

## Analysis Tools for Competing Risk Failure Data

**Cornel Bunea and Roger Cooke**

TU Delft

Mekelweg 4, Delft, The Netherlands

+31 015 2787262, [c.bunea@its.tudelft.nl](mailto:c.bunea@its.tudelft.nl)

[r.m.cooke@its.tudelft.nl](mailto:r.m.cooke@its.tudelft.nl)

**Bo Lindqvist**

NTNU Trondheim

N-7491 Trondheim, NORWAY

+47-735-93532, [bo@math.ntnu.no](mailto:bo@math.ntnu.no)

**Key words:** reliability data, competing risks, statistical tests

**Introduction** Modern Reliability Data Bases (RDB's) are designed to meet the needs of diverse users, including component designers, reliability analysts and maintenance engineers. To meet these needs RDB's distinguish a variety of ways in which a component's service sojourn may be terminated. Up until quite recently, this data was analyzed from the viewpoint of independent competing risk. Independence is often quite implausible, as eg when degraded failures related to preventive maintenance compete with critical failure. The maintenance crew is trying to prevent critical failures while losing as little useful service time as possible; and hence is creating dependence between these competing risks. We have recently learned how to use simple models for dependent competing risk to identify survival functions and hence to analyze competing risk data. This type of analysis requires new statistical tests, and/or adaptations of existing tests. Competing risk models are described in [5]. In this paper we present a number of tests to support the analysis of competing risk data.

Competing risk data may be described as a colored point process, where each point event is described by a number of properties, and where a coloring is a grouping of properties into mutually exclusive and exhaustive classes. For example, a maintenance engineer is interested in degraded and incipient failures, as they are associated with preventive maintenance. He is trying also to take the least expensive maintenance action: repair action or adjustment action are favored above replace action. Critical failures are of primary interest in risk and reliability calculations and a component designer is interested in the particular component function that is lost, in the failure mechanisms and he wishes to prevent the failure of the most expensive components of the system.

In addition to this, two other main operations may be performed on the data: superposition, pooling. Time histories having the same begin and end points may be superposed. The set of event times of the superposition is the union of the times of the superposed processes. In general, superposition is performed in order to obtain a renewal process. If the maintenance team returns components to service as good as new, then all time histories of the components should be superposed.

The pooled data are considered as multiple realizations of the same random variable or stochastic process. When time histories are pooled, these are considered as realizations of the same (colored) point process. In general, pooling is performed on identical independent point processes in order to obtain better statistical estimates of the inter-event distribution.

To perform these operations on data, a set of questions arise, requiring statistical tests to answer these questions:

*Are the time histories homogeneous and independent?*

Independence will fail, if the events for the time histories of the components tend to cluster in calendar time. If homogeneity fails, the uncolored events should not be regarded as

realizations of the same point process. If homogeneity holds, then the number of events up to time  $t$  for every component should not differ significantly.

*Is the coloring stationary?*

The pooled process is now considered as colored. The coloring is stationary if the proportion of "red" and "green" points does not vary significantly with calendar time.

*Is the process "color blind" competing risk?*

The process is color blind if the distribution of the  $i$ -th event is independent of the color of the previous event. Color blindness implies that the processes gotten by splicing together all inter event times beginning with color  $j$ ,  $j=1, \dots, n$ , are homogeneous.

*Is the uncolored process stationary competing risk?*

*Is the uncolored process renewal competing risk?*

In order to answer to this questions, a number of statistical test are presented, together with an application to the data set coming from two identical Compressor Units from one Norsk Hydro ammonia plant, for the period of observation 2-10-68 up to 25-6-89.

**Model assumption 1:** The stratum of  $w$  sockets (units) is homogeneous

Homogeneity within a group of component sockets can be assumed in the case that the components are similar in design, operating circumstances and maintenance regime. Nevertheless, this assumption should always be checked with the failure data.

Suppose we have  $w$  processes  $\{N_i(t)\}$ ,  $i = 1, \dots, w$ , where  $N_i$  is the counting process with intensity  $I_i(t)$

We shall test the hypothesis:

$$H_0: I_1(t) = \dots = I_w(t)$$

Cox and Lewis (1966) proposed the following statistic when all the processes are defined over the same interval:

$$D(N) = \sum_{i=1}^w \frac{(N_i(t) - N^*(t))^2}{N^*(t)}$$

$$\text{where } N^*(t) = \frac{\sum_{i=1}^w N_i(t)}{w}$$

We reject  $H_0$  when  $D > \chi_{w-1, \alpha}$ , where  $\chi_{w-1, \alpha}$  is the  $\alpha$  quantile of the  $\chi^2$  distribution with  $w-1$  degrees of freedom. This test is designed to detect if the  $w$  processes have the same intensity  $I(t)$  under the assumption of Poisson processes (non-homogeneous), not to detect if their intensity is constant in time (homogeneous Poisson process).

To avoid the assumption that  $\{N_i(t)\}$  are Poisson processes, the following approach is recommended by Paulsen et al [10]: Divide the interval  $[0, t]$  in two equal pieces, giving the processes:  $N_{i,1}(s)$ ,  $0 \leq s \leq t/2$  and  $N_{i,2}(s)$ ,  $t/2 < s \leq t$ ,  $i=1, \dots, w$ .

Let  $\pi \in w!$  be a permutation of  $(1, 2, \dots, w)$  and let  $N_\pi = \{N_{\pi(i),1} \otimes N_{i,2}\}$ ;  $i = 1, 2, \dots, w$

where

$$N_{\pi(i),1} \otimes N_{i,2} = \begin{cases} N_{\pi(i),1}(s) : 0 \leq s \leq t/2 \\ N_{\pi(i),1}(s) + N_{i,2}(s) - N_{i,2}(t/2) : t/2 < s \leq t \end{cases}$$

We reject  $H_0$  if  $D(N_\pi)$  is in the upper or lower 2,5% quantile.

**Model assumption 2:** The processes are independent

We assume that there are no clusters of failures in the stratum of sockets. In the extreme case all the sockets in the stratum could fail together in a small interval on the calendar time. This would indicate a strong dependency between the sockets and a possible common external cause of the failures in the sockets. This assumption should therefore always be checked.

We want a test to detect ‘clustering’ of failure events from the various processes, that is, the tendency of events in the different processes to occur close in time. Such a clustering means either that there are common peaks in the intensities  $I_i(t)$  of the processes or that the processes are stochastically dependent.

$H_0$ : ”No clustering across processes”

We choose one of the  $w$  processes  $N_i(t)$ , with a ‘medium or large, number of events. For simplicity we choose the process  $N_1$  and let the events of this process occur at times  $T_1, \dots, T_k$ . Define intervals  $(T_i - \Delta, T_i + \Delta)$  covering these events. There is thus defined a set  $S_1(2\Delta)$  on the time axis, namely the union of the intervals of length  $2\Delta$ ; let this set have length  $T_1(2\Delta)$ . If there are  $N_1(T)$  events then  $T_1(2\Delta) < N_1(T) 2\Delta$ , with equality if and only if no two events occur in  $2\Delta$  and no events occur in  $(T - \Delta, T]$  where  $T$  is the end of the period of observation.

Let

$n = N(T) = \sum_{k=2}^w N_k(t)$  is the total number of events of the processes

$N^*(T) = \sum_{k=2}^w N_k^*(t)$  is the total number falling in any of the intervals of length  $2\Delta$

For this analysis we consider the approximation in which  $N_1(T)2\Delta \ll T$  and  $T_1(2\Delta) = N_1(T) 2\Delta$ . Then under the null hypothesis  $N^*(T)$  has a binomial distribution with index  $n$  and parameter  $T(2\Delta)/T$ .

$T(2\Delta)/T = N_1(T)2\Delta/T = p_T(2\Delta)$  is the probability of any observation of the process  $N_i(t)$ ,  $i > 2$  falling in one of the intervals with length  $2\Delta$ .

$E(N^*(T)) = n p_T(2\Delta)$ ;

$\text{Var}(N^*(T)) = n p_T(2\Delta) (1 - p_T(2\Delta))$

Using a Normal approximation, we reject  $H_0$  when

$$\frac{N^*(T) - np_T(\Delta)}{\sqrt{np_T(\Delta)(1 - p_T(\Delta))}} > U_\alpha$$

where  $U_\alpha$  is the  $\alpha$  quantile of  $N(0,1)$ .

If  $N^*(T)$  is small (less than 5) we should use the binomial or Poisson distribution rather than  $N(0,1)$ . It is suggested that  $\Delta$  be chosen so that  $P_T(\Delta) \approx 0.1$

**Model assumption 3:** The colored process is stationary

The coloring is stationary if the proportion of, say, “red” and “green” points, does not vary significantly with calendar time.

We construct a test to detect whether the ratio of colored events (R and G) is constant over time. Let

$P_R(t) = P\{\text{event occurring at time } t \text{ is colored R}\}$

$P_G(t) = P\{\text{event occurring at time } t \text{ is colored G}\}$

We test the hypothesis

$$H_0 : P_R(t) = p_R \quad (\text{independent of time})$$

We test  $H_0$  by dividing the interval  $(0, T)$  in two. Let  $N_{R,1}$  = number of R events in  $(0, T/2]$  and  $N_{R,2}$  = number of R events in  $(T/2, T)$ .

Given the total number of events  $n_1 = N(T/2)$  and  $n_2 = N(T) - N(T/2)$ , we have that  $N_{R,1}$  and  $N_{R,2}$  under  $H_0$  are independent and binomial distributed with probability  $p_R$ . Using a normal approximation and the result that if  $X \sim N(\mu_1, \sigma_1^2)$  and  $Y \sim N(\mu_2, \sigma_2^2)$  then  $X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$ , we reject  $H_0$  when

$$\frac{|N_{R,1} / n_1 - N_{R,2} / n_2|}{\sqrt{(1/n_1 + 1/n_2)(1 - N_R / n)N_R / n}} > U_{\alpha/2}$$

where  $n = n_1 + n_2$ ,  $N_R = N_{R,1} + N_{R,2}$ , using  $p_R^* = N_R/n$ .

**Model assumption 4:** The process is “color blind competing risk”

The competing risk processes is describe by a n-tuple of risk or colors, and a countable set of n-tuples of random variables  $(X_{11}, \dots, X_{1n})$ ,  $(X_{21}, \dots, X_{2n})$ ,  $(X_{31}, \dots, X_{3n}) \dots$ . After the  $(i-1)$ -th event, the next event occurs at time  $\min\{X_{i1}, \dots, X_{in}\}$  and is assigned the color of the index realizing the minimum. Hence, we observe  $(\min\{X_{i1}, \dots, X_{in}\}, C(i))$ , where  $C(i)$  is the color of the  $i$ -th event.

The process is colored blind if the distribution of  $(\min\{X_{i1}, \dots, X_{in}\}, C(i))$  is independent of the color of the  $(i-1)$ -th event. Ignoring time dependence of a scale which is small relative to the

expected inter-event times, color blindness implies that the processes gotten by splicing together all inter-event times beginning with color  $j$ ,  $j = 1, \dots, n$ , are homogeneous.

Is the colored processes ‘renewed’ to the same degree by R-events and G-events? If not, the colored process might be a superposition of R and G processes (which could be renewal).

Let,

$\{X_{R,k}\}_{k=1,2,\dots}$  the R – process – the sequence of intervals starting with R  
 $\{X_{G,k}\}_{k=1,2,\dots}$  the G – process – the sequence of intervals starting with G

We test

$H_0 : \{X_{R,k}\}$  and  $\{X_{G,k}\}$  are realizations of identical processes

We reduce  $H_0$  to two ‘subhypotheses’:

1)  $H_{0,1} : I_R(t) = I_G(t)$  (the uncolored R and G processes have the same intensity)

2)  $H_{0,2} : P_R(R) = P_G(R)$  (the fraction of R-events in the R-process and G-process are the same)

1) We consider the comparison of  $k$  Poisson processes ( $k > 2$ ) and we test for equality the parameters  $\lambda_1, \dots, \lambda_k$ , no special type of alternative being specified.

Suppose that in fixed time periods  $t_{0(1)}, \dots, t_{0(k)}$ , the number of events observed are  $n_1, \dots, n_k$ , where  $n_i$  is the observed value of a Poisson variable  $N_i$ , of mean  $\mu_i = \lambda_i t_{0(i)}$

The best known test is based on the index of dispersion, i.e. on the statistic

$$d = \sum_{i=1}^k \frac{(n_i - t_{0(i)} \lambda^*)^2}{t_{0(i)} \lambda^*},$$

where  $\lambda^* = n/t_0$  and  $t_0 = \sum_{i=1}^k t_{0(i)}$  is the total time of observation  $n$

$d$  tends to a chi-square distribution with  $k - 1$  degrees of freedom [4].

In our case, we reject  $H_{0,1}$  if:

$$D = \frac{(N_R(T_R) - T_R \lambda^*)^2}{T_R \lambda^*} + \frac{(N_G(T_G) - T_G \lambda^*)^2}{T_G \lambda^*} > \chi_{1,\alpha}^2$$

where

$$\lambda^* = \frac{N_R(T_R) + N_R(T_G)}{T_R + T_G}$$

and  $N_R(T_R)$  is the number of red events in the red process and  $N_R(T_G)$  is the number of red events in the green process.

An alternative test of Lindqvist [9] could be applied to test identity within the class of homogeneous Poisson, nonhomogeneous Poisson and Weibull trend renewal process.

2) We test for

$$H_{0,2} : P_R(R) = P_G(R)$$

Is identical with the test used for ‘stationary coloring’ (see above). Let the R-process have  $n_R$  events and let  $N_{RR}$  of these be R. Let the G-process have  $n_G$  events and let  $N_{GR}$  of these be R.

We reject  $H_{0,2}$  if:

$$\frac{|N_{RR} / n_R - N_{GR} / n_G|}{\sqrt{(1/n_R + 1/n_G)(1 - N_R / n)N_R / n}} > U_{\alpha/2}$$

where

$$N_R = N_{RR} + N_{GR}, n = n_G + n_R$$

**Model assumption 5:** The process is stationary competing risk

Characteristics of a stationary series of events of importance to this work are:

- the distribution of the number of events in an interval of the time window depends only on the length of the interval;
- the expected number of events in a interval of the time window is proportional to the length of the interval;
- there exist no trend in the mean rate of occurrence of failure events throughout the length of the time window.

In our case this means that the process  $(\min\{X_{i1}, \dots, X_{in}\}, C(i))$ ,  $i = 1, 2, \dots$ , is stationary. Since the coloring has already been found to be stationary, this is equivalent to asking whether the uncolored process is stationary.

We test for trend in the rate of occurrence. The uncolored process has inter-event times  $T_1, T_2, \dots, T_n$

1) Laplace test [7]

If we consider a trend in the rate of occurrence represented by a smooth change in time

$$\lambda(t) = e^{\alpha + \beta t},$$

then we test for the null hypothesis  $\beta = 0$ . The probability density that in the interval  $(0, T]$  events occur at  $T_1 \leq T_2 \leq \dots \leq T_n$  is

$$\frac{n!}{T^n} (0 \leq T_1 \leq \dots \leq T_n \leq T)$$

To interpret this relation consider  $n$  random variables  $U_1, \dots, U_n$  independently uniformly distributed over  $(0, T]$ , each having the p.d.f.  $1/T$ . The joint p.d.f. is  $1/T^n$ . Now examine the corresponding order statistics  $U_{(1)} \leq \dots \leq U_{(n)}$ . To calculate their p.d.f. at argument  $T_1, T_2, \dots, T_n$ , note that  $n!$  different original sequences  $U_1, \dots, U_n$  each with the same p.d.f. lead to the same sequence  $U_{(1)} \leq \dots \leq U_{(n)}$ . That is, conditionally on  $n$ , the positions of the events in a Poisson process are independently uniformly distributed over the period of observation  $T$  with mean  $T/2$  and variance  $T^2/12$ .

And from central limit theorem we have

$$\frac{\sum_{i=1}^n T_i - nT/2}{T\sqrt{n/12}} \rightarrow N(0,1) \text{ when } n \rightarrow \infty$$

We reject  $H_0$  when

$$\frac{\sum_{i=1}^n T_i - nT/2}{T\sqrt{n/12}} > U_\alpha$$

2) The above test depends on taking the Poisson process as null hypothesis. We presented now a test in which the intervals  $X_1, \dots, X_n$  are independent and identically distributed, not necessarily exponentially distributed, that the series is a renewal process.

If a plausible functional form can be chosen for the distribution of the  $X_i$ 's, a special test can be constructed by taking the alternative hypothesis a model in which the parameters in the distribution are suitably chosen functions of the  $z_i$ 's.

We arrange the  $X_i$ 's in increasing order

$$X_{(1)}, X_{(2)}, \dots, X_{(n)}$$

Let  $r(i)$  be the rank of  $X_i$  and define the score for  $X_i$ :

$$S_i = 1/n + 1/(n - r(i) + 1)$$

We can use an asymptotic normal distribution for

$$\sum_{i=1}^n S_i (z_i - \bar{z})$$

with mean zero (by symmetry) and variance

$$\sum_{i=1}^n (z_i - \bar{z})^2 \left[ 1 - \frac{1}{n-1} (1/n + \dots + 1/2) \right]$$

We take  $z_i = i$  and reject  $H_0$  when:

$$|V| > U_{\alpha/2}$$

where

$$V = \sum_{i=1}^n S_i \left(i - \frac{n+1}{2}\right) / \sigma^2$$

$$\sigma^2 = \sum_{i=1}^n \left(i - \frac{n+1}{2}\right)^2 \left[1 - \frac{1}{n-1} (1/n + \dots + 1/2)\right]$$

**Model assumption 6:** The process is renewal competing risk

This model assumption implies that a socket is completely or perfectly repaired, similar to replacement to a new one. The plausibility of this model assumption can be easily questioned. Yet, we have a major modeling benefit of this assumption by the fact that the series of events is now a renewal process.

A renewal process is a process in which the intervals between events are independently and identically distributed.

It remains to test for independence of inter-event times  $X_i$ . Let  $\rho_i = \text{corr}(X_k, X_{k+i})$ . Then if  $(X_k, X_{k+i})$  are independent  $\rho_i = 0$ . But the reciprocal is not always true.

$H_0$ :  $\{X_i\}$  are independent ( $\rho_i = 0$ )

From [7] the most useful result is for large sample:

$$\text{var}(\rho^*_{j}) \approx \frac{1}{n-j} \sum_{i=-\infty}^{\infty} \rho_i^2 = \frac{1}{n-j} (1 + 2 \sum_{i=1}^{\infty} \rho_i^2)$$

For a renewal process

$$\text{var}(\rho^*_{j}) = \frac{1}{n-j}$$

$\rho^*_1 \sqrt{n-1}$  will have a unit normal distribution if  $n$  is large and  $\rho_1 = 0$ , so that we reject  $H_0$  when

$$|\rho^*_1 / \sqrt{(n-1)}| > U_{\alpha/2}$$

where

$$\rho^*_1 = \frac{\sum_{i=1}^{n-1} (X_i - X^*)(X_{i+1} - X^*)}{\sum_{i=1}^n (X_i - X^*)^2} \quad \text{and} \quad X^* = \frac{\sum_{i=1}^n X_i}{n}$$



**Model assumption 7:** The process is Poisson competing risk

Given that the  $X_i$ 's are iid, we may test whether these are exponential (thus giving the homogeneous Poisson process). We test for independent exponentially distributed times between events against the alternative of independent times between events with a Gamma distribution.

$H_0$ : Poisson process

$H_1$ : Renewal process

We reject  $H_0$  if [7]

$$l_n = \frac{2n(\log X^* - (1/n) \sum_{i=1}^n \log X_i)}{1 + \frac{n+1}{6n}} > \chi_{n-1, \alpha}^2$$

The numerator of  $l_n$  is exactly the maximum likelihood ratio criterion for homogeneity of variances, if we regard the  $X_i$ 's as variance estimates from a normal population having two degrees of freedom. We recall that an exponentially distributed random variable is proportional to a chi-squared variable with two degrees of freedom. Under the null hypothesis  $l_n$  has approximately a chi-squared distribution with  $n - 1$  degrees of freedom. The divisor was introduced by Bartlett (1937) to improve this approximation.

### Statistical Analysis of Competing Data Sets

The data set proposed for discussion of different competing risk models comes from one Norsk Hydro ammonia plant operating two identical compressor units, for the period of observation 2-10-68 up to 25-6-89. This yields 21 years of observation and more than 350 events. As every modern reliability data base, this data base has the following compressor unit history:

- Time of component failure
- Failure mode: leakage, no start, unwanted start, vibration, warming, overhaul, little gas stream, great gas stream, others
- Degree of failure: critical, non-critical
- Down time of the component
- Failure at the compressor unit: 1 - first unit failed, 2 - second unit failed, 3 - both units failed
- Failure of System and Sub-System of Compressor unit
- Action taken: immediate repair, immediate replacement, adjustment, planned overhaul, modification, others

Date	Calendar time	Operation time	Revision time	Failure mode	Degree of failure	Failure time	Degree time	Sys. Time	Sys. Degree time	Down time (h)	Unit	System	Subsystem	Subfailure mode	Action taken
10/2/68	0	0	0			0	0	0	0	0					
10/3/68	1	1	1	UST	C	1	1	1	1	1.25	1	SMO	PUM	UST	AKR
10/6/68	4	4	4	UST	C	3	3	4	4	136	2	SPE	ANN	LIG	AKR
10/6/68	4.5	4.5	4.5	UST	C	0.5	0.5	0.5	0.5	0.08	1	SPE	ANN	LIG	AKR
11/7/68	36	36	36	VIB	N	32	36	35	36	0	3	SMO	PUM	VIB	AKR
12/8/68	67	67	67	LEK	N	31	31	63	67	0	2	SPE	TAN	LEK	AKU
1/2/69	92	92	92	UST	C	25	88	25	88	5.33	2	SPE	FOR	ANN	JUS
10/2/69	112	R	E	V	I	S	I	O	N	1272	3/16/69				
4/23/69	203	150	38.2	LEK	N	58	83	150	150	0	1	KOM	ROR	LEK	AKR
4/23/69	203.25	150.25	38.4	LEK	N	0.25	0.25	0.25	0.25	0	2	KOM	ROR	LEK	AKR
4/23/69	203.5	150.5	38.6	LEK	N	0.25	0.25	150	150	0	2	INS	TRY	LEK	AKR
4/23/69	203.75	150.75	38.8	LEK	N	0.25	0.25	58	83	0	2	SPE	VEN	LEK	AKR
7/18/69	289	236	124	LEK	N	86	86	86	86	0	2	SPE	VEN	LEK	JUS
8/3/69	305	252	140	UST	C	16	160	252	252	154	1	ELM	ANN	UST	AKR
8/18/69	320	267	155	OVH	N	15	31	31	31	0	3	SPE	FIL	OVH	OVH
8/20/69	322	269	157	LEK	N	2	2	233	233	0	3	SMO	SEP	LEK	AKR
8/28/69	330	277	165	UST	C	8	25	10	185	0.5	1	SPE	VEN	LEK	AKR
8/28/69	330.5	277.5	165	UST	C	0.5	0.5	127	277	1.25	1	INS	NIV	ISI	AKU
9/4/69	337	284	172	UST	C	7	7	15	283	2.5	3	SMO	PUM	UST	JUS
10/4/69	367	R	E	V	I	S	I	O	N	720	11/3/69				
12/1/69	425	342	28	LEK	N	58	73	65	75	0	2	SPE	ROR	LEK	AKR
12/29/69	453	370	56	VIB	N	28	28	28	28	0	2	SPE	PUM	VIB	AKU
1/2/70	457	374	60	LEK	C	4	90	224	427	15	2	KOM	ROR	LEK	AKR

Failure mode	
LEK	leakage
IST	no start
UST	unwanted start
LIG	little gas stream
STG	great gas stream
VIB	vibration
VAR	warming
OVH	overhaul
ANN	others
Degree of failure	
N	non-critical
C	critical
Action taken	
AKR	immediate repairation
AKU	immediate replacement
JUS	adjustment
OVH	planned overhaul
MOD	modification
ANN	others
UNIT	
1	compressor 1 failed
2	compressor 2 failed
3	both compressor failed
System	
SMO	

Figure 1. Compressor units data

The data base indicates also, that there are 18 revision periods with a duration from 4 to 84 days.

For the analysis of this data base a software tool was developed in the higher order programming language Visual Basic and Excel. From the new friendly user interface the operator can choose two classes of competing risk for the five fields of interest: failure mode, degree of failure, action taken, system/subsystem. The operator can perform an analyze over the whole time of observation or for a certain time window by specifying the limiting dates of interest.

Figure 2. User interface

In order to increase the available observation for the theoretical processes, hence to reduce the uncertainty in model estimation, one may pool the data obtained from unit 1 and 2. This operation has to take into account the assumptions 1 and 2.

The assumption of homogeneity within the 2 units data is accepted at the level  $\alpha = 0.1931$  under the assumption of non-homogeneous Poisson processes and it is rejected at the level  $\alpha = 0.0328$  using the approach proposed by Paulsen et al [10]. The assumption of independence is also accepted at level of significance  $\alpha = 0.95$ .

Having the assumptions of homogeneity and independence validated, the next step can be approach: the statistical analysis of competing risk concept (assumptions 3 and 4). Following the choice of competing risk classes proposed in Bunea et al [3] for different fields of interest, one can obtain the significance levels for rejection the null hypothesis of assumptions 3 and 4 are:

	<b>Assumption 3</b> The colored process is stationary	<b>Assumption 4</b> The process is color blind competing risk	
		$I_R(t) = I_G(t)$	$P_R(R) = P_G(R)$
<b>Failure Mode</b> Class I: LEK,IST,UST,VIB Class II: LIG,STG,VAR,OVH,ANN	0.6312	rejected	0.7889
<b>Action taken</b> Class I: AKR,JUS,OVH,MOD,ANN Class II: AKU	0.8969	rejected	0.5616
<b>Degree of failure</b> Class I: Critical Class II: Non-critical	0.9813	rejected	0.9977
<b>System</b> Class I: SMO,KOM,INS Class II: SPE,ELM,GEA	0.9195	rejected	0.9427
<b>Sub-system</b> Class I: PUM,TAN,FOR Class II: ROR,TRY,...,ANN	0.0783	rejected	0.7152

Table 1. Significance levels of rejection the null hypothesis for model assumptions 3 and 4

Since the “color blind” assumption has been rejected, it is not necessary to check whether the uncolored process is stationary competing risk (Figure 3). However for the state of the art, one can ask whether the uncolored process is stationary. Due to the fact that the homogeneity was accepted under the assumption of Poisson processes, Laplace’s test is used. The null hypothesis is accepted at the significance level  $\alpha = 0.3957$ .

The most important assumption of this analysis, the process is renewal competing risk, is accepted at the level of significance  $\alpha = 0.5067$ .

At this point we have to find the competing risk model, which is appropriate to interpret data. In literature one can find two methods: via graphical interpretation – Bunea et al [3], Cooke et al [5] and via statistical testing – Bunea et al [2], Dewan et al [8]. We propose here to test if the renewal competing risk process is Poisson. The assumption is rejected at the significance level  $\alpha = 0.00128$ .

## Conclusions

Figure 3 shows the main assumptions that have been made through this present work. The validation/invalidation of the assumptions is marked on the graph. The analysis is not complete due to the fact that no tests are available in the literature to test the independence of the colors. One try of testing independent exponential distributions, can be find in Bunea et all [2]. At this point of color "independence" a probabilistic interpretation of data has been developed in a new type of analysis (Cooke et all – [5]). One can see that if the independence assumption holds than a constant hazard rate is characteristic for our process, hence we have to deal with a Poisson process. Nevertheless, the probabilistic analysis performed on this data (Bunea et all [3]), indicates that independent model might hold but exponential distribution not.

## Bibliography

- [1] T.J. Bedford, R.M. Cooke, Probabilistic Risk Analysis: Foundations and Methods, Cambridge University Press, 2001, Chapter 9, pg. 153-190
- [2] C. Bunea, R.M. Cooke, Bo Lindqvist, Maintenance study for components under competing risks, ESREL 2002, pg. 212-217
- [3] C. Bunea, R.M. Cooke, Bo Lindqvist, Competing risk perspective over reliability data bases, MMR 2002, pg. 131-134
- [4] H. Cramer, Mathematical methods of statistics, 1946, Princeton University Press
- [5] R.M. Cooke, The design of reliability data bases, Part I and II, Reliability Engineering and System Safety, vol 51, no 2, pg 137-146 and 209-223, 1996
- [6] R.M. Cooke, T.J. Bedford, Reliability databases in perspective, due to appear in IEEE, December 2002
- [7] D.R. Cox, P.A.W. Lewis, The statistical Analysis of series of events, 1966, Methuen & Co ltd, London
- [8] I. Dewan, J.V. Deshpande, S.B. Kulathinal, On the dependence structure between the time of failure and the cause of the failure in the competing risks model, MMR 2002, pg. 189-191
- [9] Bo Lindqvist, The trend renewal process, a useful model for reliability systems, 1993, Norwegian Institute of Technology
- [10] J. Paulsen, J. Dorrepaal, R.M. Cooke, Per Hokstadt, The design and use of reliability data base with analysis tool

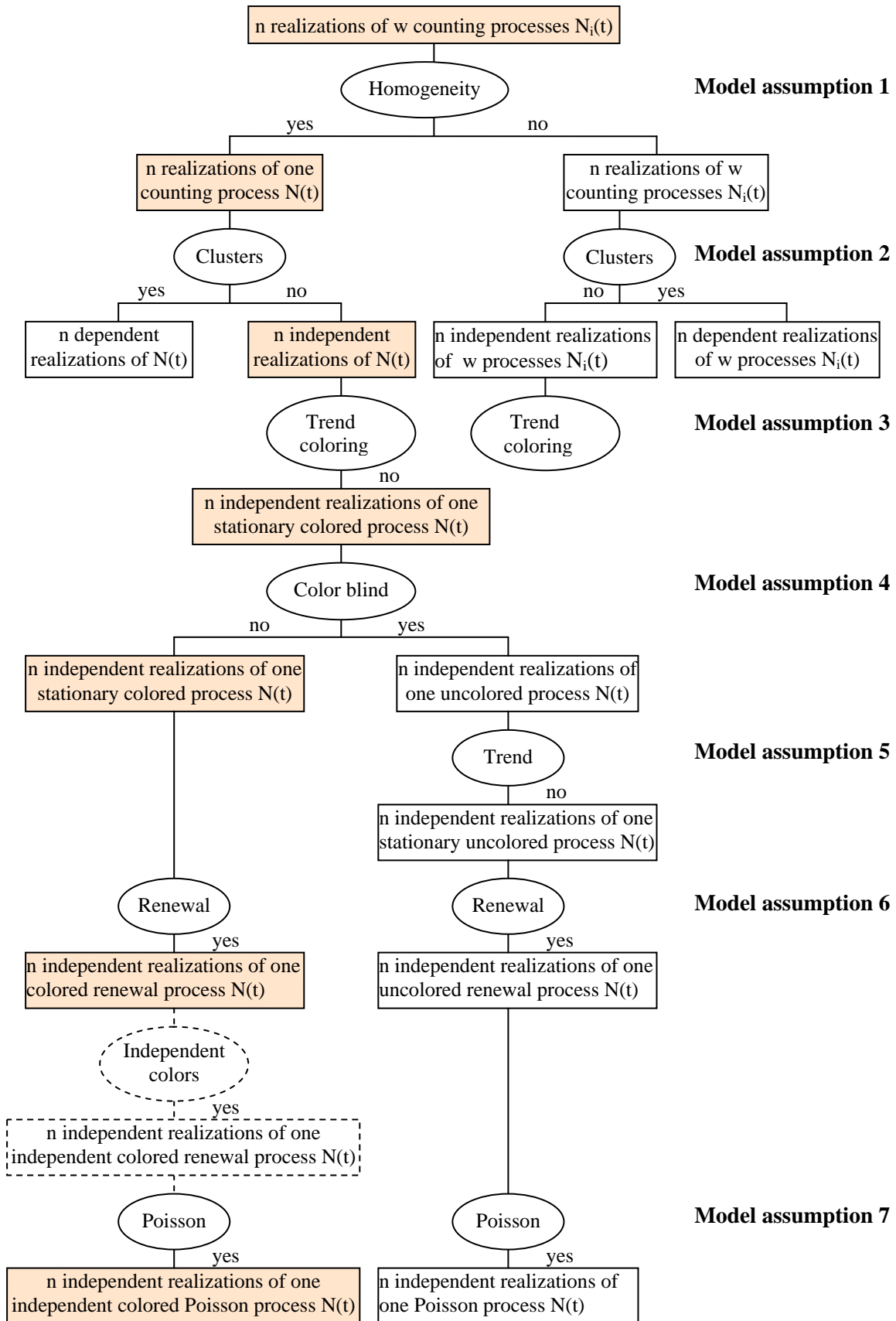


Figure 3. Model assumptions graph