CHAPTER 04 - Product limit estimator

4.1 Assumptions

I make the following assumptions in calculating product limit estimators, in order to make the model simple:

The events are independent.
No multiple events occur. That is, individuals are not allowed to have the same atom a_j and P(2 or more events at the same time)=0.
The no. of failures and the no. of trials in view at each atom a_j are DISCRETE random variables.
The no. of failures out of the no. of trials in view at each atom a_j are also independent, so that I can use the Binomial Model with logit link as a basis for logistic regression. For this purpose in GLIM, I define the no. of failures as a dependent variable, the no. of trials as a vector (For Binomial errors, GLIM needs a vector) and the failure time as an explanatory variable.

4.2 Definition

The product limit estimator (From CO, 1984) is given by :-

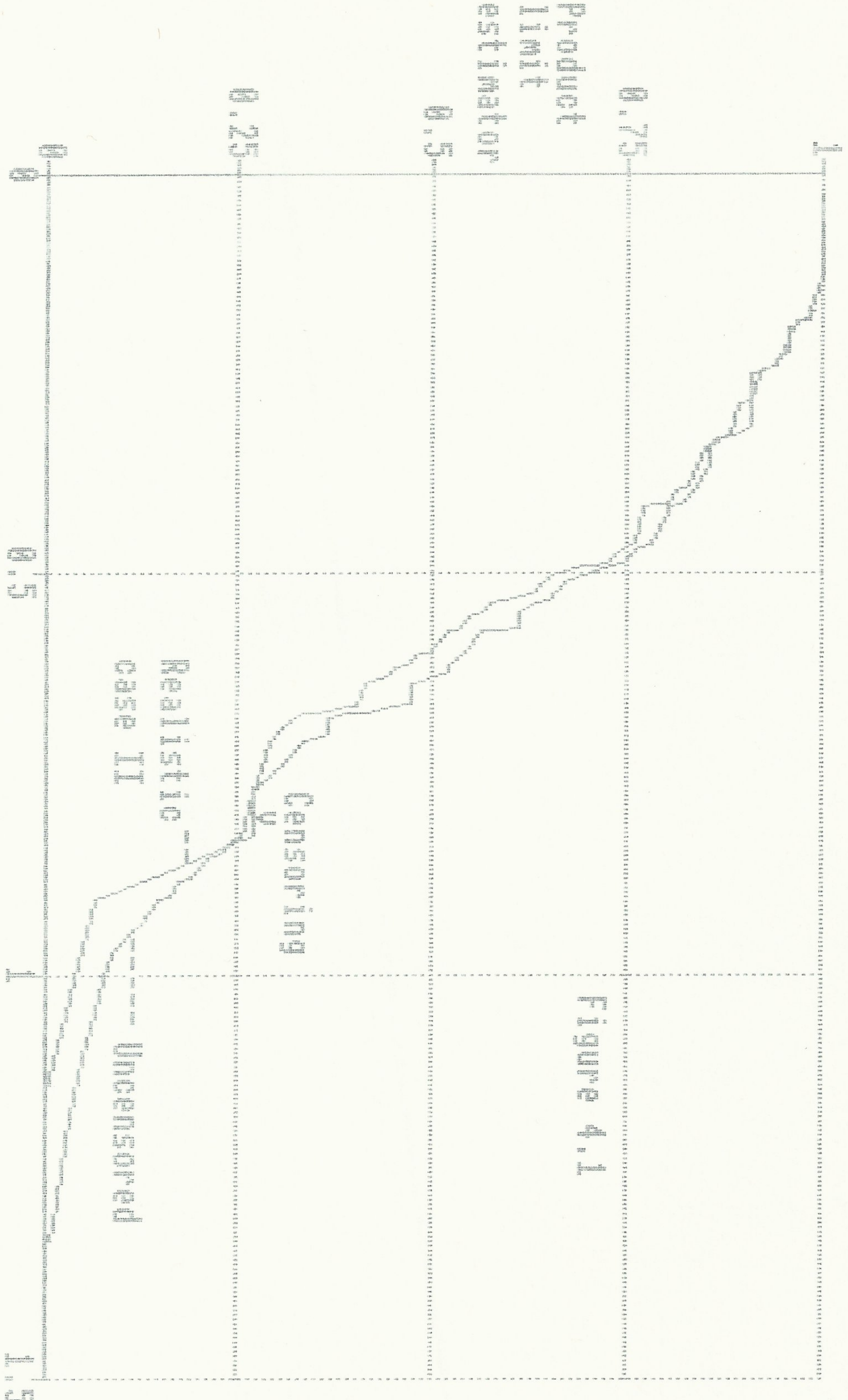
Product limit estimator = Product of $[(1-(d_i/r_i)]$
where d_i is the no. of failures at atom a_j (I have assumed it to be equal to 1 in my computer programs, if uncensored)
r_i is the no. of trials in view at atom a_j.

EG: Using the data sets and the models in Chapter 1, I obtained the following product limit estimators from my programs (The product limit estimator is for the longest failure time that is not censored) :-

Uncensored	0.0000
Constant censoring	0.0267
Realistic case I	0.0000
Realistic case II	0.0000

It is possible to plot a graph of product limit estimator against time (include any censoring). I used my data sets and calculated their product limit estimators for the uncensored and realistic case II models in Chapter 1 to obtain these graphs in Figure 4.1. I can distinguish between uncensored and censored models, using these graphs. If the product limit estimator is strictly decreasing as time increases, we have the uncensored model. If the product limit estimator is decreasing as time increases, we have the censored model.

In the next chapter, I will experiment how the product limit estimator behaves with time, when the second treatment group is accelerated by some factor. I can still distinguish between uncensored and censored models, as mentioned above.



4.3 Tests

Product limit estimator gives some idea of the survivor function. I will use the K-S test (described below) to examine whether the product limit estimator is reasonably consistent with the survivor function.

4.3.1 Description of the K-S test

For each uncensored individual, I take the absolute difference between the product limit estimator and the survivor function. Then I calculate the absolute maximum difference and compare it with the critical value of D in Table 16 of Statistical Tables (Murdoch and Barnes, 1986). If the calculated maximum difference is less than the tabulated value of D, then the product limit estimator is consistent with the survivor function.

EG: Using the data sets and models in Chapter 1 and the estimated parameters of the Weibull distribution in Chapter 3, I obtained the following results from my programs (The product limit estimator and survivor function is for the longest failure time that is not censored) :-

	Uncensored	Constant censoring	RC I	RC II
Product limit estimator	0.0000	0.0267	0.0000	0.0267
Survivor function	0.0062	0.0219	0.0054	0.0223
Calculated maximum diff.	0.0896	0.0904	0.0960	0.0783
D value at 5% level	0.1570	0.1592	0.1727	0.2178

I found that for each model, the product limit estimator was reasonably consistent with the survivor function.

4.4 Uses of product limit estimators

Product limit estimators are not only useful for medical work applications, but also for radioactive decay models in the form of a game. The algorithm for such models could be:-

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

In accelerated life and proportional hazards models (Chapter 5), we may modify step (2) for different probabilities of decay. If censoring is present, as in Figure 4.1, none of the remaining objects have decayed, because they have not obeyed the rules for decaying.

4.5 Computing

For the product limit estimator only, I used the following algorithm:-

(1) Set the product limit estimator to 1.
(2) Simulate the Weibull distribution with some starting values of the parameters. In the censoring models, take the failure times as the minimum of the censor times and the times simulated from the Weibull distribution.
(3) Sort the failure times in ascending order.
(4) Calculate the product limit estimator.

For the goodness of fit test, I used the following algorithm:-

(1) Follow the procedure in estimation, described in Chapter 3.
(2) Sort the failure times in ascending order.
(3) Calculate the product limit estimator and survivor function. Take the absolute differences between these values. Hence, find the maximum absolute difference.
(4) Find the critical value of D, corresponding to the no. of failures. Use linear interpolation, if necessary.
(5) Compare the calculated maximum difference with the critical value of D. If the calculated maximum difference is less than this critical value, the product limit estimator is consistent with the survivor function.

4.6 Summary

The product limit estimator gives some idea of the survivor function.
To distinguish between uncensored and censored models, look at the graphs as shown in Figure 4.1. If the product limit estimator is strictly decreasing as time increases, we have the uncensored model. If the product limit estimator is decreasing as time increases, we have the censored model.

4.7 Later Work

In accelerated life models (Chapter 5), I will experiment how the product limit estimator behaves with time, when the second treatment group is accelerated by some factor.
I will be using Binomial models with logit link as a basis for logistic regression.