5.1 Introduction

"In the previous chapters we have discussed models and analysis for relatively simple problems involving a single distribution. When, as this is the case, two or more data sets have to be compared, this is best sometimes best done by estimating survivor functions etc for each set of data separately and then making a qualitative comparison, either directly or via summary statistics.

In the present chapter, we review some of the many possible models that may be used to represent the effect on the failure time of explanatory variables. For this we suppose that for each individual there is defined a q*1 vector z of explanatory variables. The components of z may represent various features thought to affect failure time, such as:

(1) Treatments;
(2) Intrinsic properties of the individuals;
(3) Exogenous variables.

An example of explanatory variables:-

In a simple comparison of 2 treatments, eg of a 'new' treatment with a 'control', we consider a binary explanatory variable equal to 1 for individuals receiving the treatment, and equal to 0 for those receiving the control. If the treatment is specified by a dose or (stress) level the corresponding explanatory variable is dose or log dose." (CO, 1984)

5.2 Accelerated life model

5.2.1 Introduction

The accelerated life model may provide the answers to such questions like 'Explain why the normal age for a man to die is 65, whereas in the near future a man will die when he reaches the age of 130.' and 'How do I improve the design life of electronic components?'. To answer these questions, we have to know what the rate parameter is going to represent. For example, in the first question the rate parameter may represent the no. of deaths per year.

5.2.2 Simple form

Suppose there are 2 treatments represented by values 0 and 1 of the explanatory variable z. Let the survivor function at z=0 be $F_0(t)$, in the accelerated life model there is a constant w such that the survivor function at z=1, written variously $F_1(t)$ or $F(t,1)$, is given by

$F_1(t)=F_0(wt)$, (5.1)

so that $h_1(t)=wh_0(wt)$ and $f_1(t)=wf_0(wt)$ (5.2) (CO, 1984)

Example :-

For Weibull $F_0(t)=\exp(-(pt)^k)$, under normal conditions and $F_1(t)=\exp(-(pwt)^k)$. Also we have $h_0(t)=kp(pt)^{k-1}$, under normal conditions and $h_1(t)=wkp(pwt)^{k-1}=k(pw)^k(t)^{k-1}$. Having multiplied F_1 by h_1, $f_1(t)=wkp(pwt)^{k-1}\exp(-(pwt)^k)=k(pw)^k(t)^{k-1}\exp(-(pwt)^k)$. (5.3)

"More generally, put $w=w(z)$ in (5.1), (5.2) and (5.3).
If $F_0(.)$ refers to the standard conditions $z=0$, then $w(0)=1$.
A representation in terms of random variables is

$$T=T_0/w(z) \quad (5.4)$$

where T_0 has survivor function $F_0(.)$.
If $(mu)_0=E(\ln T_0)$, we can write this as:

$$\ln T=(mu)_0 - \ln w(z)+e \quad (5.5)$$

where e is a random variable of mean=0 and with a distribution not depending on z.
In problems with a limited no. of distinct values of z, it may not be necessary to specify $w(.)$ further. In other contexts, a parametric form for $w(.)$ may be needed, we then write $w(z;b)$.
Since $w(z;b)>=0$, $w(0;b)=1$, a natural candidate is

$$w(z;b)=\exp(b^T z) \quad (5.6)$$

where now the parameter vector b is q*1, (Where q=p+1)
Then (5.5) can be written as

$$\ln T=(mu)_0 - b^T z+e, \quad (5.7)$$

a linear regression model. Note that for the comparison of 2 groups, with a single binary explanatory variable, we get (5.1) with $w=\exp(b)$." (CO, 1984)

5.2.3 Some consequences useful for model checking

"The central property of the accelerated life model can be re-expressed in various ways that can be used as a basis for testing the adequacy of the model. Thus from (5.5) the distributions of ln T at various values of z differ only by translation. In particular var(ln T) is constant.

Alternatively in the 2-sample problem we can compare quantiles, we define $t_0^{(a)}$, $t_1^{(a)}$, for $0<a<1$, by

$$a=F_0(t_0^{(a)}), \quad t_0^{(a)}=F_0^{-1}(a)$$
$$a=F_1(t_1^{(a)}), \quad t_1^{(a)}=F_0^{-1}(a) \quad (5.8)$$

so that under (5.1) $t_1^{(a)}=t_0^{(a)}/w$
ie the so-called Q-Q plots (quantile-quantile plots) are straight lines through the origin." (CO, 1984).
Cox and Oakes (1984), uses the inversion method to obtain (5.8).
The GLIM macro RESPLOT produces two Q-Q plots (see APPENDIX 04) and I will be using this macro in most of my GLIM programs.
According to Cox and Oakes (1984), it is assumed that
$F_0(.)$ is strictly decreasing (ie the survivor function at its previous time is less than the survivor function at its future time, or the survivor function at its past time is less than the survivor function at its present time), so that the quantiles are uniquely defined.

5.2.4 Parametric version

The Weibull distribution has parameters (p_0w,k), where p_0 refers to standard conditions.
Survivor function is $\exp(-(p_0wt)^k)$
Probability density function is given by
$wkp_0(p_0wt)^{k-1}\exp(-(p_0wt)^k)=k(p_0w)^k(t)^{k-1}\exp(-(p_0wt)^k)$
Hazard function is given by $wkp_0(p_0wt)^{k-1}=k(p_0w)^k(t)^{k-1}$
These results agree with the simple form. Let's compare them with the results obtained from CHAPTER 2.

5.2.4.1 Comparisons

5.2.4.1.1 Introduction

I now find out what happens to the survivor functions and various quantities, when the treatments are subject to an acceleration factor of w.

Figures 5.1, 5.2, 5.3, 5.4 and 5.5 provide the graphs of survivor functions, hazard functions, pdfs, integrated hazards and product limit estimator against time. In each case, the index parameter is 3 and the rate parameter (under standard conditions) is 1.25. Some or all of the data from the data sets are taken from F77 program 21, except Figure 5.5, whose data sets are taken from F77 program 23.

5.2.4.1.2 Statistics

We can use equations (2.6) to (2.10). The Weibull distribution has survivor function $F(t)=\exp(-(p_o w t)^k)$

To find the median, put survivor function=0.5 and use (2.6)
$0.5=\exp(-(p_o w t)^k)$
$\ln(0.5)=-(p_o w t)^k$
$0.693=(p_o w t)^k$
$(0.693)^{1/k}=(p_o w t)$
$t_m=(0.693)^{1/k}/p_o w$ (5.9)

If w>1, the median of the Weibull distribution under the accelerated life model, is less than that obtained under standard conditions.

Similarly, lower quartile : $t_{lq}=(0.288)^{1/k}/p_o w$ (5.10)
upper quartile : $t_{uq}=(1.386)^{1/k}/p_o w$ (5.11)

To find the mid-quartile range, take the average of (5.10) and (5.11).
Mid-quartile range : $t_{midqr}=(0.837)^{1/k}/p_o w$ (5.12)
To find the inter-quartile range, subract (5.10) from (5.11).
Inter-quartile range : $t_{iqr}=(1.099)^{1/k}/p_o w$ (5.13)

Using the results in 3.5.1 and 5.2.4, we find that the mean and variance of the Weibull distribution are k/pw and $k/(pw)^2$ respectively.
The coefficients of variation and skewness are CONSTANT.

5.2.5 Uses of the accelerated life model

In Figure 5.1, the graphs can be used in modelling the no. of insects killed. The survivor function is proportional to the percentage of insects not killed, the rate parameter is proportional to the concentration of insect repellant and the index parameter is the order of reaction.
If the distribution is exponential, we can model radioactive decay. The survivor function is proportional to the no. of radioactive nuclei and the rate parameter is the decay constant.

In Figure 5.2, the graphs have many uses. Here are some examples:-
Modelling the RPM of a car engine. The PDF is proportional to the angular velocity or RPM and the rate parameter is proportional to the speed of the car. The radius may not be constant.

Study of heart rates. In accordance with Figure 5.2, a man does some exercises until the PDF is a maximum and then takes a break. The PDF is proportional to the no. of heart beats per minute and the rate parameter is proportional to the speed of the exercises.

Vertical motion under gravity. A ball falls from the top of a building to the ground and then bounces upwards. The PDF is proportional to the velocity of ball and the rate parameter is proportional to the angle of the ball.

Modelling temperature rises when running a race. In accordance with Figure 5.2, a typical runner has no temperature rise at the start of a race and has his own maximum temperature rise when he finishes his race. He then has a drink and/or sponge after his race to cool down. The PDF is proportional to the temperature rise of a body and the rate parameter is proportional to the velocity of the runner.

In Figure 5.3, the graphs have many uses. Here are some examples:-

Modelling the velocity-time graph of objects. The hazard function and rate parameter are proportional to the velocity of object. In the exponential distribution, the acceleration (gradient of velocity-time graph) is zero. In the Weibull distribution, the acceleration is non-zero. If k is at least 2, the acceleration is positive (A special case when k=2 is constant acceleration). If k is less than 1 the acceleration is negative, because the velocity is DECREASING as time increases.

Study of heart rates during continuous exercise. The PDF is proportional to the no. of heart beats per minute and the rate parameter is proportional to the speed of the exercises.

In Figure 5.4, the graphs have many uses. Here are some examples:-

Modelling the temperature rise of a body, when running a race. The integrated hazard is proportional to the temperature rise and the rate parameter is proportional to the speed of the runner. Assume that the runner does not have a drink and the temperature rise is proportional to the kinetic energy of a body.

Study of the reaction kinetics in chemistry. pt is proportional to the concentration of a substance, assuming that the other substances are not allowed to have varying concentrations. The rate parameter is proportional to the rate constant and the index parameter represents the order of reaction. The gradient of the graph in Figure 5.4 gives the rate of reaction.

In Figure 5.5, the graphs have many uses. Here are some examples:-

Modelling radioactive decay using dice or random numbers. The product limit estimator is proportional to the no. of objects left. The rate parameter
represents the decay constant. For the dice experiment:-

(1) Roll a certain no. of dice.
(2) Discard the dice that show '6', say.
(3) GOTO (1), until there are no dice left.

For another group, we have to modify (2) to increase or reduce the probability of decay. If the product limit estimator is constant for any time interval, then none of the objects have decayed, because none of them have obeyed the rules for decaying to occur.

The random numbers experiment is similar to the dice experiment, but we have to modify (2), eg Discard the random numbers less than 0.400. This experiment is useful for modelling the students taking and/or resitting exams. If the random numbers are 0.400 or greater, say, they have passed and must be discarded from resitting exams.

5.3 Proportional hazards model

5.3.1 Simple form

"A second broad family of models that has been widely used in the analysis of survival data is best specified via the hazard function. Suppose that the hazard is $h(t;z)=w(z)h_o(t)$ (5.14)
Survivor function and density is thus:

$[F_o(t)]^{w(z)}$,

$w(z)[F_o(t)]^{w(z)-1}f_o(t)$" (CO, 1984)

EG for Weibull:

$h(t;1)=kw(z)(pt)^{k-1}$,

$F(t;1)=(\exp[-(pw(z)t)^{k}])$,

$f(t;1)=w(z)[\exp[(-pw(z)t)^{k}]kp(pt)^{k-1}[\exp((-pt)^{k}\ )]$

"Reasons for considering this model are that:

(a) There is a simply understood interpretation to the idea that the effect of, say, a treatment is to multiply the hazard by a constant factor;

(b) There is in some fields empirical evidence to support the assumption of proportionality of hazards in distinct treatment groups;

(c) Censoring and the occurence of several types of failure are relatively easily accommodated within this formulation and in particular the technical problems of statistical inference when $h_o(t)$ is arbitary have a simple solution." (CO, 1984)

5.3.2 Relation with accelerated life model

"For constant explanatory variables the question naturally arises as to when the proportional hazards model is also an accelerated life model. For this we need a function chi(z) such that

$[F_o(t)]^{w(z)}=F_o\ [t*chi(z)]$

Write $G_o(tau)=\ln[-\ln\ F_o\ (\exp(tau))]$

Then $\ln\ w(z)\ +\ F_o(tau)=G_o((tau)+(lambda(z)))$,

where $lambda(z)=k^{-1}\ \ln(chi(z))$

For this to hold for all t and for some non-zero lambda(z), ie non-unit chi(z), we need

$G_o(tau)=(k*tau)+(a)$, $lambda(z)=\ln\ w(z)$, where a and k are constants.

Thus on writing $p=\exp(a/k)$ we have that

$F_o(t)=\exp[-(pt)^{k}]$

That is, Weibull distribution is the only initial distribution for which, with constant explanatory variables, the accelerated life and proportional hazards models coincide.

It follows directly from the definition of the Weibull distribution that the accelerated life model with 'scale' parameters has survivor function and hazard

$\exp[-(pw_{a1}(z)t)^{k}]$, $k[pw_{a1}(z)]^{k}t^{k-1}$

That is, a proportional hazards model defined by:

$w_{ph}\ =\ [w_{a1}]^{k}$" (CO, 1984)

We can compare the proportional hazards model with the accelerated life model, by means of graphs.
Figures 5.1 to 5.5 provide the graphs under standard conditions, accelerated life and proportional hazard models.
In each case, k=3 and p=1.25, under standard conditions, the hazard is doubled and the rate parameter is accelerated by a factor of 2 (equivalent to the hazard being multiplied by 8). The comparisons and uses for the accelerated life model also apply to the proportional hazards model, but in Equations (5.9) to (5.13), substitute $w_{ph} = [w_{a1}]^k$.

Index	Rate	Median	Lower-Q	Upper-Q	IQ Range	Mid-Q
3.00000	1.25000	0.70800	0.52811	0.89202	0.36391	0.71007
3.00000	1.57490	0.56194	0.41916	0.70800	0.28883	0.56358
3.00000	2.50000	0.35400	0.26406	0.44601	0.18195	0.35503

Table 5.1 - Comparisons of the statistics for standard, proportional hazards and accelerated life models

---

From Table 5.1, we find that Equation 5.4 applies for the median, etc. The area of the curve in Figure 5.1 is proportional to the spread.

k	p	Mean	Var.	CV	Skew
3.00000	1.25000	2.40000	1.92000	0.57735	3.66329
3.00000	1.57490	1.90488	1.20952	0.57735	3.66329
3.00000	2.50000	1.20000	0.48000	0.57735	3.66329

Table 5.2 - Comparisons of the mean, variance, CV and skewness for standard, proportional hazards and accelerated life models

---

From Table 5.2, If we accelerate by a factor of 2 the mean is doubled and the variance is quadrupled. Equation 5.4 applies for the mean and NOT the variance. The coefficients of variation and skewness are constant. Although the coefficient of variation measures the relative spread, it is useless at comparing standard, proportional hazards and accelerated life models.

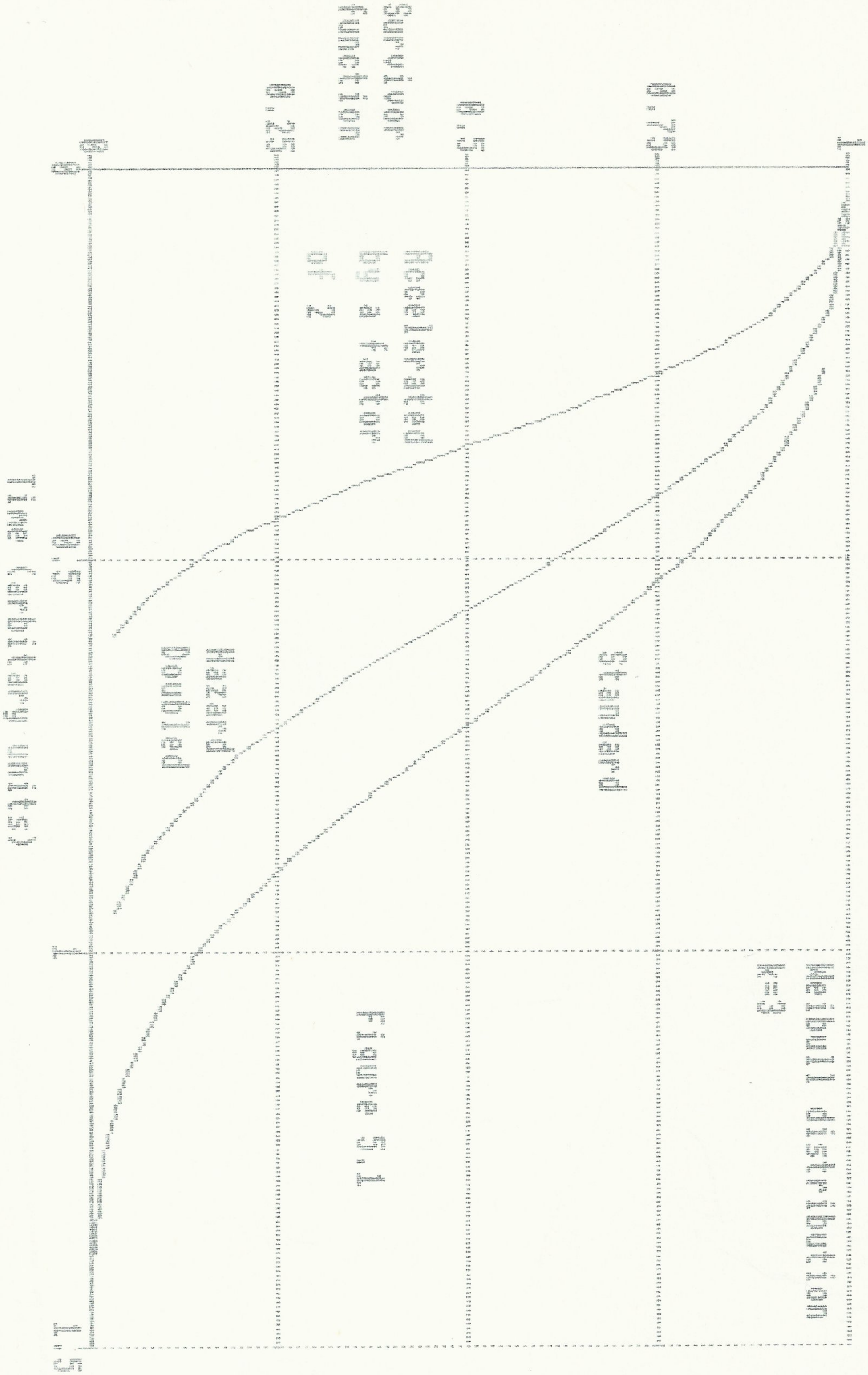

Survivor function
0
0.25
0.5
0.75
1
Time (arbitrary units)
0
0.5
1
1.5
Accelerated by a factor of 2
Hazard doubled
Standard
Weibull distribution
k=3

Figure 5.1

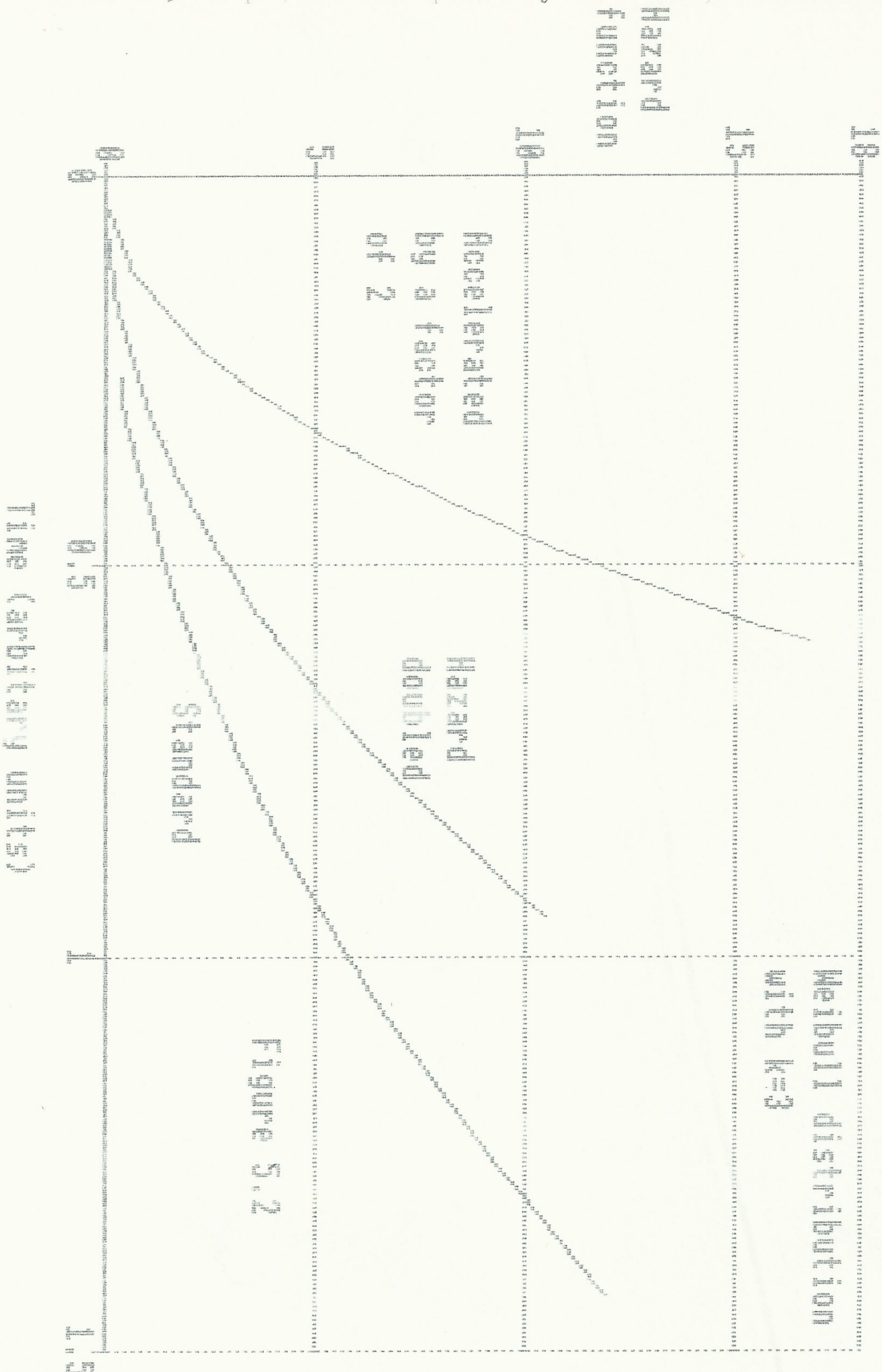
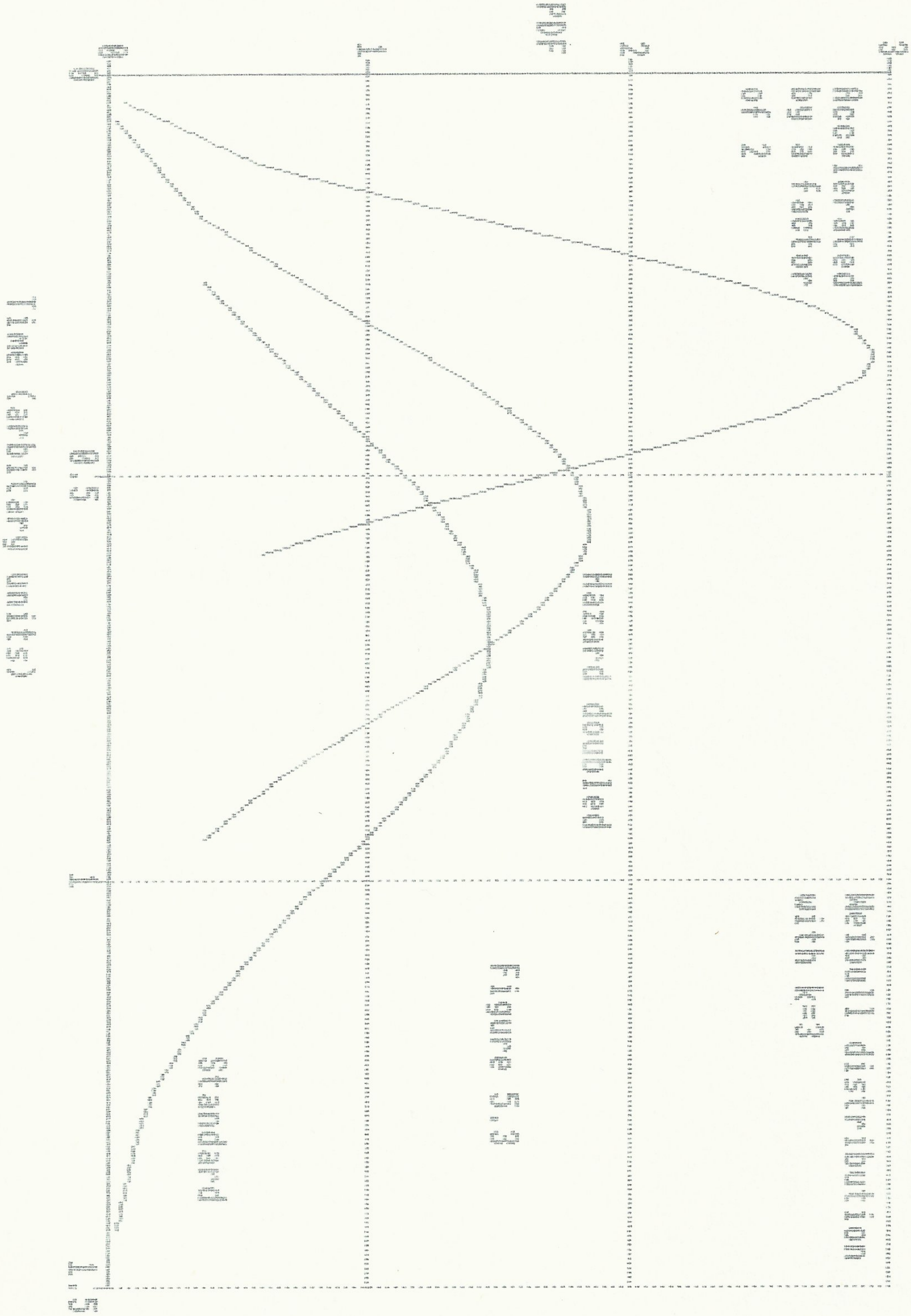
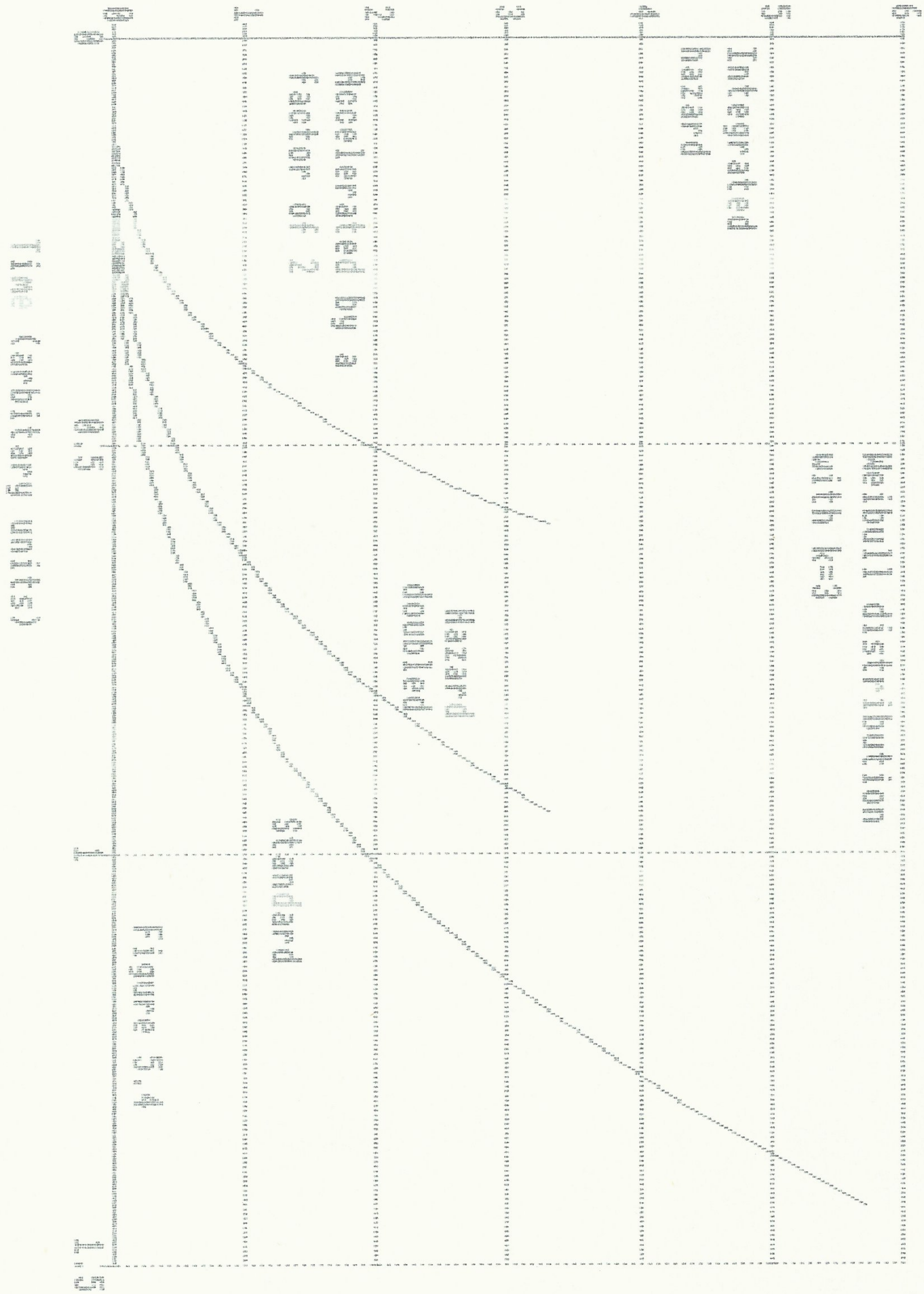
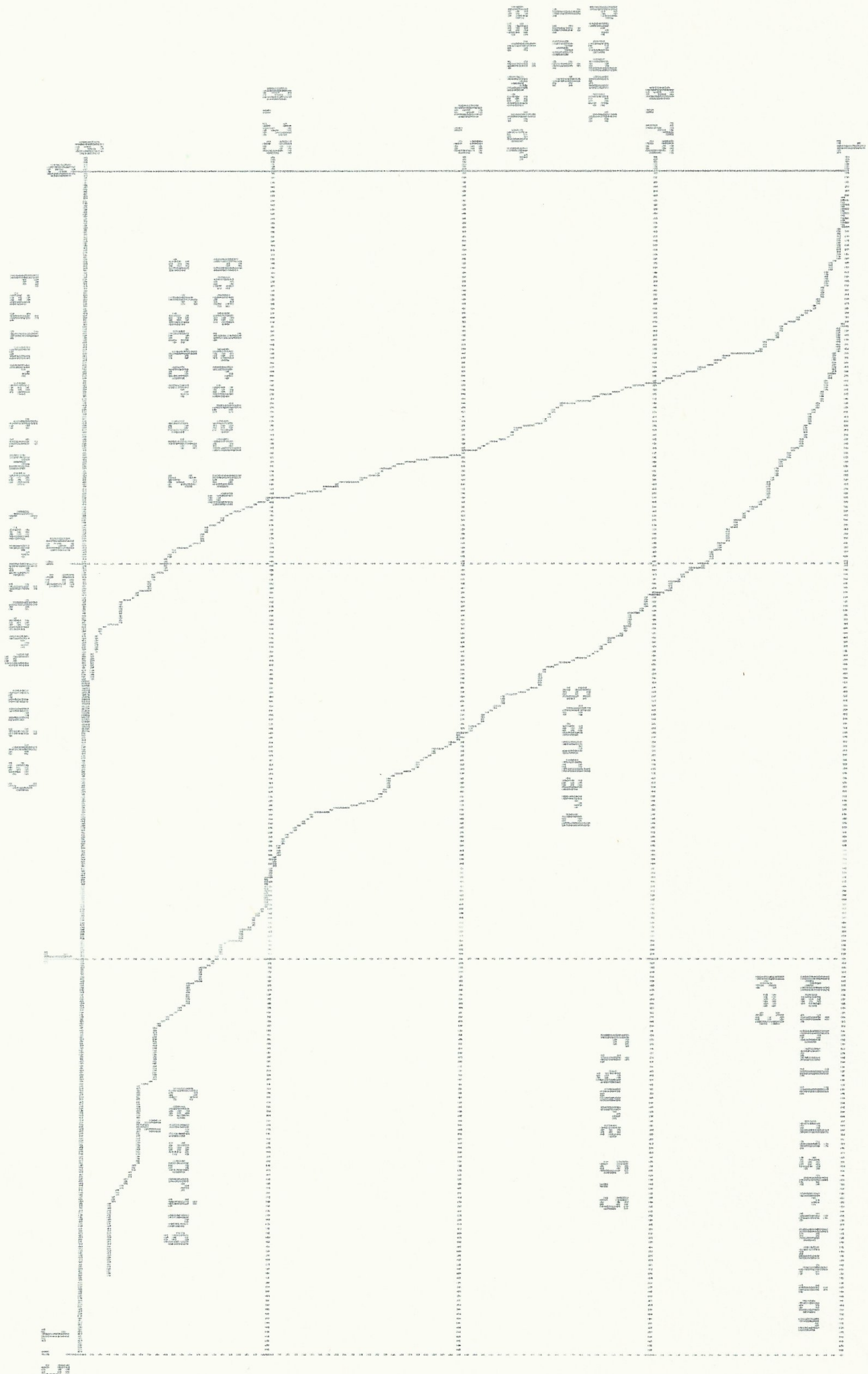

Weibull distribution
with k=3
Hazard
function
Accelerated
by a factor
of 2
Hazard
doubled
Standard
0
5
10
15
18
0
0.5
1
1.5
Time (arbitary units)

Figure 5.2

Dependence on explanatory variables ; Model formulation







5.4 Computing

Assume that the data (no. of individuals per treatment, index and rate parameters, seeds, conditions for censoring (if applicable) and acceleration factor) have been entered by the user.

5.4.1 Survivor functions, etc (With subroutines)

(01) Simulate the Weibull distribution, with some rate parameter. In addition, where appropriate, if the individual is censored in accordance with some censoring rule, the no. of failures is decremented by 1.
(02) Obtain the survivor functions, etc
(03) If more treatments, GOTO (01)

5.4.2 Product limit estimator (With subroutines)

(01) Simulate the Weibull distribution, with some rate parameter. In addition, if the individual is censored in accordance with some censoring rule, the no. of failures is decremented by 1.
(02) Sort the x values in ascending order.
(03) Obtain the product limit estimator.
(04) If more treatments, GOTO (01)

5.5 Summary

This chapter relates to most of my earlier work.
In the Weibull distribution, the parametric version of the accelerated life model is equivalent to the simple form.
In the Weibull distribution, the accelerated life and proportional hazards models coincide. I can use the same programs for the accelerated life model in order to use the proportional hazards model. (Cox and Oakes, 1984)
Accelerated life and proportional hazards models are used in many medical and scientific applications. EG Modelling temperature rises of bodies.

I have to modify my existing programs and have at least 2 treatments. I may include subroutines or each explanatory variable has its own indicator variable.

5.6 Later Work

In the next chapter, we will be studying the fully parametric analysis of dependence. It can be shown that $p_0=\exp(b_0)$.