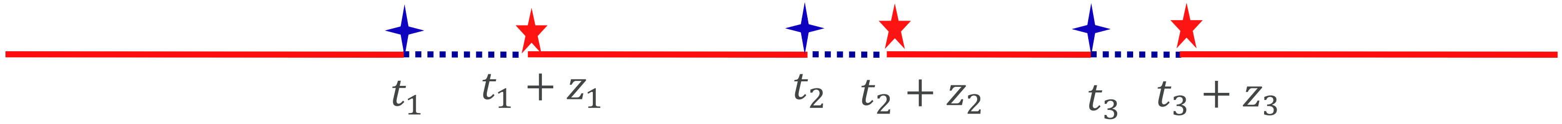
Modelling the failure process of a multi-component system

Prof Shaomin Wu

Kent Business School

Repairable systems: failures, repair,

- Analysis of recurrent event data:

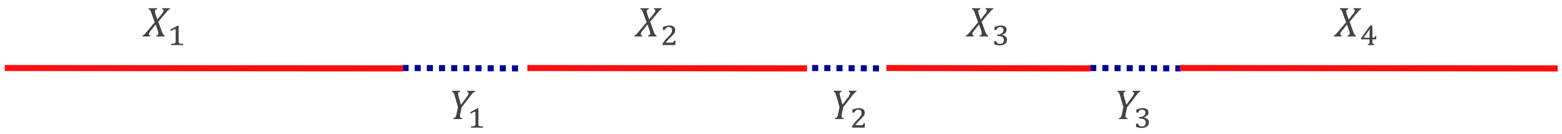


: failed

: repaired

t_k: time when the kth failure occurred

$t_k + z_k$: time when the kth failure is fixed



X_k: the kth working time, or gap time

Y_k: time duration of the kth repair,

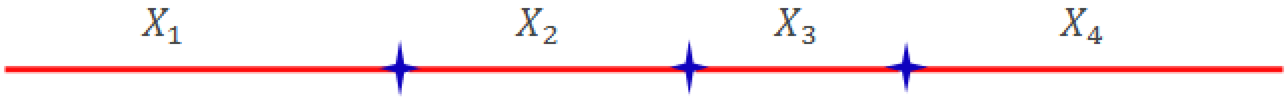
Questions & Applications

- Recurrent event analysis! Below are some applications
  - modelling times between failures of technical systems (software or hardware)
  - modelling the number of insurance/warranty claims
  - modelling the numbers of patient visits to their doctors

- We are interested in the following questions:
  A. How can we estimate the length of working times/gap times, i.e., X_k $(k = 1,2, \dots)$?
  B. How many occurrences are there within a given time period?

Questions A & B

- Recurrent events:



- Notations:

$$S_n = \sum_{i=1}^{n} X_i$$

$$Z_t = \sum_{n=1}^{\infty} \chi\{S_n \le t\} = \sup\{n: S_n \le t\}$$

$$m(t) = E[Z_t]$$

We are interested in

1. **What are the distributions of the gap times? (Question A)**
2. **How many events occurred within a given time? (Question B)**

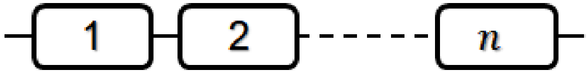
Equivalently,

1. **What are the distributions of X_i?**
2. **How can we estimate $m(t)$?**

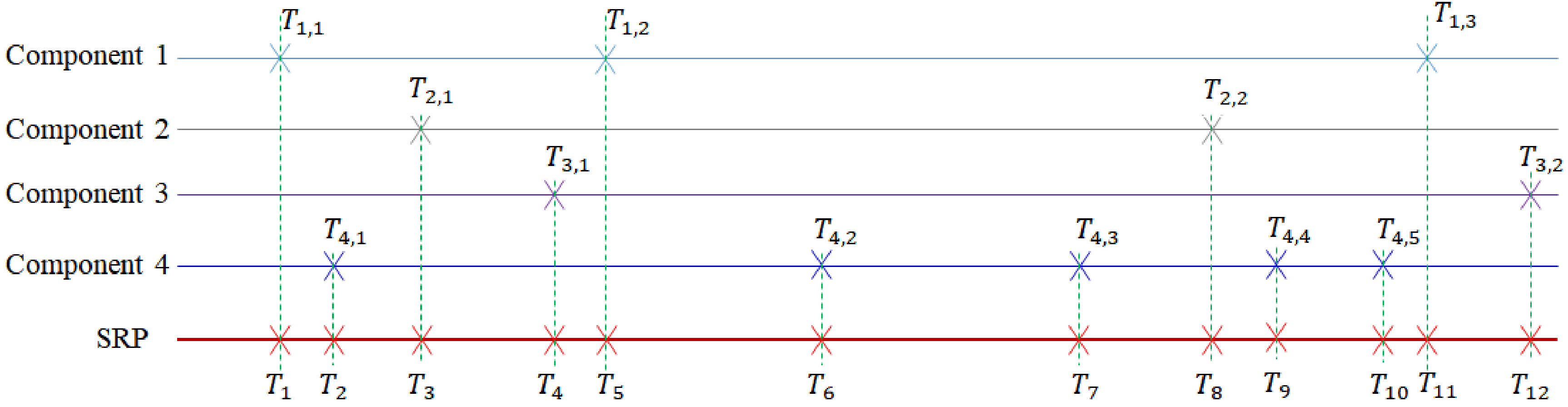
The estimation becomes more complicated for a multi-component system!

The challenge

- Consider a system composed of n sockets in series, there is a component in each socket

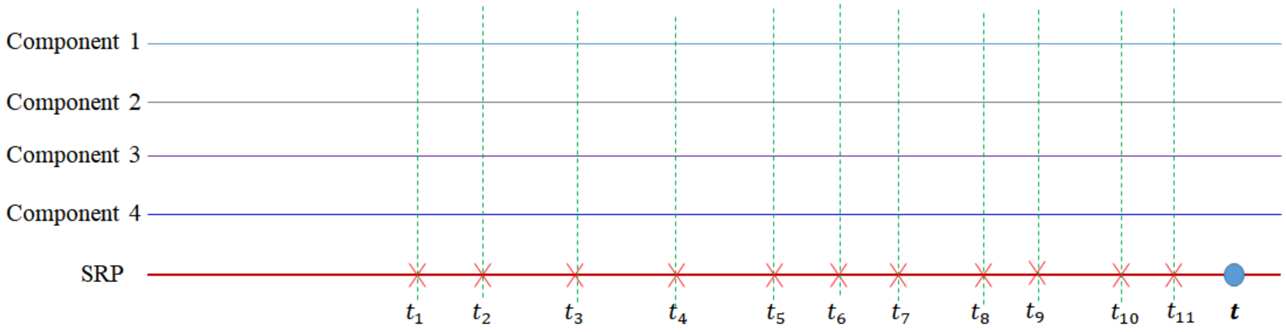


- Superposition of renewal processes:
  - If a component fails, it is replaced with a new, identical one
  - The failure process of the system is a superimposed renewal process (SRP)



The challenge—cont'd

- In the real world
  - The total number of failures of the SRP can be obtained if the reliability of each component is given;
  - Nevertheless, if the times between failures of the system are available, but the components that cause the system to fail are unknown, then the failure data are masked failure data
- However, the SRP requires a model of the failure process of each individual component. In practice,
  - We may not have historical data of which component causes the system to fail
  - We do not have many failure data



- Question: how can we build a model merely based on times-between-failures of the system, without knowing the failure process model of each individual component?

Literature review

- Cox and Smith (1954) prove that the observed failure times of an SRP tend to be distributed with a homogeneous Poisson process when the number of components goes to infinity and the time is far from the origin;

- Khinchin (1956) further clarifies that the SRP tends to be a non-homogeneous Poisson process (NHPP);

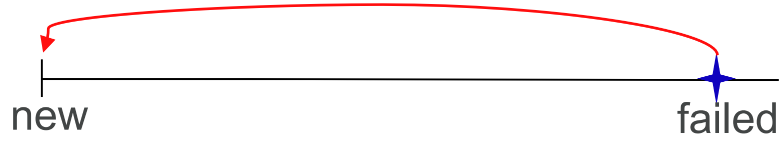
- Drenick (1960) proves the same property as that of Cox and Smith (1954) even if the failure processes of the multi-sources follow heterogeneous renewal processes;

- Recent publications considers methods to approximate the SRP, considering different scenarios (see Wu (2019 b) for a review)

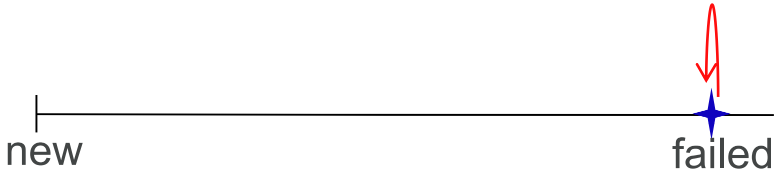
- Conclusion: when the number of components goes to infinity and the time is far from the origin, the SRP can be approximated by NHPP

Effectiveness of a repair

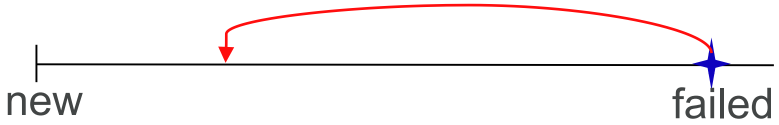
Scenario 1: Perfect repair, i.e., as good as new (AGAN)



Scenario 2: Minimal repair, i.e., as bad as old (ABAO)



Scenario 3: Imperfect repair, between AGAN and ABAO

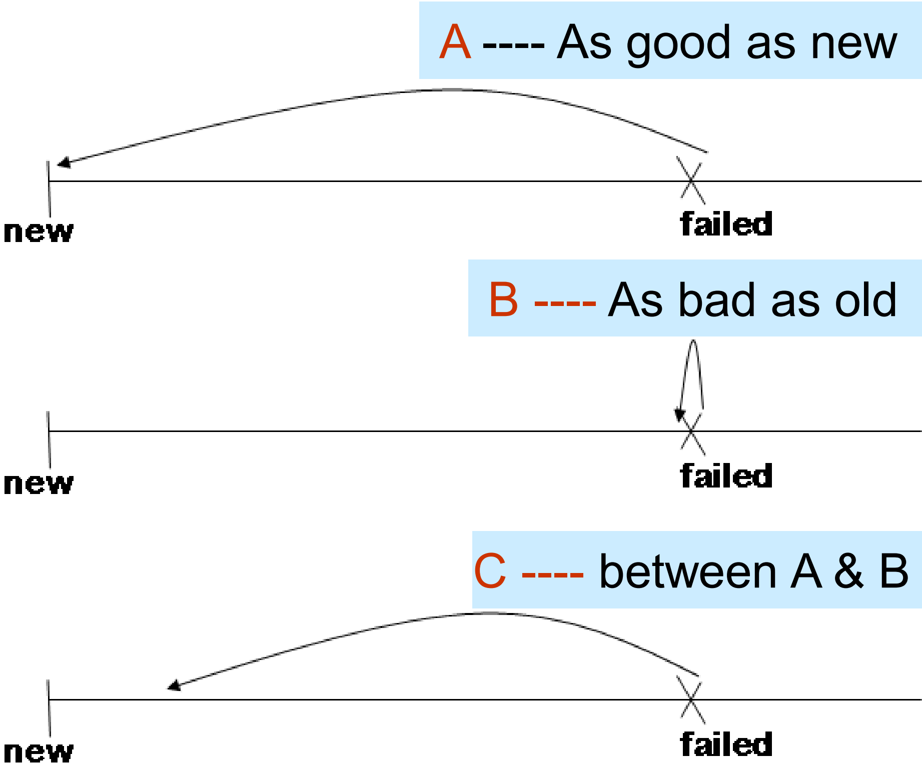


Scenario 4: Better than new repair;
Scenario 5: Worse than old repair

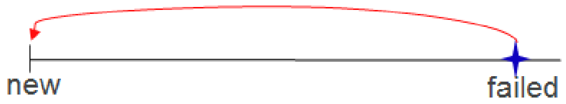
To estimate the distribution of X_k, one can use different methods of modelling the gap times, including renewal processes, NHPP, etc

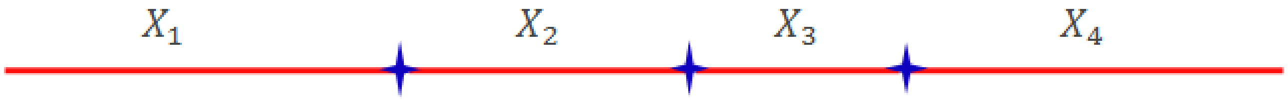
Existing modelling methods

- Differs from life data analysis, where events are assumed to be iid
- Non-parametric methods
- Parametric methods, for example
  - Renewal process (RP)
  - Nonhomogeneous Poisson process (NHPP)
  - Geometric process
  - Cox process (DPP)



Renewal process: The model for **perfect repair**



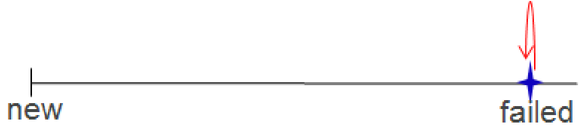


- Given a sequence of random variables $\{X_k, k = 1,2, \ldots\}$, if they are independent and the cdf of X_k is given by $F(x)$ for $k = 1,2, \ldots$, then $\{X_k, k = 1,2, \cdots\}$ is called a **renewal process** (RP)

- The expected number of jumps observed up to some time t, i.e., renewal function, is

$$m(t) = F(t) + \int_0^t m(t-u) f(u) du$$

NHPP: The model for minimal repair



- Nonhomogeneous Poisson process (NHPP). Below are the assumptions
  - $N(0) = 0$ (where $N(t)$ is the number of failures and is a random variable)
  - The number of events in disjoint intervals are independent
  - No events happen simultaneously
- Probabilities of a given number of failures for the NHPP model are calculated by

$$P(N(t) = k) = \frac{\Lambda(t)^k e^{-\Lambda(t)}}{k!}$$

where $\Lambda(t)$ is the cumulative failure intensity, i.e., $\Lambda(t) = \int_0^t \lambda(u)du$

  - If $\Lambda(t) = \lambda_0 t$, then the NHPP reduces to the HPP
  - The most widely used failure intensity function is the power law

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

Models for imperfect repair

- Age-modification methods. Let V_k be the virtual age after the kth maintenance
  - $V_k = V_{k-1} + A_k X_k$, or $V_k = A_k(V_{k-1} + X_k)$

- Intensity-modification methods. Let $\lambda_k(t)$ be the failure intensity of the kth maintenance (A_{k-1} is the effectiveness of repair)
  - $\lambda_k(t) = A_{k-1}\lambda_{k-1}(t)$

- A hybrid method*
  - Given a sequence of non-negative random variables $\{X_k, k = 1,2, \ldots\}$, if they are independent and the cdf of X_k is given by $F(a^{k-1}x^{h(k)})$ for $k = 1,2, \ldots$, where a is a positive constant, $h(k)$ is a function of k and the likelihood of the parameters in $h(k)$ has a known closed form, and $h(k) > 0$ for $k \in N^+$, then $\{X_k, k = 1,2, \ldots\}$ is called a **doubly geometric process** (DGP)
  - $\lambda_k(t) = a^{k-1}\lambda_{n-1}(a^{k-1}t)$ if $h(k) = 1$

Model I

- Consider a series system with three components with hazard functions $0.4t$, $0.45t$, and $0.5t$, respectively. If a component fails and then is replaced, at least a failure intensity of $\inf\{0.4t, 0.45t, 0.5t\} \equiv \phi_1(t)$ is retained. As such, we may assume the intensity function of the system is the sum of two sub-functions

  – Sub-function 1: this part, or$\phi_1(t)$, does not change if a failed component is repaired

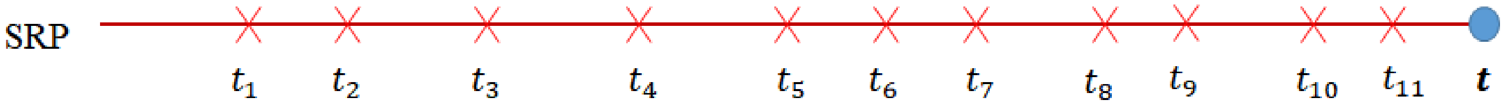
    - We may use the non-homogeneous Poisson process to model this part

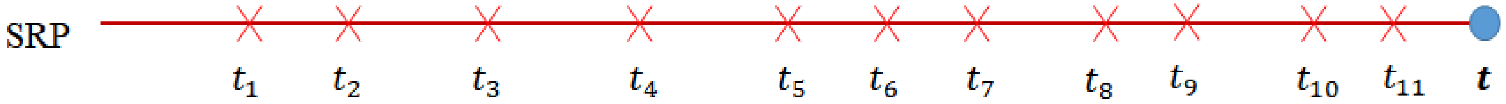
  – Sub-function 2: this part, or $\phi_2(t)$, models the changing intensity function upon repair

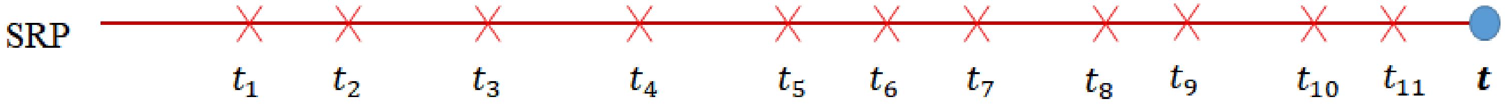
Wu, S. and Scarf, P., 2017. Two new stochastic models of the failure process of a series system. *European Journal of Operational Research*, *257*(3), pp.763-772.

Sub-function 2: Modelling the changing intensity function

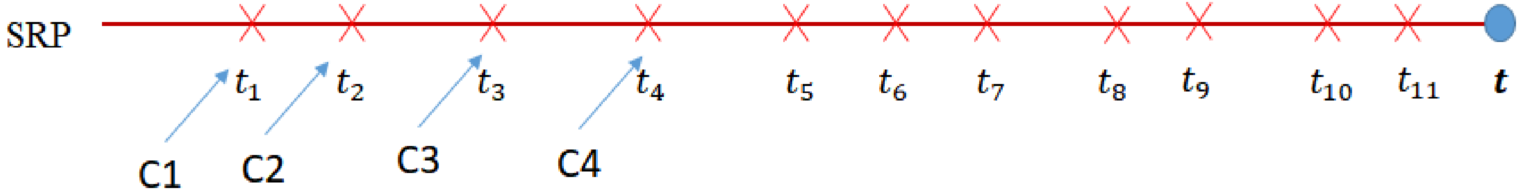
- Assume
  - the number of components in a system is m, which is known
  - the failure rate function of each component is $\frac{1}{m}\lambda_2(t)$
- Modelling the changing intensity function $\phi_2(t)$. If the components are identical at time $t = 0$, a convenient method is to assume that the components fail in turn



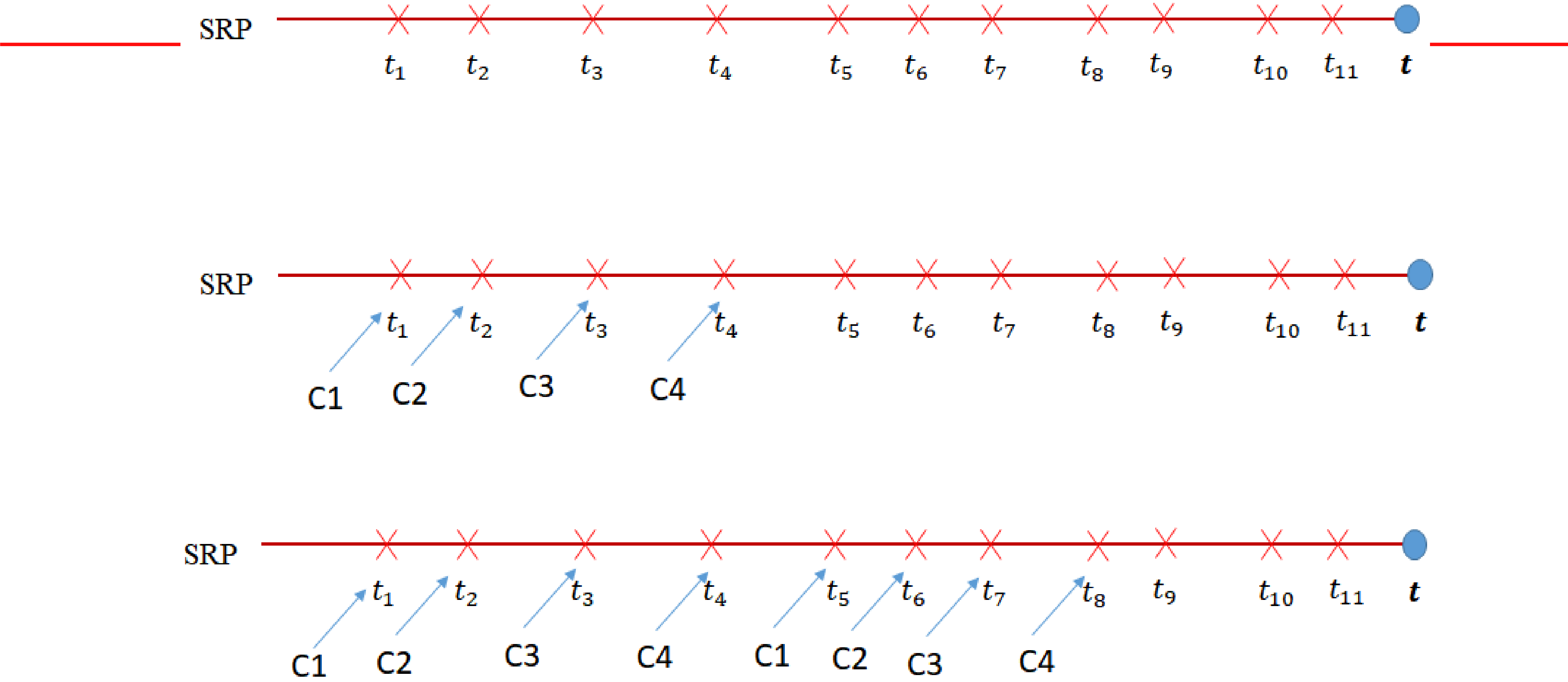

SRP
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} $\boldsymbol{t}$

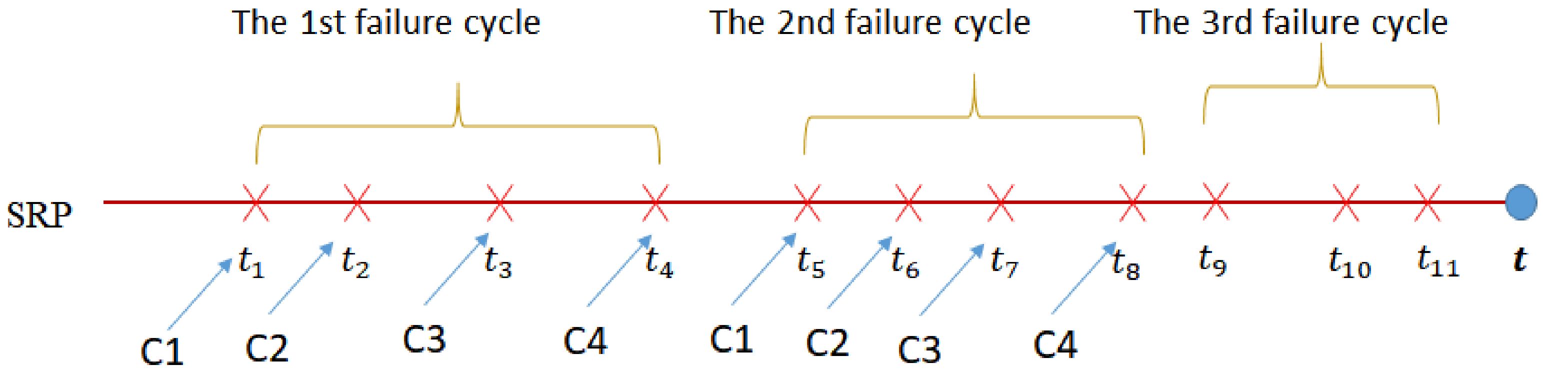


SRP
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} $\boldsymbol{t}$



SRP
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} $\boldsymbol{t}$
C1
C2
C3
C4


SRP
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} $\boldsymbol{t}$
SRP
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} $\boldsymbol{t}$
C1 C2 C3 C4
SRP
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} $\boldsymbol{t}$
C1 C2 C3 C4 C1 C2 C3 C4



The 1st failure cycle
The 2nd failure cycle
The 3rd failure cycle
SRP
t_1
t_2
t_3
t_4
t_5
t_6
t_7
t_8
t_9
t_{10}
t_{11}
t
C1
C2
C3
C4
C1
C2
C3
C4

The model

- Mathematically, the failure intensity of the two parts can be expressed by

$$\phi_1(t|H_{t^-}) = \lambda_1(t)$$

and

$$\phi_2(t|H_{t^-}) = \begin{cases} \lambda_2(t) & \text{if } N_t = 0 \\ \frac{1}{m}\left(\sum_{k=0}^{N_t-1} \lambda_2(t-T_{N_t-k})+(m-N_t)\lambda_2(t)\right) & \text{if } 1 \leq N_t < m \\ \frac{1}{m}\sum_{k=0}^{N_t-1} \lambda_2\left(t - T_{N_t-k}\right) & \text{if } N_t \geq m \end{cases}$$

respectively

Simple moving average method

- Given a time series $x_1, x_2, \dots, x_t$, you are asked to provide a forecast value at time $t+1$, i.e., $\hat{x}_{t+1}$, where $\hat{x}_{t+1}$ denotes a forecast of x_{t+1}

- Let $\hat{x}_{\mathrm{t+1}} = \frac{1}{p}\sum_{i=0}^{p-1} x_{t-i}$, then $\hat{x}_{\mathrm{t+1}}$ is a forecast with the ***simple moving average*** method

Numerical examples: settings

- Settings: Assume
  - There are M identical systems
  - Each system may experience M_j failures (times-between-failures of different systems may be different)
  - Each system is composed of m components, the lifetime distribution of each component follows $F(t) = 1 - \exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right)$, where α is randomly chosen from $(12,60)$ (i.e., $\alpha \in (12,60)$) and β is randomly chosen from $(0.5, 4)$, (i.e., $\beta \in (0.5,4)$)
- RP: renewal process;
- GP: geometric process;
- NHPP-PL: NHPP with the power law;
- GRP: Kijima model I

Numerical examples

			AIC values					
M	M_j	m	RP	GP	NHPP-PL	GRP	Model I	[redacted]
		5	855.73	856.42	856.48	853.53	851.81	840.95
			(77.25)	(74.92)	(73.45)	(75.20)	(75.33)	(74.83)
	15	10	670.60	672.39	663.45	663.22	662.98	659.31
			(50.69)	(50.68)	(48.25)	(48.13)	(47.91)	(47.88)
		15	562.87	564.77	553.70	554.04	553.75	551.34
			(36.72)	(36.79)	(34.74)	(35.33)	(35.18)	(34.96)
10		5	1126.03	1125.65	1125.50	1121.52	1119.72	1109.49
			(107.52)	(103.85)	(102.58)	(105.03)	(104.70)	(105.42)
	20	10	877.23	877.36	868.31	867.82	867.69	862.78
			(68.15)	(65.43)	(64.89)	(64.72)	(64.32)	(65.03)
		15	711.76	711.08	699.44	700.16	700.09	695.76
			(57.32)	(58.90)	(57.95)	(57.82)	(58.04)	(57.09)

Simple exponential method

- Given a time series $x_1, x_2, \dots, x_t$, you are asked to provide a forecast value at time $t+1$, i.e., $\hat{x}_{t+1}$, where $\hat{x}_{t+1}$ denotes a forecast of x_{t+1}

- Let $\hat{x}_{\mathrm{t+1}} = \sum_{i=0}^{p-1} \alpha^{t-i} x_{t-i}$, then $\hat{x}_{\mathrm{t+1}}$ is a forecast with the ***simple exponential smoothing*** method

Wu, S., 2019 (a). A failure process model with the exponential smoothing of intensity functions. *European Journal of Operational Research*, 275(2), pp.502-513.

Assumptions

- Suppose the failure process of a real series system of multiple components. Once the system fails, the failed component can be immediately identified and replaced with a new identical one;
- Assume that there are m components, which have failure rate functions $\frac{1}{m}\lambda_0(t)$, $\frac{1}{m}\rho\lambda_0(t), \frac{1}{m}\rho^2\lambda_0(t), \dots, \frac{1}{m}\rho^{m-2}\lambda_0(t)$, and $\frac{1}{m}\rho^{m-1}\lambda_0(t)$, respectively;
- The effectiveness of repair is: once a repair is conducted, the component with intensity function $\frac{1}{m}\lambda_0(t)$ is replaced, and the intensity function of other component changes from $\frac{1}{m}\rho^{m-k}\lambda_0(t)$ to $\frac{1}{m}\rho^{m-k-1}\lambda_0(t)$

Model II: failure process model with the exponential smoothing of intensity functions

- The model

$$\lambda(t|\mathscr{H}_{t-}) = \begin{cases} \frac{1}{m}\sum_{k=0}^{m-1}\rho^{m-k-1}\lambda_0(t), & \text{if } N_t = 0, \\ \frac{1}{m}\left(\sum_{k=0}^{N_t-1}\rho^{m-k-1}\lambda_0(t-T_{N_t-k}) + \sum_{k=N_t}^{m-1}\rho^{m-k-1}\lambda_0(t)\right), & \text{if } 1 \le N_t < m, \\ \frac{1}{m}\sum_{k=0}^{m-1}\rho^{m-k-1}\lambda_0(t-T_{N_t-k}), & \text{if } N_t \ge m. \end{cases}$$

- ESI: Exponential Smoothing of Intensity
- MAI: Moving Average of Intensity for the model with $\rho = 1$

Experimental results on simulated data

The means and standard deviations of $(-\log(L))$ from 30 repetitions.

n	m	Estimated values of $(-\log(L))$ Nine existing models									New models	
		$q=2$		$q=3$					$q=4$		$q=3$	$q=2$
		RP	NHPP-PL	GP	Kijima I	Kijima II	ARI_m	ARA_m	BBIP	Model II	ESI	MAI
	5	45.33	42.40	43.08	41.65	41.17	40.82	41.20	42.18	40.60	40.82	41.00
		(5.07)	(4.47)	(4.21)	(4.25)	(4.35)	(4.21)	(4.29)	(4.21)	(4.23)	(4.25)	(4.28)
15	15	28.73	24.79	26.72	24.18	23.96	24.15	24.25	24.32	23.14	23.84	24.02
		(4.21)	(3.38)	(3.23)	(3.57)	(3.68)	(3.72)	(3.72)	(3.23)	(3.65)	(3.77)	(3.67)
	25	23.16	19.35	21.62	18.32	18.33	18.36	18.31	18.96	17.53	18.01	18.24
		(3.95)	(3.58)	(3.04)	(3.84)	(3.74)	(3.83)	(3.80)	(3.47)	(3.96)	(3.89)	(3.87)
	5	91.80	86.67	88.15	86.02	85.46	85.30	85.51	86.34	85.09	85.22	85.33
		(9.10)	(8.13)	(8.10)	(8.08)	(7.99)	(8.33)	(8.16)	(8.10)	(8.10)	(8.06)	(8.17)
30	15	57.69	50.53	53.60	49.96	49.34	50.02	49.92	49.98	48.42	49.06	49.23
		(7.70)	(6.03)	(5.99)	(6.19)	(6.14)	(6.15)	(6.12)	(5.98)	(6.23)	(6.13)	(6.18)
	25	45.03	36.95	41.76	36.46	35.97	36.26	36.26	36.52	35.22	35.67	35.84
		(7.32)	(5.57)	(5.45)	(5.66)	(5.81)	(5.92)	(5.79)	(5.44)	(5.83)	(5.93)	(5.95)
	5	136.12	127.52	130.17	126.61	125.76	125.60	125.71	127.20	125.15	125.41	125.56
		(13.97)	(11.82)	(12.19)	(11.69)	(11.63)	(11.75)	(11.66)	(11.84)	(11.31)	(11.68)	(11.62)
45	15	84.94	74.17	79.01	73.65	73.12	73.99	73.52	73.49	72.15	72.75	72.96
		(11.66)	(7.87)	(8.77)	(7.95)	(8.75)	(8.24)	(7.93)	(7.86)	(7.73)	(8.06)	(8.02)
	25	64.92	52.24	59.85	51.72	51.31	51.69	51.62	51.68	50.57	51.09	51.30
		(12.03)	(6.77)	(9.13)	(6.94)	(7.07)	(7.20)	(7.11)	(6.71)	(7.00)	(7.02)	(7.09)

Experimental results on real datasets

- The model outperforms many other existing models on real-world datasets and does not need to assume "a failure component is replaced/renewed"

The real-world datasets.

No.	Dataset	n	Data source	Model
1	Hydraulic system (LHD 1)	23	Kumar and Klefsjö (1992)	NHPP-PL
2	Hydraulic system (LHD 3)	25	Kumar and Klefsjö (1992)	NHPP-PL
3	Hydraulic system (LHD 9)	27	Kumar and Klefsjö (1992)	NHPP-PL
4	Hydraulic system (LHD 11)	28	Kumar and Klefsjö (1992)	NHPP-PL
5	Hydraulic system (LHD 17)	26	Kumar and Klefsjö (1992)	NHPP-PL
6	Hydraulic system (LHD 20)	23	Kumar and Klefsjö (1992)	NHPP-PL
7	Air conditioner (TBF 7909)	24	Proschan (1963)	HPP
8	Air conditioner (TBF 7912)	30	Proschan (1963)	HPP
9	Air conditioner (TBF 7913)	27	Proschan (1963)	HPP
10	Air conditioner (TBF 7914)	23	Proschan (1963)	HPP
11	Compressor	24	Yanez et al. (2002)	Kijima I
12	Main propulsion motor	24	Yanez et al. (2002)	Kijima I
13	Powertrain System 510	55	Guida and Pulcini (2009)	BBIP
14	Powertrain System 514	35	Guida and Pulcini (2009)	BBIP
15*	Diesel engine	56	Lee (1980)	NHPP-WLL

* In dataset 15, there is a value 0, which is replaced with 0.5 in this paper.

Comparison on 15 real-world failure datasets

$-\log(L)$ of each model on the real-world datasets.

No.	Estimated value of $(-\log(L))$ Nine existing models									New models	
	$q=2$		$q=3$					$q=4$		$q=3$	$q=2$
	RP	NHPP	GP	Kijima I	Kijima II	ARI_m	ARA_m	BBIP	Model II	ESI	MAI
1	129.99	128.50	129.50	128.46	128.50	128.44	128.46	129.09	128.02	128.32	129.38
2	148.72	146.96	148.72	146.96	145.47	145.31	145.32	146.22	144.94	144.20	144.94
3	166.55	163.64	165.39	163.52	163.65	163.48	163.49	164.36	163.96	163.45	163.97
4	158.05	157.09	157.99	157.09	155.89	155.54	155.86	156.23	155.21	154.91	155.44
5	151.20	149.33	150.96	149.32	149.12	148.38	148.36	149.81	148.98	148.87	149.14
6	137.27	136.86	137.12	136.73	135.65	135.55	135.65	136.61	134.77	135.29	135.80
7	125.37	126.30	124.48	125.37	125.37	125.05	125.37	125.96	125.37	125.37	125.37
8	151.94	150.43	151.14	150.41	150.42	150.37	150.42	150.64	150.20	150.24	151.33
9	143.96	144.22	143.10	143.96	143.10	141.19	142.75	143.81	141.80	142.07	142.15
10	119.60	119.66	119.21	119.52	119.56	118.68	119.57	119.47	117.98	118.58	119.55
11	191.06	189.30	190.95	189.32	188.90	188.70	187.82	189.22	188.78	188.12	188.85
12	183.88	182.44	182.70	181.63	182.45	181.55	181.85	183.29	182.37	182.45	182.91
13	543.26	543.57	543.19	542.28	541.88	543.32	543.58	542.52	542.35	542.35	542.87
14	356.18	357.05	356.06	355.88	355.07	354.55	357.06	353.77	354.74	353.98	355.04
15	369.29	368.31	369.14	368.31	368.00	367.06	367.68	367.70	367.89	367.75	368.07
$-\log(L)^{*}$ [a]	205.09	204.24	204.64	203.92	203.53	203.14	203.55	203.91	203.16	203.06	203.65
AIC*	414.17	412.48	415.29	413.83	413.07	412.29	413.10	415.83	414.32	412.13	411.31
AIC_c^{*} [a]	415.20	413.50	417.07	415.62	414.85	414.08	414.88	418.64	417.13	413.91	412.33
BIC*	416.89	415.19	419.36	417.90	417.14	416.36	417.16	421.25	419.74	416.20	414.02

[a] The value with * on its right upper corner represents the mean of the value.

Performance on the real world datasets

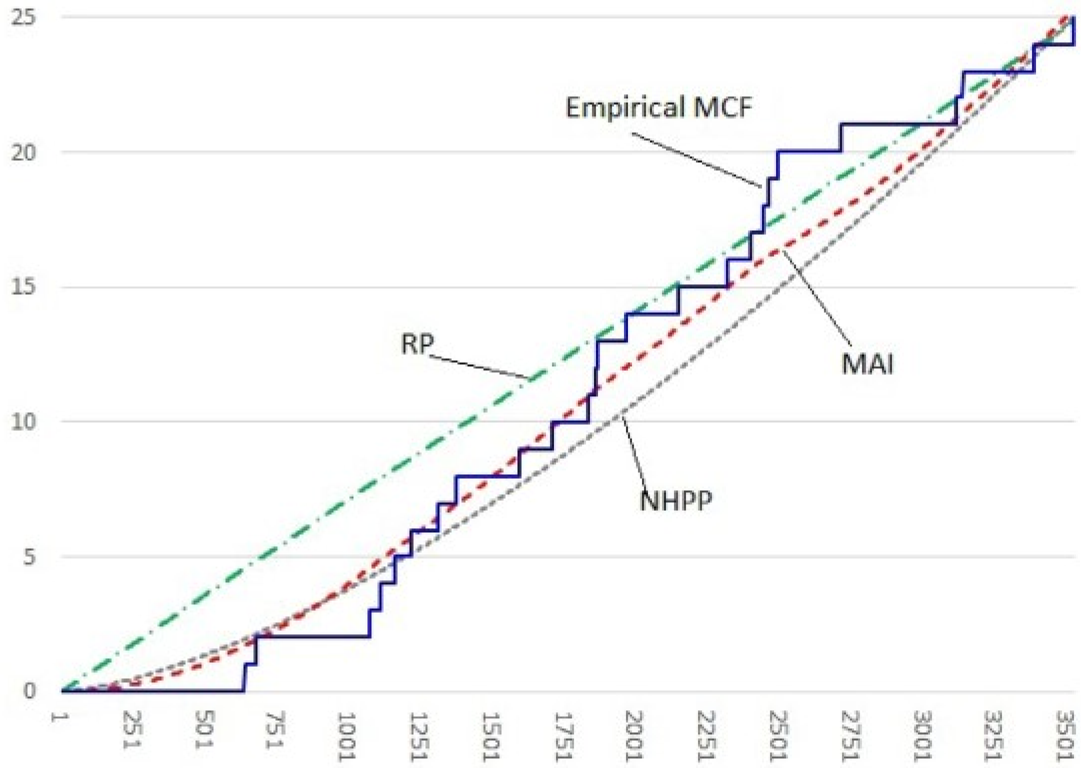
Results of the performance comparison from Table 7.

No.	ESI		MAI	
	$-\log(L)$	$-\log(L)$ of the "$q = 3$" models	AIC	AIC_c & BIC
1	Model II	ESI	NHPP-PL	NHPP-PL
2	ESI	ESI	MAI	MAI
3	ESI	ESI	NHPP-PL	NHPP-PL
4	ESI	ESI	MAI	MAI
5	ARA_m	ARA_m	MAI	MAI
6	Model II	ESI	MAI	MAI
7	GP	GP	MAI	MAI
8	Model II	ESI	NHPP-PL	NHPP-PL
9	ARI_m	ARI_m	MAI	MAI
10	Model II	ESI	MAI	MAI
11	ARA_m	ARA_m	MAI	MAI
12	ARI_m	ARI_m	NHPP-PL	NHPP-PL
13	Kijima II	Kijima II	MAI	MAI
14	BBIP	ESI	ESI	MAI
15	ARI_m	ARI_m	MAI	MAI
Frequency	4 x Model II	8 x ESI	10 x MAI	11 x MAI

- If comparing the values of the AICc and BIC of the 15 models, then one can find that the MAI has the smallest values of the AICc and BIC in 11 out of the 15 cases. That is, the MAI outperforms the other models in terms of both AICc and BIC

Performance on a real-world dataset

- Cumulative air-conditioning failures and estimated mean cumulative functions



Conclusion and further development

- Conclusions
  - Development of methods to estimate the failure process of a series system with masked failure data can find applications in practice
  - Sample size is usually small; there is a need to develop models with a small number of parameters
  - Real datasets are needed to validate any proposed methods

- Further development
  - The above developments were on approaches to approximating the SRP (superposition of renew processes)
  - One may extend the above methods to the situation in which a failed component is maintained with imperfect repair

References

- Cox, D.R. and Smith, W.L., 1954. On the superposition of renewal processes. *Biometrika*, *41*(1-2), pp.91-99.
- Drenick, R.F., 1960. The failure law of complex equipment. *Journal of the Society for Industrial and Applied Mathematics*, *8*(4), pp.680-690.
- Khinchin, A.I., 1956. On Poisson streams of random events. *Theory of Probability and Its Applications*, *1*, pp.248-255.
- Wu, S., 2017. Doubly geometric processes and applications. *Journal of the Operational Research Society*, pp.1-13.
- Wu, S. and Scarf, P., 2017. Two new stochastic models of the failure process of a series system. *European Journal of Operational Research*, *257*(3), pp.763-772.
- Wu, S., 2019 (a). A failure process model with the exponential smoothing of intensity functions. *European Journal of Operational Research*, 275(2), pp.502-513.
- Wu, S., 2019 (b). Superimposed Renewal Processes in Reliability. *Wiley StatsRef: Statistics Reference Online*, pp.1-12.
- Wu, S., 2021. Two methods to approximate the superposition of imperfect failure processes. *Reliability Engineering & System Safety*, *207*, p.107332.

This is a keynote talk given at *2021 IEEE Global Reliability & Prognostics and Health Management Conference (PHM-Nanjing 2021)*, Nanjing, China, October 15-17, 2021.

Thanks the organiser for their invitation