より抜粋

JMP による寿命データの最尤法による解析の縁の下の力持ちは、`Maximize()` 関数である。筆者も新しい統計手法を習得するために、この関数を使って数値計算を繰り返し、その経験をもとに Excel で数値計算を行ってきた。JMP の画面操作による各種の最尤法を用いた統計計算も、`maximize()` 関数が下支えしているに違いない。

`Maximize` 関数および `Minimize` 関数は、式を最適化（最大化または最小化）する因子の値を求めます。式は、因子の値に対して連続関数でなければいけません。

この関数を呼び出す方法は、次のとおりです。

```
result = Maximize(objectiveExpression,{list of factor names}, <<option(value))
result = Minimize(objectiveExpression,{list of factor names}, <<option(value))
```

`objectiveExpression` は、最適化する対象の式です。式そのもの、または式が保存されているグローバル変数名のどちらかを指定します。

許容値の上限または下限だけを指定する場合は、以下の例のように片側を欠測値にします。

```
{beta} //制限なし
{beta (0,1)} //0以上1以下
{beta (.,1)} //1以下
{beta (0,.)} or {beta (0)} // 0以上
```

因子の値は、数値または行列のいずれかです。

また、以下のオプションも使用できます。デフォルト値は、以下のとおりです。

```
<< tolerance(.00000001) //収束基準  
<< maxIter( 250) //最大反復数  
<< limits() //  
<< Method(NR|BFGS) // Newton-Raphson法またはBFGS更新の準Newton法を指定する
```

関数を呼び出す前に、因子には初期値が割り当てられているものとします。

これらの関数は、複数の局所解を持つ関数において、大域的な最適解を見つけるとは限りません。指定された初期値から、ある局所的な最適解（問題によっては、大域的な最適解かもしれませんが）を求めることしかできません。

現在のところ、戻り値は目的関数の値です。

JMP10 p240 より抜粋

Maximize() 関数は、Excel のソルバーと同様に、分散共分散行列の出力はないので、derivative() 関数で、パラメータに関して 2 階の偏微分式を自ら求めて、Excel と同様に行列計算を行う必要がある。

文 献

- 1) 大橋靖雄・浜田知久馬 (1995), 生存時間解析, SAS による生物統計, 東大出版会.
- 2) 奥野忠一監訳, 柴田義貞, 藤野和健, 鎌倉稔成訳 (1988), 寿命データの解析, 日科技連出版社.
- 3) Kalbfleisch, J.D. and Prentice, R.L. (2002), The Statistical Analysis of Failure Time Data 2nd ed., Wiley.
- 4) 中村慶一訳 (1968), 応用回帰分析(原著 1st ed.), 森北出版.
Draper, R.N. and Smith, H. (1998), Applied Regression Analysis 3rd ed., Wiley.
- 5) 廣野元久 (2000), JMP4 ハンズオンセミナー ～生存時間分析の実行～,
<http://www1.odn.ne.jp/~gengen525/index1.files/jmpv4user.pdf>
- 6) SAS Institute (2012), 品質管理および信頼性/生存時間, SAS Institut Inc.
- 7) SAS Institute (2012), スクリプト言語 (JSL) ガイド, SAS Institut Inc.
- 8) 高橋行雄 (2011), JMP による各種分割実験入門 –変量効果モデルの基礎–,
<http://www.yukms.com/biostat/takahasi2/rec/001.htm>
- 9) 高橋行雄 (2013), 非線形回帰の基礎応用, SAS ユーザー総会論文集, 439-460,
<http://www.sascom.jp/download/pdf/SUGJ2013.pdf>
- 10) Meeker and Escobar (1998), Statistical Method for Reliability Data, Wiley.
- 11) 芳賀敏郎 (2004), 最小 2 乗法, 最尤法, 線形モデル, 非線形モデル,
<http://www.yukms.com/biostat/haga/download/archive/likelihood/Likelihood.pdf>
- 12) 鈴木和幸, 増田明彦, 石田勉 (2009), 信頼性データの解析, 日科技連出版社.
- 13) 広瀬英雄 (2004), 新段階昇圧法を用いた低破壊値の推定法について, 統計数理,
Vol52:175-187. <http://www.ism.ac.jp/editsec/toukei/pdf/52-1-175.pdf>
- 14) 高橋倫也, 志村隆彰 (2004) 「特集 極値理論」について, 統計数理, Vol52:1-4.
<http://www.ism.ac.jp/editsec/toukei/pdf/52-1-001.pdf>
- 15) 清水貴宏 (2011), 加速試験の寿命予測に必要な試験水準の設定と故障データに関する考察,
http://www.i-juse.co.jp/statistics/xdata/reliability2-shimizu_doc.pdf

索引

数字			
1 階の微分	122		
2 階の微分	122		
2 階の偏微分	79		
2 次形式	101		
63.2 パーセント点	41		
95%信頼区間	22		
95%信頼区間	30		
A--F			
Arrhenius ケルビン	56		
Arrhenius 摂氏	48		
Arrhenius 華氏	56		
CHISQ.DIST()	54		
Derivarive()	123		
Devalt サンプル	48		
Excel アドイン	17		
Excel ソルバー	17		
Excel 分散共分散行列	91		
Excel の Weibull 関数	13		
Excel のソルバー	3		
Excel の行列関数	3		
Failure Data	1		
Fun.jmp	6		
JMP			
JMP モデルのあてはめ	98		
JMP 最小極値分布	88		
JMP 生存時間パラ	95		
JMP 接着剤	95		
JMP 等高線図	115		
JMP 微分関数	123		
JMP 偏微分	121		
JMP アレニウス変換温度	47		
JMP イベントプロット	29		
JMP サンプル	28		
JMP のサンプル	6		
JMP の寿命の二変量	55		
JMP の操作マニュアル	6		
JMP 寿命の一変量	19		
JMP 寿命の一変量	20		
JMP 寿命の二変量	41		
JMP 寿命の二変量	43		
JMP 寿命の二変量	48		
JMP 生存時間パラメトリック	11		
JMP 生存時間分析	9		
JMP 生存時間分析	22		
JMP 二変量の関係	5		
JMP 二変量の関係	8		
		JMP 二変量の関係	38
		L---S	
		Life Data	1
		LIFEREG プロシジャ	2
		likelihood 尤度	19
		Maximize()	124
		Meeker プロファイル尤度	109
		Meeker ら (1998)	105
		Mekker ら (1998)	36
		Minverse()	69
		Minverse() 関数	78
		Nelson 1982	29
		Nelson 原著者	5
		Newton-Raphson 法	76
		Newton-Raphson 法 手作業	79
		RELIABILITY プロシジャ	2
		SAS ユーザー総会	2
		SAS/QC	2
		SEV 最小極値分布	61
		Smallest Extreme Value Distribution	61
		Survival Time	1
		W	
		Wald	109
		Wald 法ワイブル	31
		Weibull	48
		Weibull α	14
		Weibull α 3 元のセル	96
		Weibull α 分散	103
		Weibull α, β 群ごとに異なる	40
		Weibull β	14
		Weibull β 分散	103
		Weibull β を共通	52
		Weibull β を共通	42
		WEIBULL()	16
		WEIBULL()	40
		α—	
		μ_{SEV}	61
		σ_{SEV}	61
		σ_{NOR}	86
		φ_{NOR}	86
		Φ_{NOR}	86
		σ_{SEV}	87
		φ_{SEV}	87
		Φ_{SEV}	87

あ			
アレニウス摂氏 パラメータ推定	58	カプラン・マイヤー曲線	22
アレニウス変換温度	10	カプラン・マイヤー法	1
アレニウス変換温度	46	カプラン・マイヤー法	22
アレニウス変換温度	56	幾何平均	31
アレニウス変換温度	94	逆行列 対角要素	83
位置	53	逆推定 分散	100
位置 別々	56	行列関数 Excel	79
位置 μ	20	極値パラメータ δ	14
位置パラメータ μ	14	極値パラメータ λ	14
因子グリッドの設定	115	近似計算 (Wald 法)	34
因子の設定	115	偶発故障型	70
打ち切り	7	繰り返し測定	93
打ち切り δ	24	グリッドテーブルの出力	115
打ち切りデータ	20	形状パラメータ m	14
打ち切りデータ	24	形状パラメータ β	14
打ち切りデータ	2	検定 包括モデルの	53
応答局面解析	99	検定 モデルの比較	53
大橋靖雄・浜田知久馬 (1995)	1	検定 尤度比	53
奥野監訳本	76	検定統計量 モデル	53
奥野監訳本	78	交互作用プロット	97
奥野監訳本	90	格子点 メッシュ状	113
奥野監訳本	104	合成分散 デルタ法	104
奥野監訳本(1988)	5	合成分散 一般式	100
奥野監訳本(Nelson)	63	合成分散の一般式	84
奥野忠一監訳 (1988)	35	誤差 ワイブル分布	55
温度 反応の関係	95	誤差 ワイブル分布	90
温度の回帰式	45	誤差 基準化	88
		故障確率の 95%信頼区間	70
か		故障データ	1
カイ 2 乗分布 下側確率	113	故障データ t	24
カイ 2 乗分布の下側確率	117	故障のメカニズム	70
回帰式 最尤法	87	故障までの時間	28
回帰直線 アレニウス摂氏温度	57	故障率 ノンパラメトリック	37
回帰直線 ワイブル分布を誤差	57	故障率 推定	20
回帰分析 最小極値分布を誤差	92	故障率 0.10	38
回帰分析 最尤法	54	故障率曲線	8
回帰分析の拡張	8	故障率の 95%信頼区間	72
回帰分析の拡張	85	今後の課題	99
各種の推定	59		
確認試験 寿命の	99	さ	
加速(過酷)寿命試験	1	最小 2 乗法	2
加速因子	11	最小 2 乗法	8
加速係数プロファイル	58	最小 2 乗法	100
加速試験 回帰分析	85	最小 2 乗法 回帰分析	85
加速試験 最小極値分布	87	最小極値パラメータの推定	34
加速試験 接着剤	93	最小極値分布	3
加速試験データ	7	最小極値分布	10
加速試験データの統計解析	35	最小極値分布	14
加速寿命試験	36	最小極値分布	55
加速条件	86	最小極値分布	61
活性化エネルギー	94	最小極値分布 基準化	86
稼働率	21	最小極値分布 ディーゼル発電機ファ	
		極値分布	64



















最小極値分布 加速試験	88	信頼性データ	2
最小極値分布 最尤法	64	信頼度	22
最小極値分布 対数尤度	122	鈴木ら (2014)	22
最小極値分布 必須の予備知識	90	正規確率プロット	4
最小極値分布 パーセント点	74	正規分布 ニュートン・ラフソン法	82
最小極値分布 確率密度関数 f	62	正規分布 最尤法	81
最小極値分布 累積分布関数 F	62	正規分布 最尤法	83
最小極値分布	31	正規分布 偏微分	81
最大化 対数尤度	17	正規方程式	85
最尤解	82	生存関数	65
最尤推定量	19	生存曲線	1
最尤推定量	25	生存時間データ	1
最尤法	2	生存率 推定	20
最尤法	24	生存率	21
最尤法 Newton-Raphson 法	79	積和	39
最尤法 パラメータ推定	66	摂氏温度	11
最尤法 ワイブル分布	17	摂氏温度での推定曲線	57
最尤法 加速試験データ	35	接着剤 寿命データの解析	94
最尤法 回帰分析	86	接着剤 新たな解析手順	95
最尤法 合成分散の一般式	102	セミパラメトリック	1
最尤法 正規分布	17	線形	48
最尤法 正規分布	83	全数打ち切り 10°C	45
最尤法 打ち切りデータ	25	相対湿度	93
最尤法による回帰分析	54	ソルバー	39
算術平均	15		
視覚的判断	26	た	
時間 t での故障率	72	対数正規	48
時間ごとの故障率	28	対数正規分布	3
試験研究段階	99	対数正規分布	31
指数分布	3	対数変換 2重	25
下側確率 0.632	13	対数変換 最小極値分布	61
尺度 σ	20	対数尤度 $\ln L$	16
尺度	53	対数尤度 差の2倍	53
尺度パラメータ η	14	対数尤度 最大化	76
尺度パラメータ σ	14	対数尤度 (-2)	20
尺度パラメータ α	14	対数尤度のピーク	117
寿命試験 接着剤	93	対数尤度の和	19
寿命データ	1	多因子 寿命試験	93
寿命データ	15	互いに独立	100
寿命分位点	60	多群のデータ	37
瞬間故障率	22	多重打ち切りデータ	37
瞬間故障率	37	逐次的 変化過程	83
初期値	16	逐次的に変化	16
信頼区間 カプラン・マイヤー	33	調整故障率	25
信頼区間 パーセント点	75	直線 フリーハンド	38
信頼区間 パラメータ	69	釣鐘型の形状	31
信頼区間 ワイブル確率プロット	73	ディーゼル発電機	4
信頼区間 故障確率	33	ディーゼル発電機のファンデータ	29
信頼区間 故障時間	59	デザイン行列 X	89
信頼区間 時間 t	32	デバイス A	7
信頼区間 時間 t	74	デバイス A	35
信頼区間 尤度	110	デルタ法	111
信頼性工学	22	デルタ法 合成分散	101

デルタ法 知る人ぞ知る	104	分散共分散の推定	55
統計ソフト JMP	1	分散 合成の一般式	74
統計ソフト SAS	2	分散の合成	61
等高線図 下側確率	119	分散分析 3元配置	93
同時尤度関数	24	平均 μ	18
		平均値 標準誤差	67
な		変化量 0	80
中村訳 (1968)	105	変化量 パラメータ	80
日本の信頼性	63	偏差平方	15
ニュートン・ラフソン法	3	偏微係数	79
		偏微分 2階	68
は		偏微分 2階	77
芳賀敏郎 (2004)	17	偏微分 正規分布	81
ハザード	22	偏微分係数行列 Z	77
パラメータ 線形	101	偏微分係数行列 Z	83
パラメータ 非線形	101	偏微分式 相互検証	123
パラメータ 微分	102	包括モデルの検定	56
パラメータ同時推定	18	ボトム点 下側確率	117
ピーク 対数尤度	117	ボルツマン定数	46
歪んだ等高線	120		
非線形回帰	121	ま	
非線形モデル	17	マトリクス展開 分散	105
一山型の形状	14	密度曲線の追加	48
微分 JMP	122	メディアン・ランク法	25
微分 数値計算	122	目の子	15
表記法の混在	63	目盛 ワイブル確率	30
標準正規分布	85	目盛 対数正規確率	31
標準偏差	15	目盛ノンパラメトリック	31
標準偏差 σ	18		
標準偏差の標準誤差	67	や	
比例ハザードモデル	1	尤度 likelihood	19
廣野 (2000)	93	尤度	15
品質管理データ	2	尤度曲線	20
ファンの故障時間	4	尤度等高線	112
不偏分散	15	尤度の積	19
ブラック・ボックス	79	尤度の最大化	76
プロファイル曲線	115	尤度比カイ 2乗	53
プロファイル曲線 対数尤度	20	尤度比カイ 2乗	56
プロファイル尤度	109	尤度比検定	94
プロファイル尤度法	33	尤度比検定	115
	33		
分位点の追加	48	ら	
分散 回帰の推定値	100	量産品 品質の担保	99
分散 故障確率	103	累積故障率	15
分散 合成分散の一般式	70	歴史的背景	35
分散 最小極値分布パーセント点	104	ログランク検定	1
分散 算術平均	102		
分散 平均値	100	わ	
分散共分散 Weibull パラメータ	76	ワイブル確率プロット 温度 x を連続	
分散共分散行列 パラメータ	66	尺度	54
分散共分散行列 Σ	77	ワイブルプロット	9
分散共分散行列パラメータ	32	ワイブル回帰	94
分散共分散行列を Σ	71	ワイブル回帰 セルごと	96

ワイブル回帰 伝統的な	95	ワイブル分布 回帰パラメータ	54
ワイブル確率プロット 群ごとに異なる	51	ワイブル分布 群ごとに異なる	38
ワイブル確率紙	1	ワイブル分布 群ごとに異なる	50
ワイブル確率紙	8	ワイブル分布 対数尤度	76
ワイブル確率紙	15	ワイブル分布のパーセント点	27
ワイブル確率紙	26	ワイブル分布の確率密度関数	2
ワイブル確率紙	41	ワイブル分布の確率密度関数	13
ワイブル型累積ハザード紙	22	ワイブル分布 分散共分散行列	78
ワイブル型累積ハザード紙	1	ワイブル分布の累積分布関数	2
ワイブル型累積ハザード紙	6	ワイブル分布の累積分布関数	13
ワイブル型累積ハザード紙	27	ワイブル分布を最小極値分布へ変換	61
ワイブル分布	1	ワイブル累積分布	30
ワイブル分布 下側確率	119		
ワイブル分布 回帰パラメータ	44		

Excel, JMP ファイル一覧

	13 KB	寿命_1_2章_ファン_デバイスA_データリスト
	8 KB	寿命_2_2章_ファン_故障_打切_2群比較
	7 KB	寿命_3_2章_デバイスA_多群比較
	6 KB	寿命_4_2章_デバイスA_回帰直線
	8 KB	寿命_5_2章_デバイスA_ワイブル_故障率
	11 KB	寿命_6_2章_デバイスA_正規分布最尤法
	11 KB	寿命_7_2章_デバイスA_アレニウス
	61 KB	寿命_8_3章_ワイブル_確率密度数_累積分布
	224 KB	寿命_9_3章_ファン_各種_分析
	9 KB	寿命_10_3章_ファン_故障データ_正規分布
	7 KB	寿命_11_3章_ファン_カプラン_マイヤー
	8 KB	寿命_12_3章_ファン_ワイブル_確率プロット
	28 KB	寿命_13_3章_サンプルライブラリ_Fan_解析事例
	131 KB	寿命_14_4章_デバイスA_各種解析
	13 KB	寿命_15_4章_デバイスA_各種解析
	15 KB	寿命_16_4章_デバイスA_各種解析 2
	8 KB	寿命_17_5章_デバイスA_サンプルデータ
	12 KB	寿命_18_5章_デバイスA_解析
	14 KB	寿命_19_5章_デバイスA_解析2
	61 KB	寿命_20_6章_最小極値_確率密度数_累積分布

	159 KB	寿命_21_6章_最小極値_各種解析
	9 KB	寿命_22_6章_ファン_最小極値_微分_表6_5検証
	25 KB	寿命_23_7章_ファン_ワイブル_NR法
	9 KB	寿命_24_7章_ファン_微分_ワイブル
	76 KB	寿命_25_7章_ファン_故障のみ_正規_NR法
	10 KB	寿命_26_7章_ファン_故障のみ_微分_正規(分散)
	10 KB	寿命_27_7章_ファン_故障のみ_微分_正規(SD)
	93 KB	寿命_28_8章_デバイスA_各種解析
	11 KB	寿命_29_8章_デバイスA_最小極値
	20 KB	寿命_30_8章_デバイスA_回帰_微分
	11 KB	寿命_31_9章_接着剤_各種解析
	4 KB	寿命_32_9章_接着剤_3元配置
	33 KB	寿命_33_9章_接着剤_各種解析
	13 KB	寿命_34_11章_ファン_プロフィール尤度
	802 KB	寿命_35_11章_最小極値_ワイブル_尤度等高線
	9 KB	寿命_36_11章_最小極値_ μ ピーク
	9 KB	寿命_37_11章_最小極値_ σ ピーク
	11 KB	寿命_38_12章_デバイスA_回帰_微分

非売品, 無断複製を禁ずる

第4回 続高橋セミナー 寿命試験データの統計解析

BioStat 研究所(株)

〒105-0014 東京都 港区 芝 1-12-3

2015年4月10日 高橋 行雄

takahashi.stat@nifty.com , FAX : 03-342-8035