

ERIVIEW 2000 - A TOOL FOR THE ANALYSIS OF FIELD STATISTICS

Patric Oscarsson and Örjan Hallberg

Department of Hardware Performance, Ericsson Telecom AB,
S - 126 25 Stockholm, Sweden

ABSTRACT

This paper presents a new parametric optimisation method (EriView 2000) used for the analysis of field failure data. This method requires only a minimum of raw data. The aim for this analysis method is to find a statistical cumulative distribution function that can describe the studied product's failure rate. The theory behind this method is described in detail. As it is implemented as an application i Microsoft Excel it is very user-friendly. Examples are given on the analysis of car accidents in Sweden and on the field reliability performance of printed circuit boards within the telecommunication market.

KEYWORDS

Reliability, quality, failure intensity, distribution function, parameteric optimisation, statistical analysis

INTRODUCTION

Today the technical development within most areas goes very fast which means that a company has to be very flexible. A demand for survival in a the long run is to continuously invest in research & development and to constantly work in an improvement process.

An improvement process depends on facts, data and information. This may be customer notices and complaints, measurements of significant parameters and knowledge about environmental conditions for the product.

The total area from which information on product performance information can be retrieved can be divided in to two subareas; internal and external. See Figure 1 below.

Reprinted from Proc. of ESREL 1997, Soares, Authors: Oscarsson, Hallberg, EriView2000, a tool for the analysis of field statistics. pp 1083-1090, Lisbon, 1997 ISBN 0-08-042835-5 with permission from Elsevier

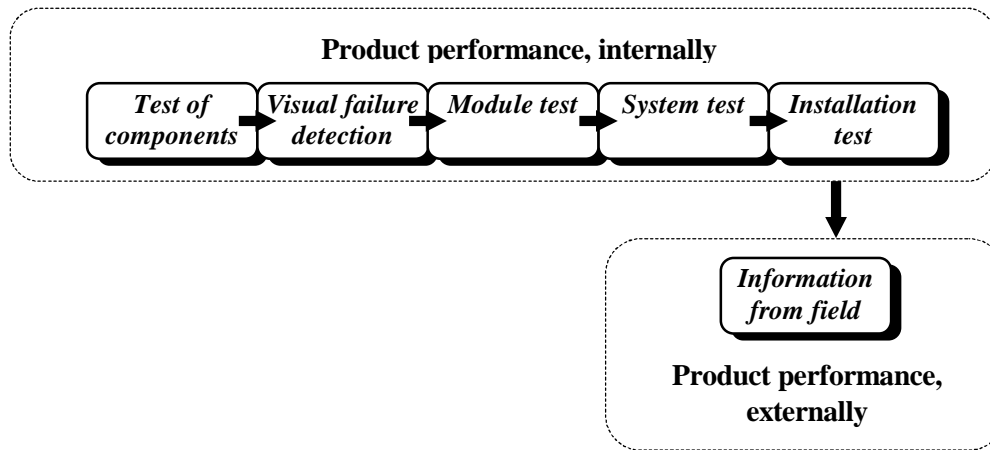


Figure 1: Examples of areas from where it is possible to retrieve product performance information

It is relatively easy to collect and to analyse quality information from the different internal tests. But, the environment in internal tests only simulates the real environment. Thus it is realised that it is very important to collect data and analyse how the product performs in its real environment in service.

There are different methods available to perform these kinds of quality analysis. Roughly the methods can be divided in to non-parametric methods and parametric optimisation methods.

THEORY

In this sector a brief description of the non-parametric analysis method is given and then is described the parametric analysis method (*EriView 2000*) developed at Ericsson Telecom AB, Sweden.

The non-parameteric method

This method requires a lot of detailed information about each failed product. The exact number of products put in to service at different calendar time points must be known. The lifelength of each failed product has to be calculated. This means that for every single product, it is necessary to be able to trace its location in service and its start time.

All population information is then compiled so that products put in to service at different calendar time points will have the same starting time. A similar compilation is then made for the failure information. It is then possible to plot the number of products in service and the number of failures vs use time. The measured failure rate can then be calculated and plotted¹ vs field use time. Frequently used is also a simpler variant of the non-parametric method, namely to calculate the average failure intensity by dividing the total number of faults by the total number of component-hours.

The parametric optimisation method

The central point in this method is to record the total number of products that are brought in to service each time period (e.g. month or quarter). It is also necessary to record the total number of returned products from field service by each time period. So, the required input for the parametric optimisation method will be:

- ? Information about the number of products put into service each time period
- ? Information about the number of products failed in service each time period.

The exact information about *when* and *where* a failed product was put in to service is not necessary for this method. Table 1 below gives an example of the minimum information needed for this parametric method.

TABLE 1
A FICTIVE EXAMPLE OF REQUIRED INPUT

Time (Year & Month)	No. of products put in to service	Number of failed products
90Q1	1390	1
90Q2	829	43
90Q3	8702	174
90Q4	8966	232
91Q1	15392	234
91Q2	18733	647
91Q3	22816	481
91Q4	24333	341
:	:	:
94Q3	2109	249
94Q4	1119	198
95Q1	693	206
95Q2	2143	230
95Q3	2277	195
95Q4	2147	155
96Q1	1040	64

From table 1 the following conclusions can be drawn:

- ? The faulty product detected 90Q1 can only belong to those 1390 products which were installed 90Q1
- ? The 43 faulty products detected 90Q2 belong to those 1390 products which were installed 90Q1 and/or to those 829 products which were installed 90Q2
- ? Failures detected in 90Q3 belong to products which were installed during 90Q1-90Q3 and so on...

The aim for this analysis method is to find a statistical cumulative distribution function for the studied product that gives the best fit to the measured data such as in table 1.

The central point is:

By using a statistical cumulative distribution function, $F(t)$, it is possible to, out of a population in service, calculate the estimated number of failed products after a certain time (t) in service.

Cumulative *Distribution Function* will from here be denoted as *c.d.f.*

It is necessary to distinguish between service time and calendar time because products are put in to service at different calendar times and the $F(t)$ intends service time.

Notations

The following notations are used:

T = calendar time

T_i = calendar time, period i

t = service time

N_i = No. of products put in to service in period T_i

u_i = No. of detected failures in period T_i

U_n = accumulated No. of detected failures during T_1, T_2, \dots, T_n

$$\text{so: } U_n = \sum_{i=1}^n u_i \quad (1)$$

Where:

The length of 1 period of calendar time = $t_i - t_{i-1}$

A generally statement will be:

$N_i F(t)$ = Estimated No. of failed products, of those that were put in to service in period T_i , before service time t .

The following notation:

$F_i = F(t_i) - F(t_{i-1})$ - the probability that a product should fail in the interval (t_{i-1}, t_i)

results in:

$N_i F_j$ = Estimated No. of failed products, of those that were put in to service in period T_i , in the service time interval (t_{j-1}, t_j) . Note that during the first service interval, t_1 , the effective number of parts put in to use is only $N_i/2$ as an average. This must be considered in real computation but is not shown in Table 2 below.

From the notations above the following can be concluded:

For N_1 : T_1 corresponds to $t_1 - t_0$; T_2 to $t_2 - t_1$ and so on.

For N_2 : T_2 corresponds to $t_1 - t_0$; T_3 to $t_2 - t_1$ and so on.

•
•

For N_i T_i corresponds to $t_1 - t_0$; T_{i+1} to $t_2 - t_1$ and so on.

The following table (table 2) can then be used to calculate the total No. of estimated failures per time period:

TABLE 2
THE TABLE DESCRIBES A THEORETICAL NOTIFICATION OF
TOTAL NO. OF ESTIMATED FAILURES

		T								
T	N	T₁	T₂	T₃	T₄	T₅	T₆	T₇	.	.
T₁	N₁	N₁F₁	N₁F₂	N₁F₃	N₁F₄	N₁F₅	N₁F₆	N₁F₇	.	.
T₂	N₂		N₂F₁	N₂F₂	N₂F₃	N₂F₄	N₂F₅	N₂F₆	.	.
T₃	N₃			N₃F₁	N₃F₂	N₃F₃	N₃F₄	N₃F₅	.	.
T₄	N₄				N₄F₁	N₄F₂	N₄F₃	N₄F₄	.	.
T₅	N₅					N₅F₁	N₅F₂	N₅F₃	.	.
T₆	N₆						N₆F₁	N₆F₂	.	.
T₇	N₇							N₇F₁	.	.
.	.								.	.
.	.									.

Out of table 2 it can be read that the estimated No. of failures in period T_1 is $N_1 F_1$. In period T_2 it is $(N_1 F_2 + N_2 F_1)$ and so on.

Therefore the following notation is stated:

b_i = Estimated No. of failures in period T_i

Then, generally:

$$b_i = \sum_{j=1,2,3,\dots}^n N_j F_k \quad (2)$$

Then the following equation:

$$B_n = \sum_{i=1}^n b_i \quad (3)$$

where:

B_n = accumulated No. of estimated failures during the periods T_1, T_2, \dots, T_n

We have now notations for both detected, U_n , and estimated, B_n , No. of failures after n periods of calendar time. Obvious, the aim is to *minimise* the difference between U_n and B_n . For that purpose the notation SS_n (Sum of Squares) is stated:

$$SS_n = \sum_{i=1}^n (U_i - B_i)^2 \quad (4)$$

where:

SS_n = (The sum of squares of the differences between the periodical values of the accumulated No. of failures during n time periods)

Note 1:

It is also possible to use periodic values, i.e. u_i and b_i , instead of accumulated values U_i and B_i

Then the notation LSS_n (Least Sum of Square) is stated where:

$LSS_n = \min(SS_n)$ for all possible values/combinations of the c.d.f.'s parameters.

That means that the aim of this parametric optimisation is to find the c.d.f., and its combination of the parameters, which gives the smallest deviation between detected and calculated/estimated No. of failures.

By this it is realized that when $\sum_{i=1}^n (U_i - B_i)^2 \approx 0$ the reliability for the product can be exactly described (at least within the time interval (T_1, T_n)) by the current c.d.f..

IMPLEMENTATION OF THE THEORY

The theory described for the parametric optimisation method is implemented as a Microsoft Excel application by Visual Basic programming. This application is named *EriView 2000*.

EriView 2000 is using Excel's built-in optimisation tool, the Solver, and by this tool it is possible to rather quickly achieve the parameter combination that gives the best fit to the measured raw data.

EXAMPLES

In this section examples are given to describe the parametric optimisation method and the application called *EriView 2000*.

Example 1; Field performance of a Printed Circuit Board

The inputs for this example is taken from table 1. Figure 2 below shows the result of a parametric optimisation analysis with *EriView 2000* on the data given in table 1. The calculated/estimated No. of returned products is compared with the measured No. of returns.

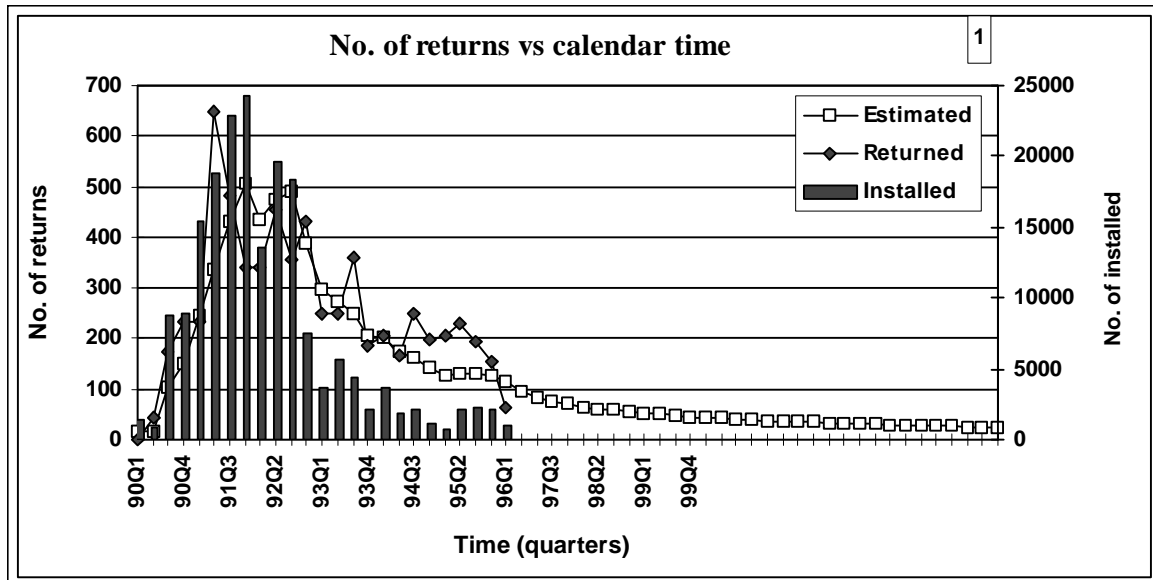


Figure 2: The results of a parametric optimisation, performed by *EriView 2000*, of the data given in table 1.

The optimal c.d.f. is then used to plot the calculated failure intensity and failure percentage vs service time. See figure 3 below.

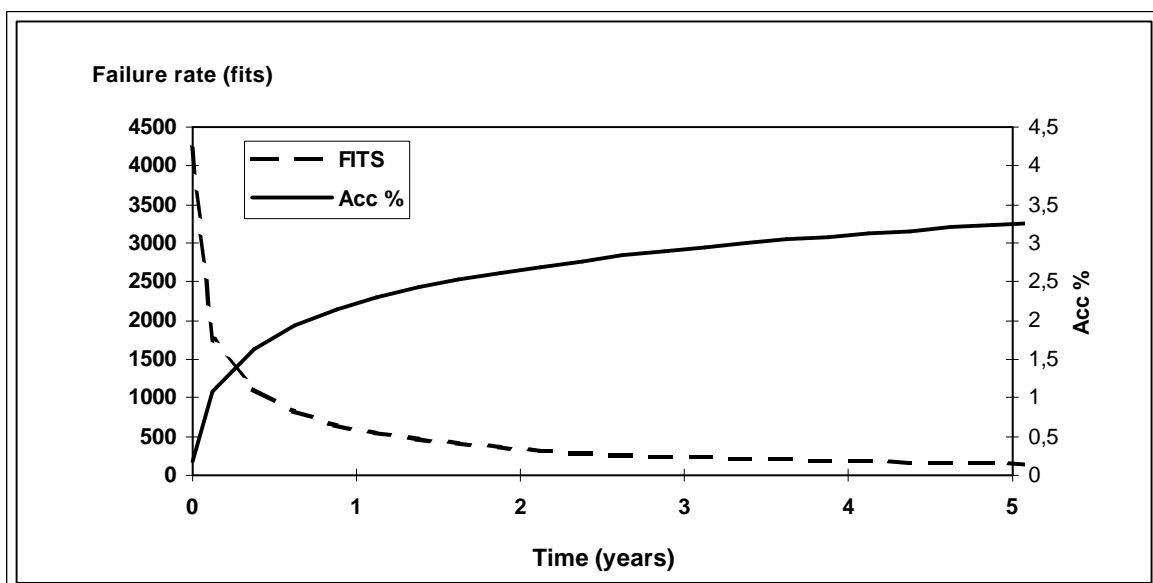


Figure 3: The failure intensity and the failure percentage vs service time as calculated by the optimal c.d.f.

Example 2; An analysis of mortal traffic accidents in Sweden

In this example we are using available statistics on the number of registered cars and the number of people killed by traffic accidents in Sweden. The input to the system being analysed is the annual *increase* of the total population of registered cars that are used in the traffic. The total transport system is assumed to be responding to an increase of the number of cars put into use as there inevitable will have to be about the same amount of new, inexperienced drivers that are exposed to the system. For this analysis it has been assumed that the average driver is active for 25 years.

The result of this 'disturbance' is then measured as the number of killed people per year. An analysis based on this raw material will give a good base for predictions of future rates of mortal accidents taking projected car sales into account.

Figure 4 gives the result after an optimisation for the time period 1976 to 1995. That means that the computer finds a c.d.f. that gives an optimum fit to the data in this time span. Obviously earlier mortal rates have been worse.

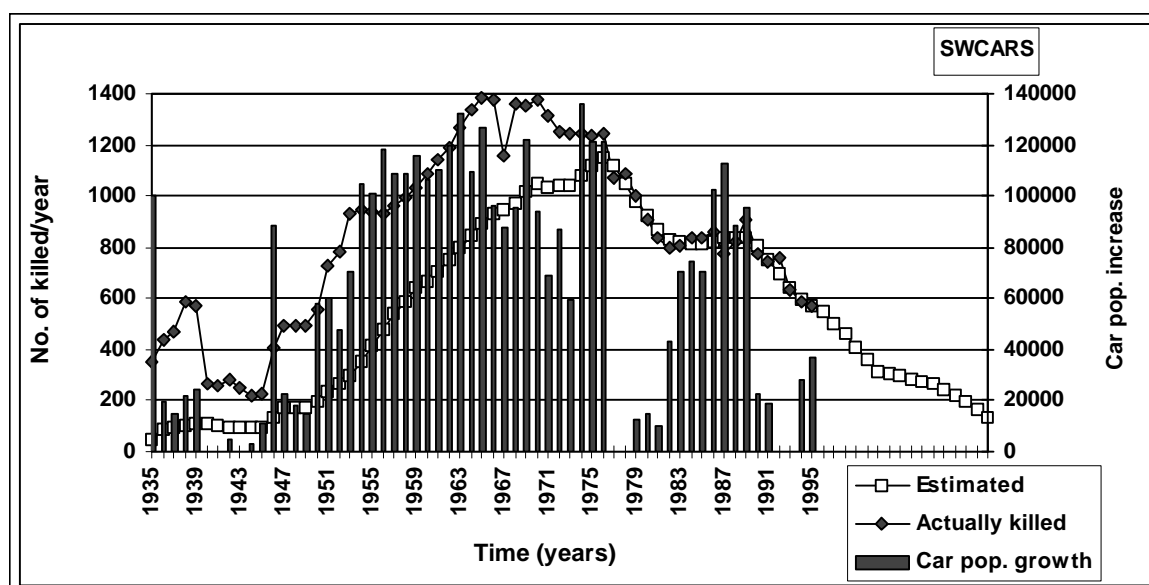


Figure 4: Reported and calculated mortal rates optimised for the period 1976 to 1995

In order to look at the traffic safety development over time, the relation between reported and calculated rates has been plotted in figure 5. Some major milestones in the history of traffic safety work has been added to explain the different trend breaks.

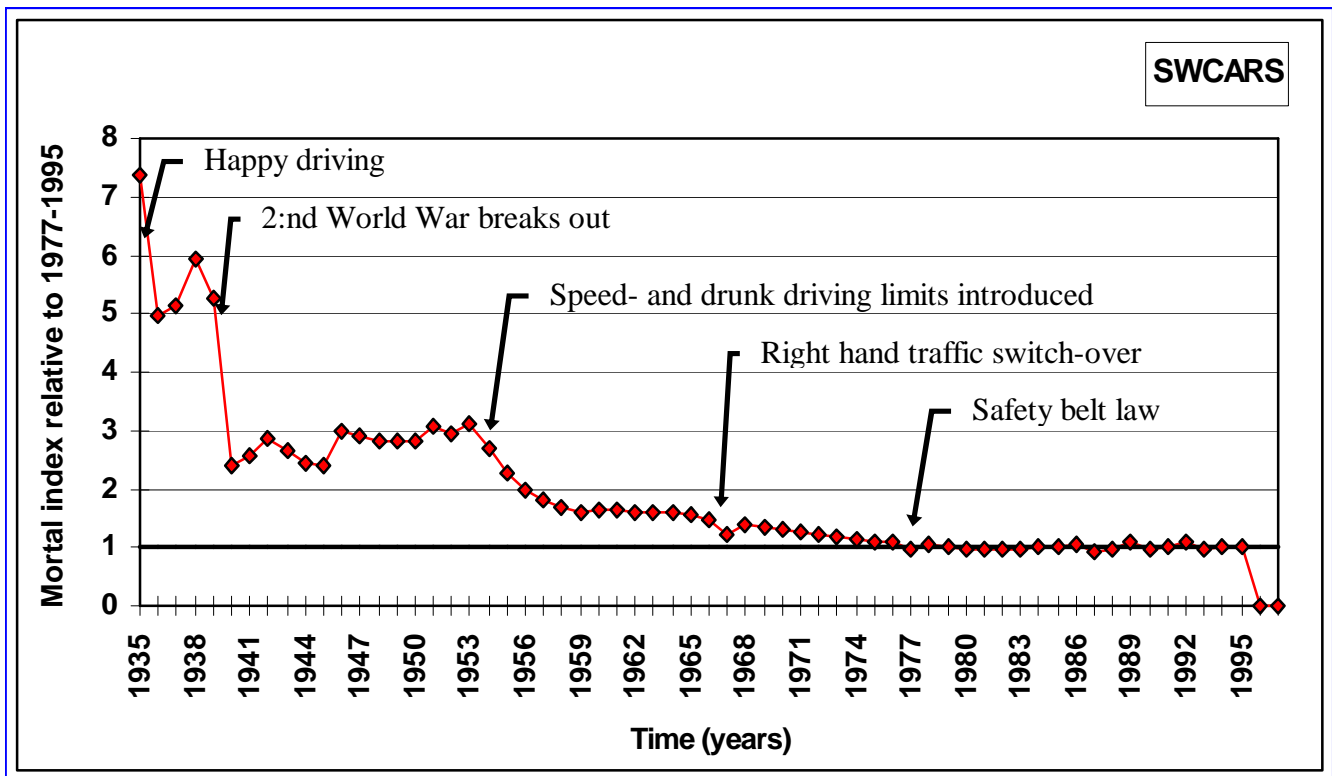


Figure 5: The mortal index vs time relative to the killing rate in 1977-1995.

The first significant drop is due to the out break of the second world war. Probably only professional and experienced drivers were driving due to petrol restrictions. In 1954 an increased emphasis was put on reducing drunk driving. Speed limits were also introduced for citydriving. This seems to have had a dramatic impact on the statistics. In 1967 Sweden gradually changed from left-hand to right-hand driving (between 2-3 o'clock one Sunday morning). The massive propaganda had only a slight effect on the mortal rate. The safety belt law in 1977 seems to be the last change that is barely noticeable in this graph.

CONCLUSIONS

In this paper a new method for the analysis of field data has been presented . Depending on the amount of data and the degree of details that is available different methods can be used for the analysis. The authors would like to draw the following conclusions from the material presented.

1. The method is of general nature and can be used in a wide range of areas.
2. The software is developed as an EXCEL application and can easily be installed and applied on most PC's
3. The trend analysis presented gives a more sensitive response to changes in quality and performance than the traditional cumulative average charts often used.

REFERENCES

- Hallberg, Ö. (1996). Facts and Fiction about the Reliability of Electronics. *Proc. ESREF 1996*, 39-46.
- Oscarsson, P. (1995). *Analysverktyg för driftstatistik*, University of Växjö, Sweden.