Wu, Shaomin (2021) *Warranty data analysis for quality improvement and economic benefits.* In: 1st International Conference on EMMA-2021 (Engineering, Medicine, Management, Arts and Sciences), 24--26 Dec, 2021, Chennai, India. (Submitted)

Warranty data analysis for quality improvement and economic benefits

Prof Shaomin Wu
Kent Business School
University of Kent

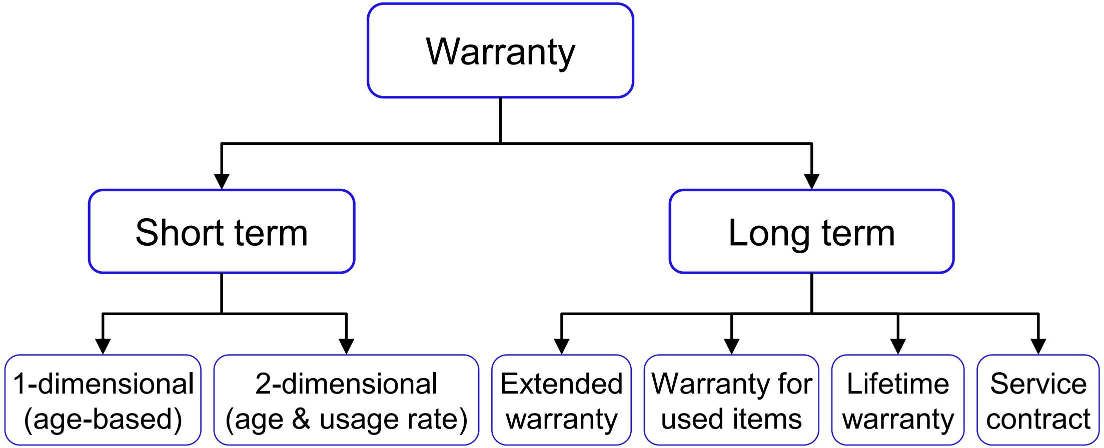
Outline

Definition → Types of Warranty → Warranty Data Analysis → Warranty Policy Optimisation → Conclusions & Future Work

What is warranty?

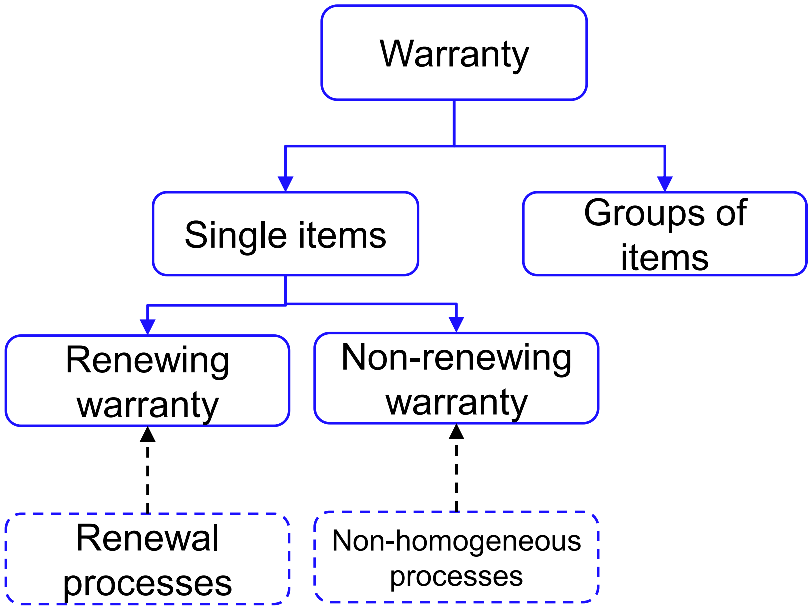
- A warranty is a contractual agreement between the buyer and warrantor (e.g., the manufacturer, retailer, etc) entered into upon the sale of the product or service

- Classification:
  - Base warranty and
  - Extended warranty

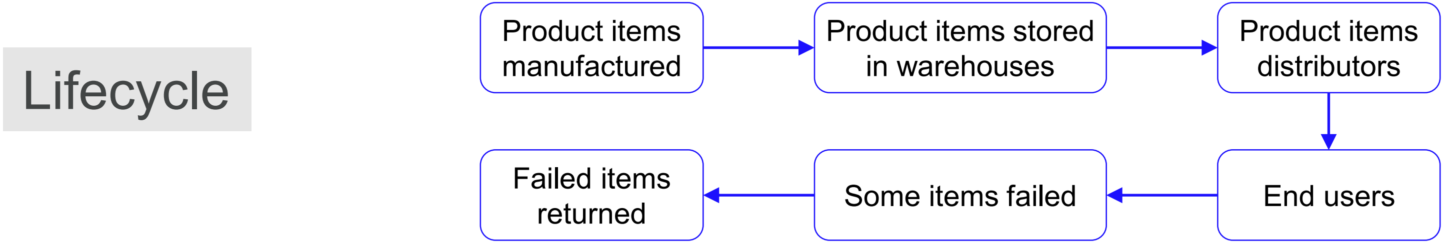


Classification of warranty

- Two types:
  - For single item sales, and
  - For groups of items
- Renewing and non-renewing
  - Under a renewing warranty, a failed item within its warranty duration is replaced by a new one, the warranty is renewed at no charge to or at a partial cost to the buyer.
  - Under a non-renewing warranty, a failed item is replaced/repaired by the warrantor within the original warranty duration, and the original warranty is not renewed.

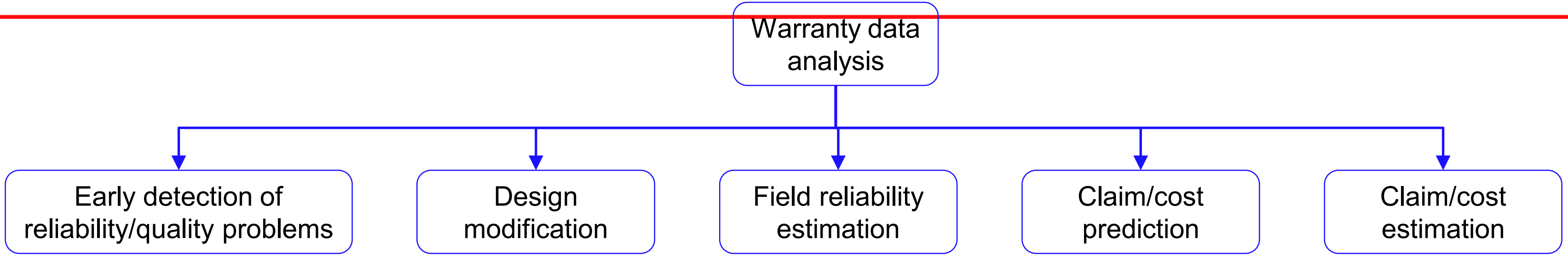


Types of warranty claim data



- Warranty claims data can be grouped into the following four categories:
  - Product related – make, model, failure(s), etc.
  - Service agent related – names, ID numbers, etc.
  - Cost related – materials, repair expenses, etc
  - Customer related –contact details, usage mode and intensity, operating environment, etc

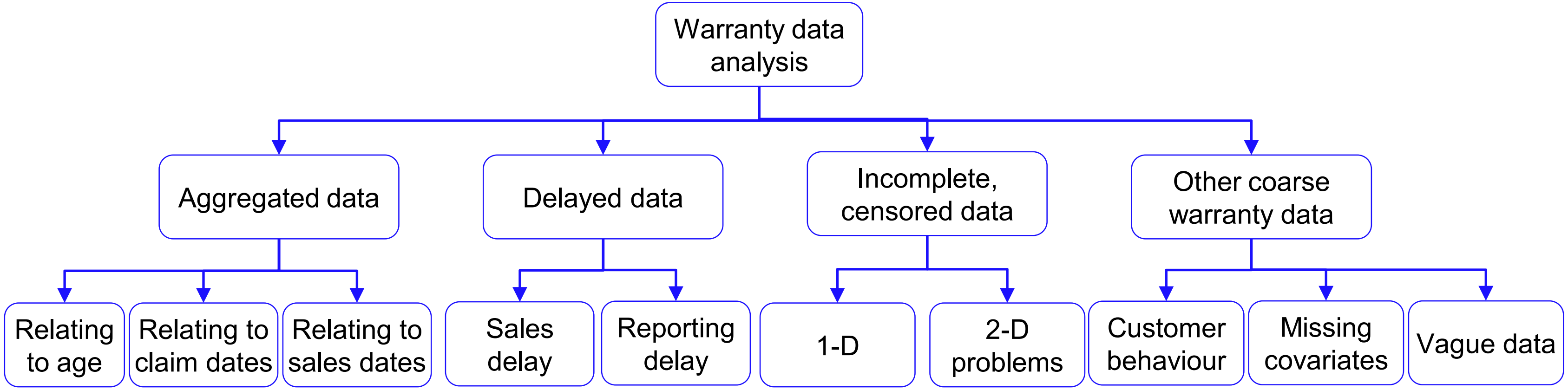
Warranty data analysis



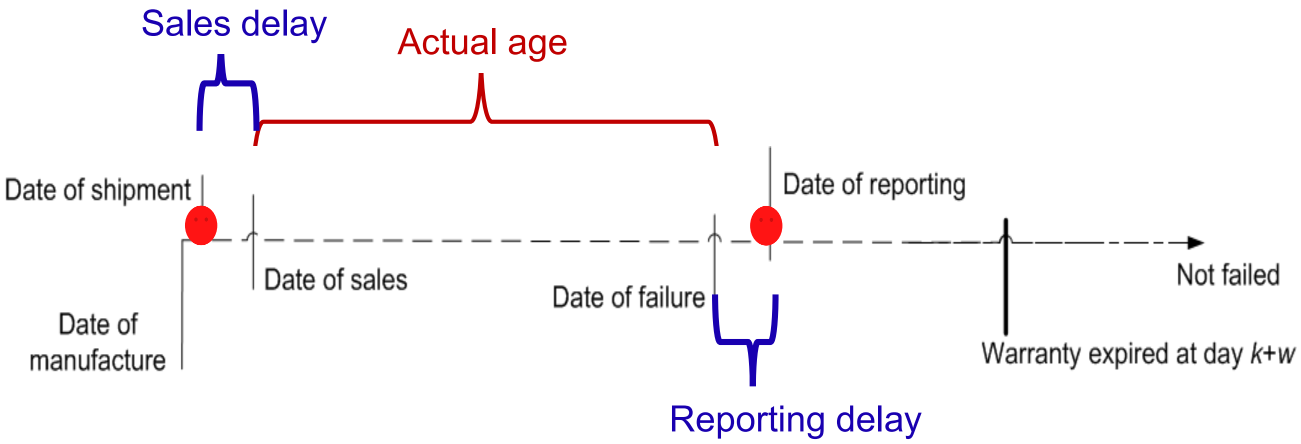
- Early detection of reliability/quality problems: to discover early indications of unexpected quality and reliability problems, where Statistical Process Control may be used
- Design modification: to detect abnormalities from warranty databases, data mining or text mining may be used
- Field reliability estimation: for selecting warranty policy, planning maintenance regimes and preparing spare parts
- Claim/cost prediction: to predict the expected number of claims and/or the respective warranty cost at the warranty coverage
- Claim/cost estimation: warranty claim estimation assumes an infinite population of items, whereas in warranty claim prediction, the population of items that is eventually sold is finite.

Challenges: Data quality-

- Warranty data are usually coarse
  - Aggregated data: data are aggregated, but each individual claim is unavailable
  - Delayed: sales delay and reporting delay
  - Incomplete: Failed but not reported (FBNR); reported but not failed (RBNF)
  - Censored: warranty length vs. lifetime
  - …

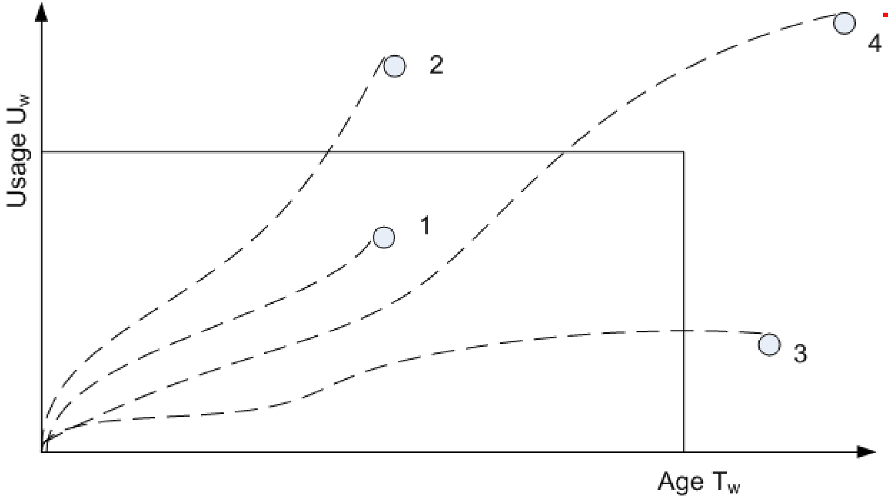


Data quality: sales delay and report delay



Reported age: from the date of shipment to the date of reporting

Data quality--- Incomplete data

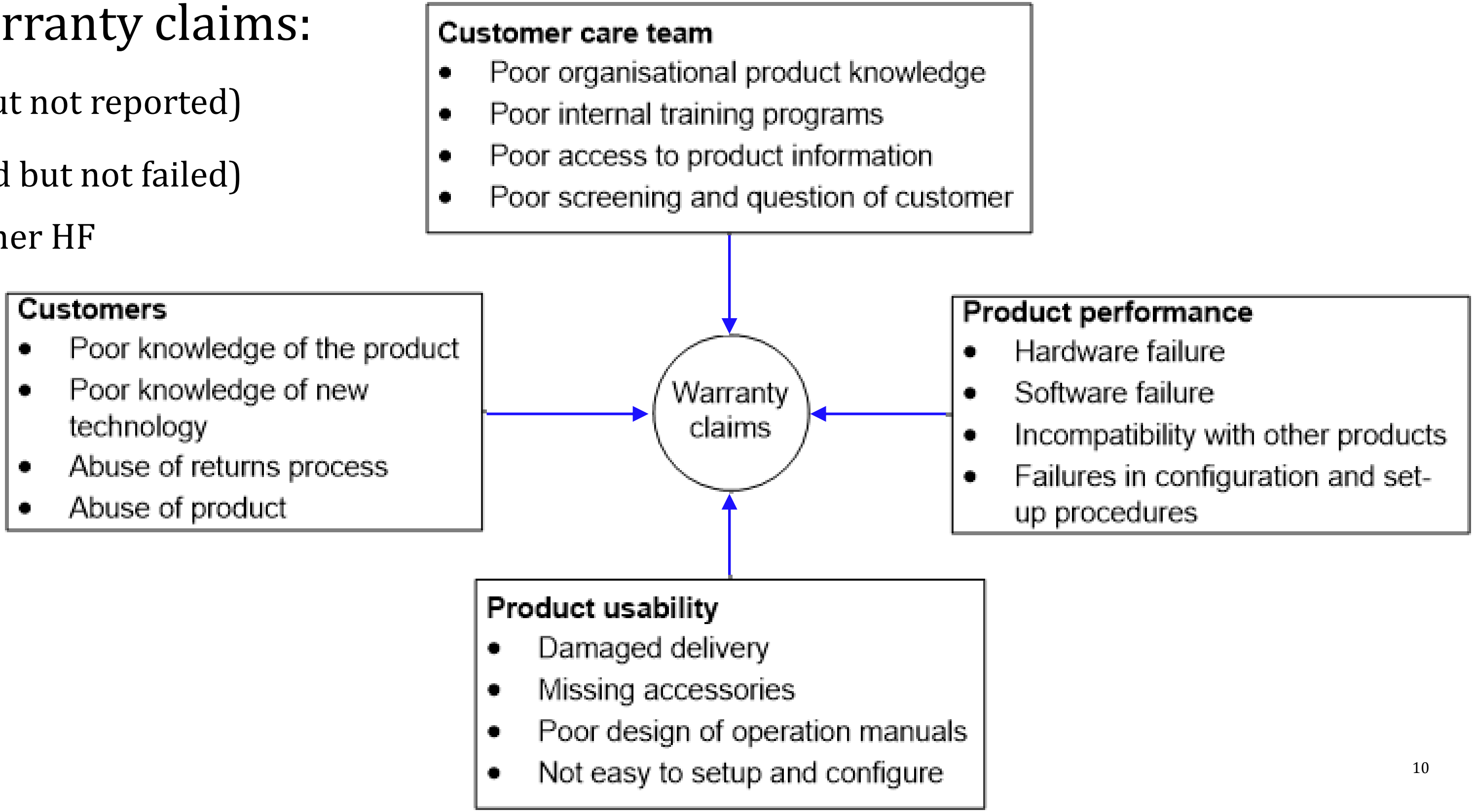


- In lifetime data analysis, both times to failure and times to termination should be known

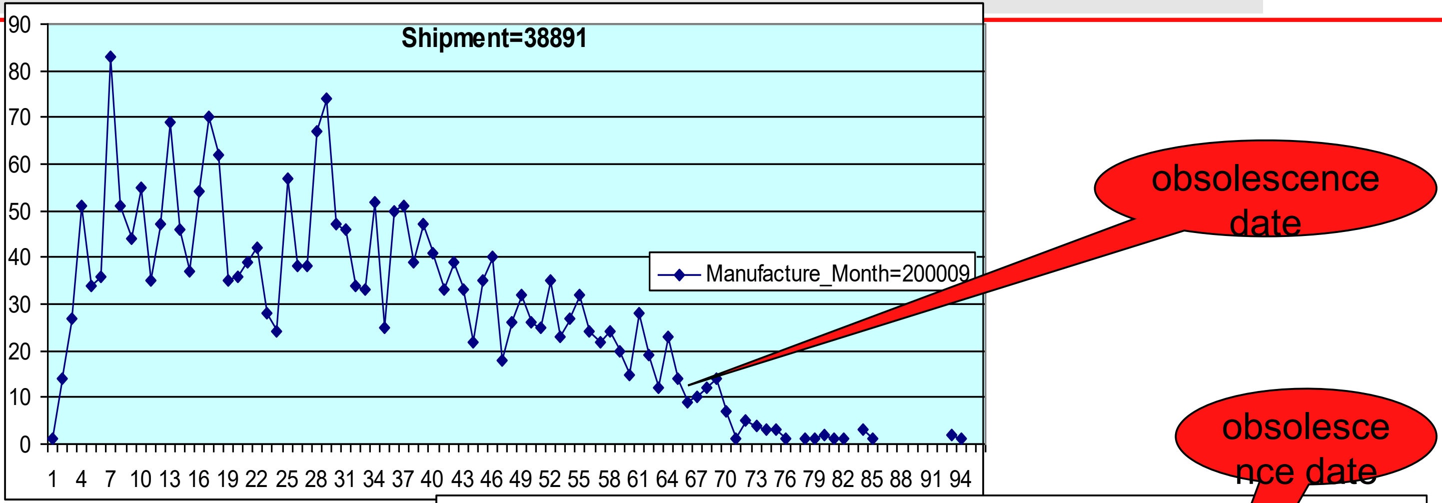
- Item 1 failed within both its age and usage limits and it may be reported to the warrantor
- Item 2 failed within the age limit but beyond the usage limit and its warranty expired;
- Item 3 failed within the usage limit but beyond the age limit, and its warranty expired.
- Item 4 has both the age and usage at failure above the age limit and the usage limit.

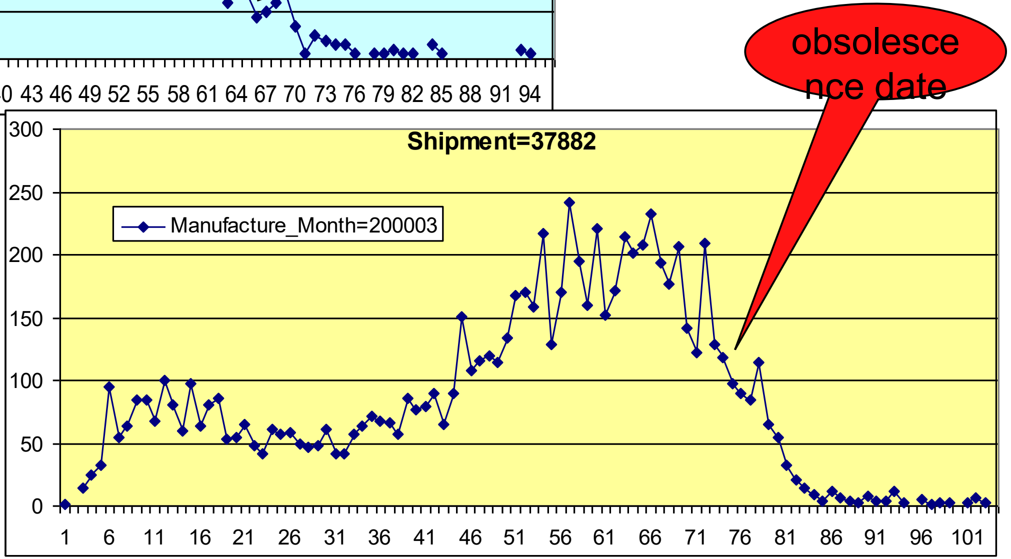
Human factors

- Human factors (HF) can influence on warranty claims:
  - FBNR (failed but not reported)
  - RBNF (reported but not failed)
  - Failure due to other HF



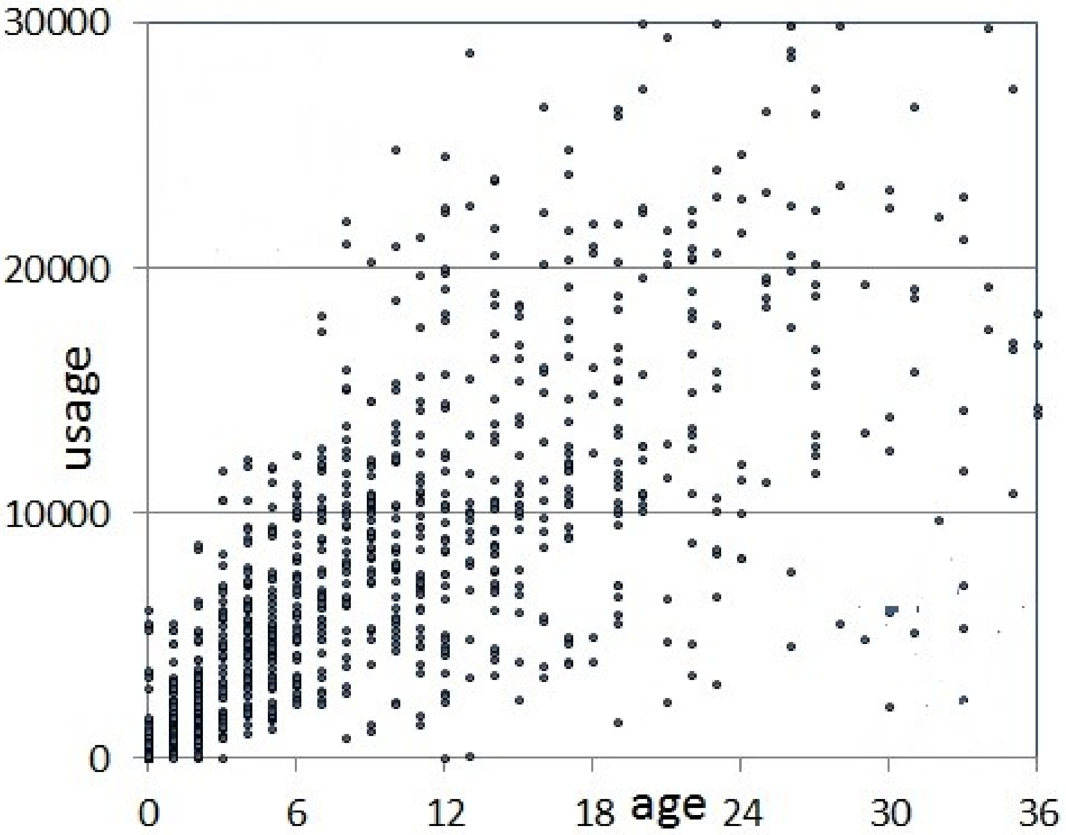
More challenges: design modification, obsolescence date



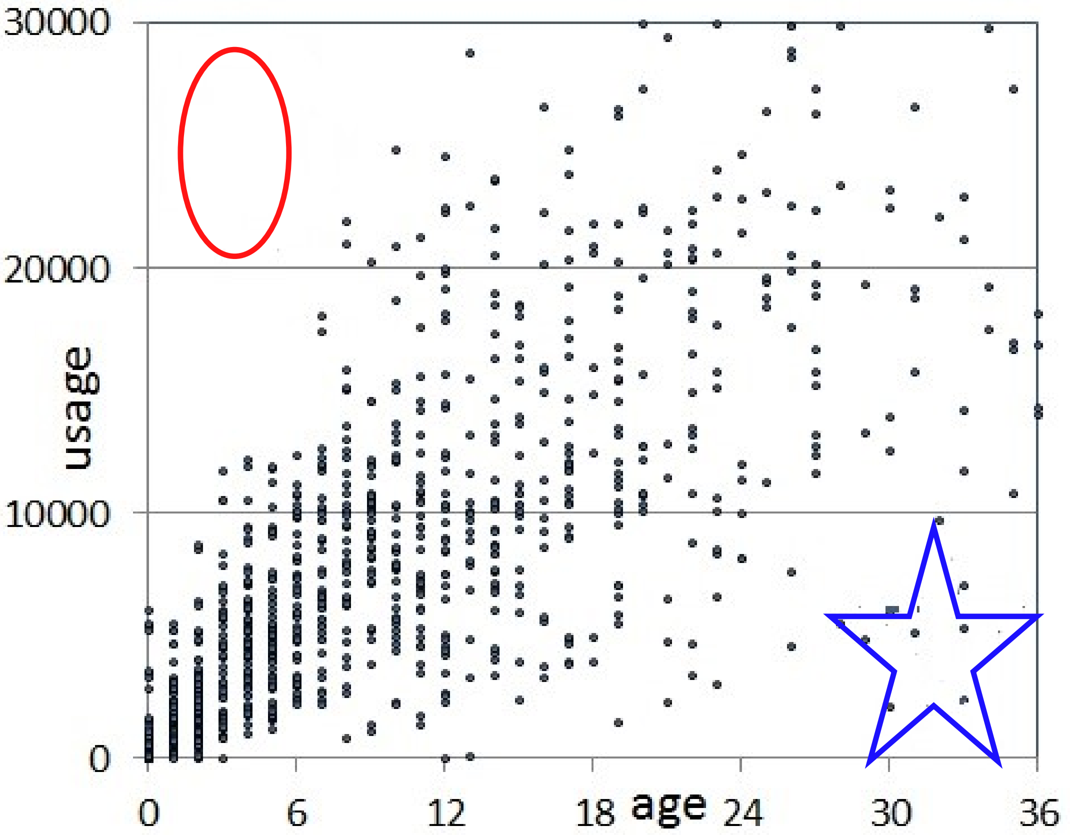


Some research outputs

Warranty claim data



Warranty claim data: asymmetric phenomenon



Interpretation of the asymmetric phenomenon

- The relationship between age and usage
  - If the age of a product is small, its usage should be small. This is because the age is the calendar time and it is not possible to develop large cumulative usage within a short period of the calendar time. Another reason is due to the operating limit, for example, a car usually cannot be driven faster than 100 miles per hour, hence the usage within a time interval is limited.
  - If the age is large, on the other hand, the usage can be small. For example, some cars are not frequently used. Hence, although they are very old, their mileage can be very small.

Copula functions

- Given a random variable X with probability distribution $F_X(X)$. Then $u = F_X(X)$ is uniformly distributed in [0,1]. Likewise, we have $v = F_Y(Y)$ uniformly distributed.

- The joint distribution of X and Y can be written

$$F(X,Y) = P(X < x, Y < y) = P(X < F_X^{-1}(u), Y < F_Y^{-1}(v)) = F(F_X^{-1}(u), F_Y^{-1}(v)) = C(u,v)$$

where $F_Y^{-1}(u) = x, F_Y^{-1}(v) = y$

Copula function

- Sklar theorem: each joint distribution $F(X, Y)$ can be written as a copula function $C(F_X, F_Y)$ taking the marginal distributions as arguments (Sklar, 1959)

- A copula function $z = C(u, v)$ is defined as

  1. $z, u, v \in [0,1]$
  2. $C(0, v) = C(u, 0) = 0, C(1, v) = v$ and $C(u, 1) = u$
  3. For every $u_1 > u_2$ and $v_1 > v_2$ we have

$$V_C(\boldsymbol{u}, \boldsymbol{v}) \equiv C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0$$

Clayton copula and Gumbel copula

- Clayton copula

$$C(v_1, v_2) = \left(v_1^{-\theta} + v_2^{-\theta} - 1\right)^{-1/\theta}$$

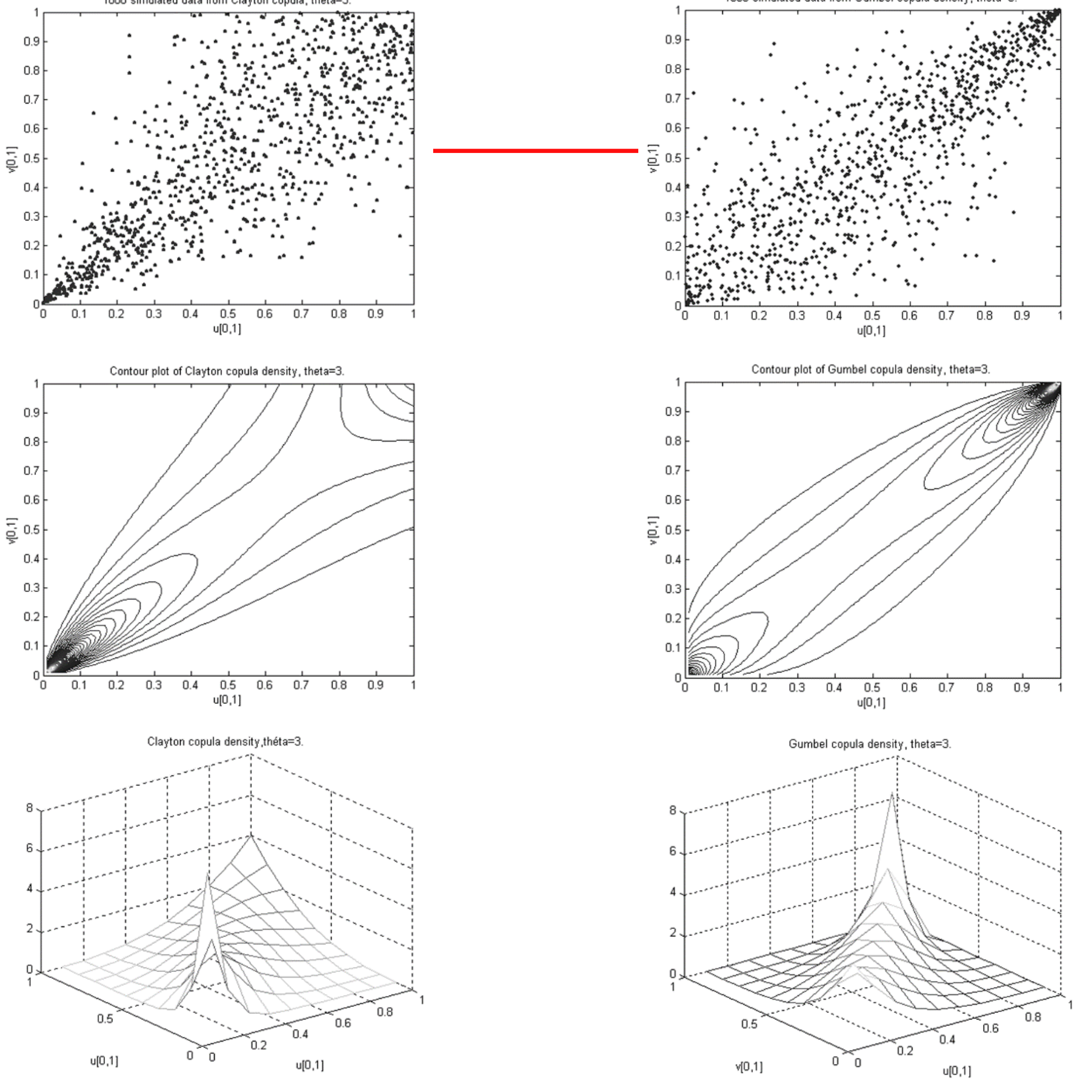
- Gumbel copula

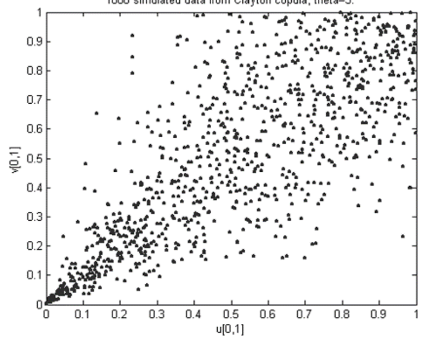
$$C(v_1, v_2) = \exp\left\{-\left[(-\ln v_1)^{1/\theta} + (-\ln v_2)^{1/\theta}\right]^{\theta}\right\}$$

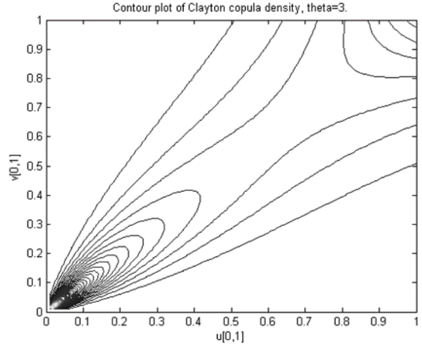
What do they look like?

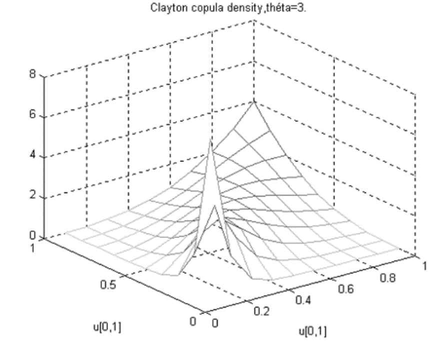
Middle figures: Contour plots of the bivariate copula densities

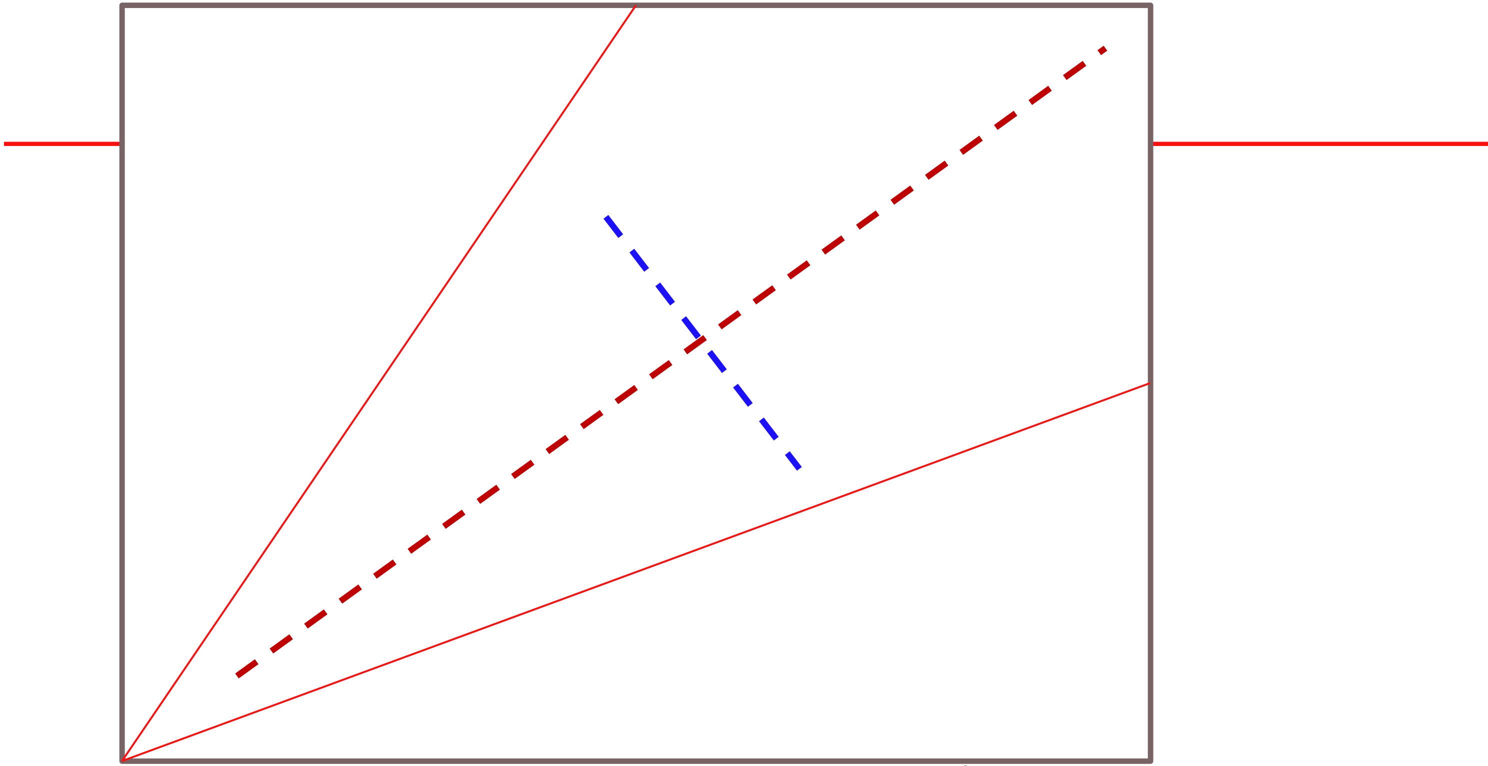
- Clayton copula (left)
- Gumbel copula (right)









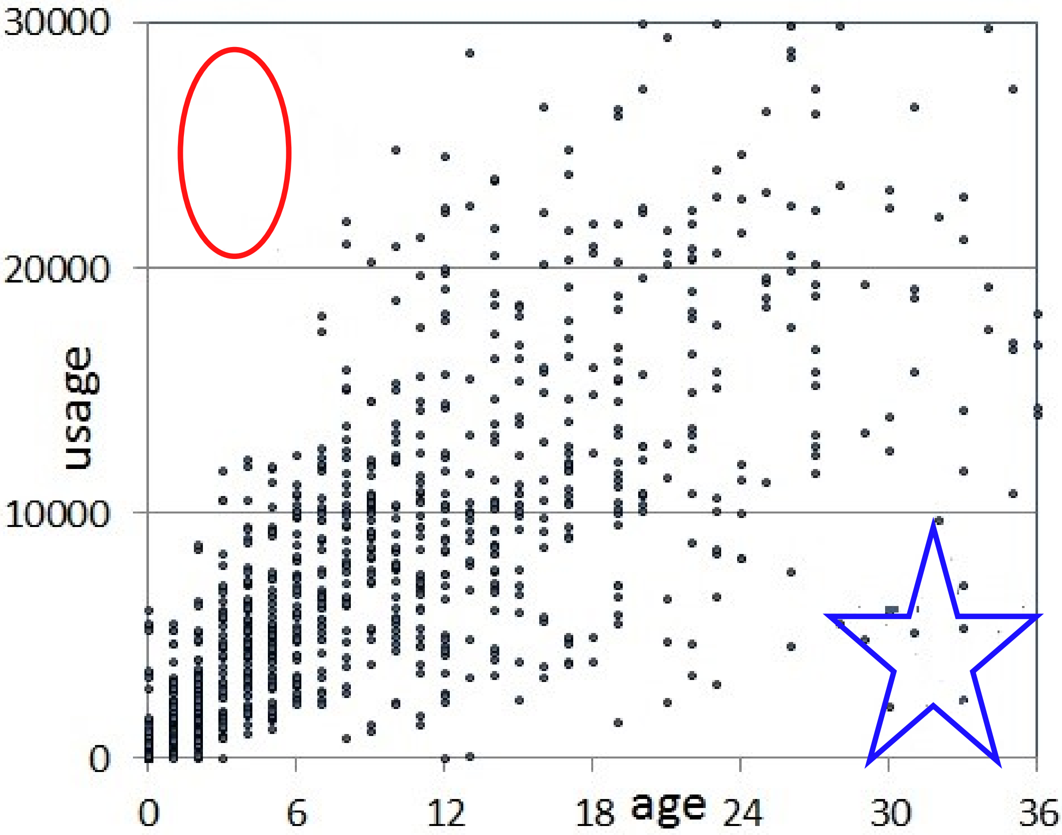


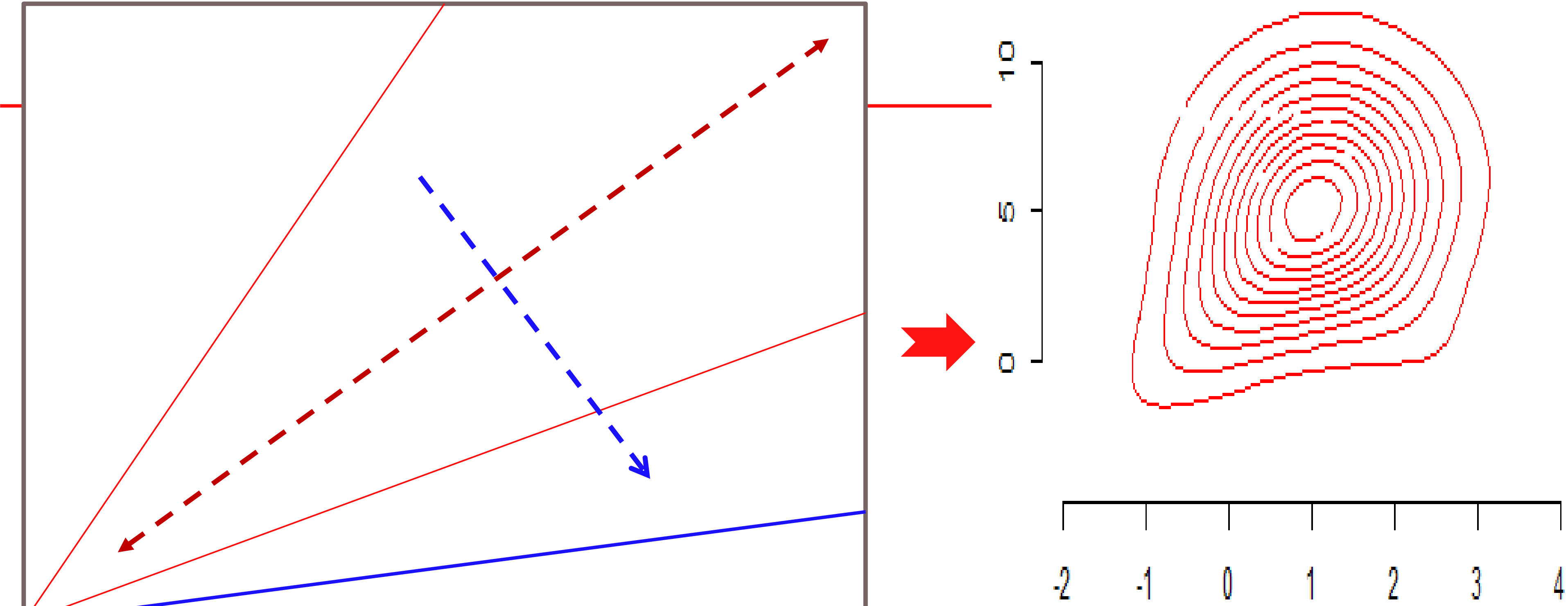
$$C(v_1, v_2) = \left(v_1^{-\theta} + v_2^{-\theta} - 1\right)^{-1/\theta}$$

$$\boldsymbol{C(v_1, v_2) = C(v_2, v_1)}$$

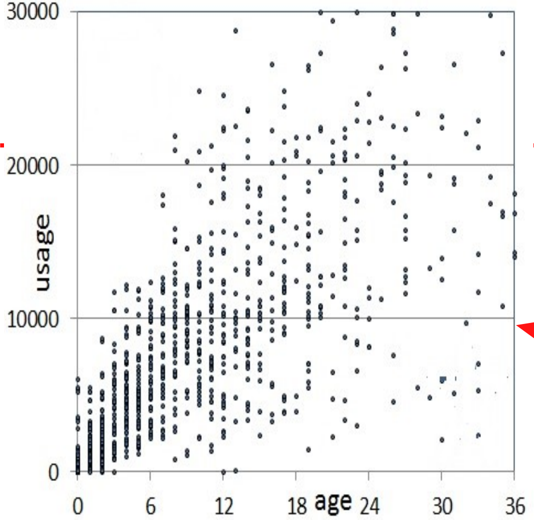
Symmetric

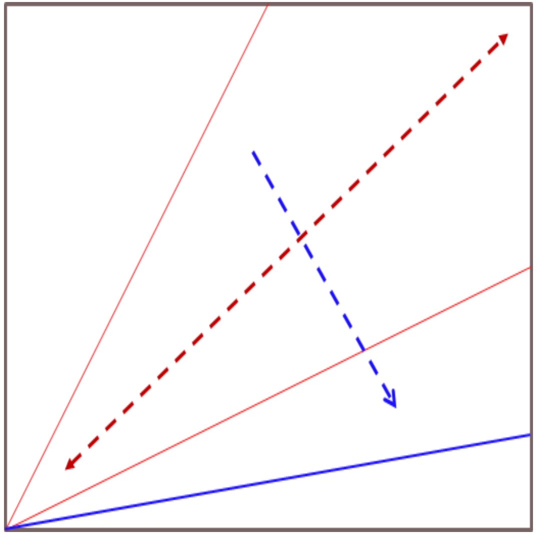
Warranty claim data: asymmetric phenomenon

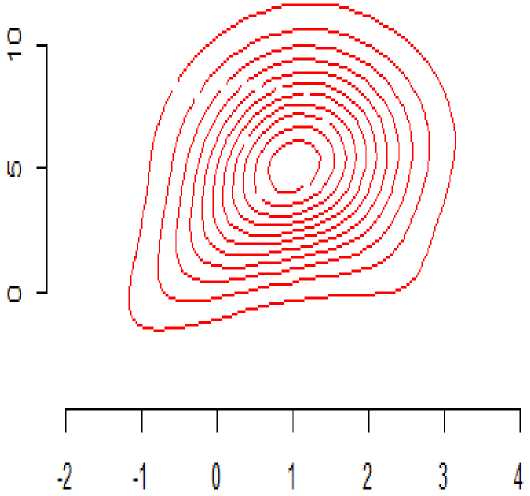




$$C(v_1, v_2) \neq C(v_2, v_1)$$

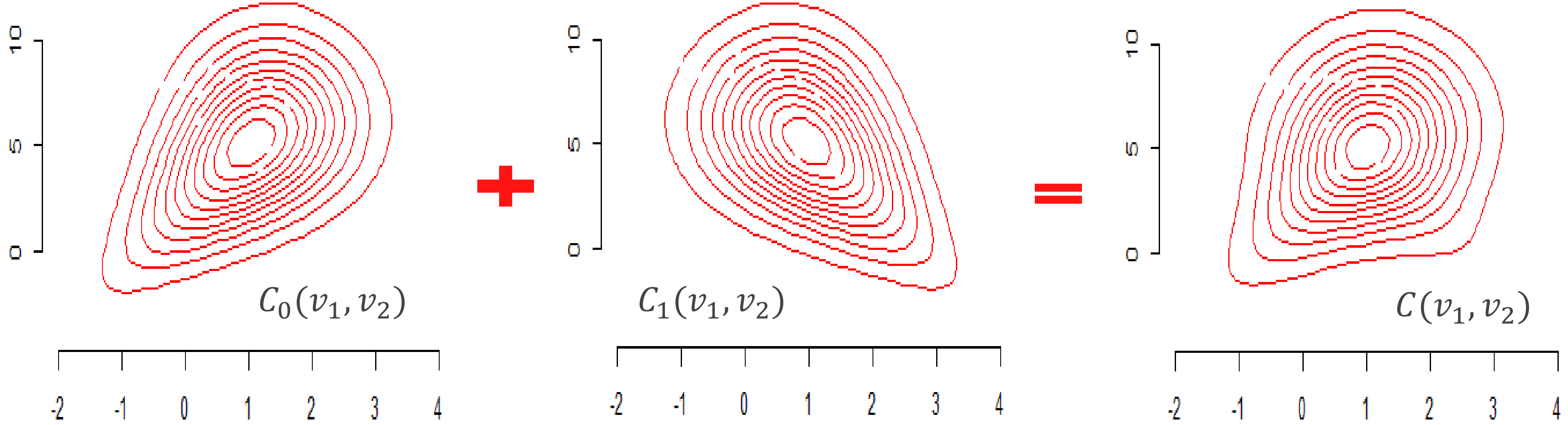






$$C(v_1, v_2) \neq C(v_2, v_1)$$

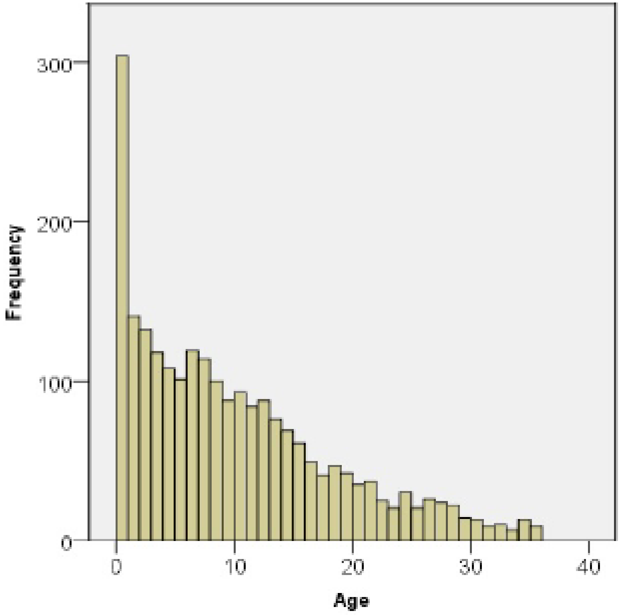
Construction of asymmetric copulas



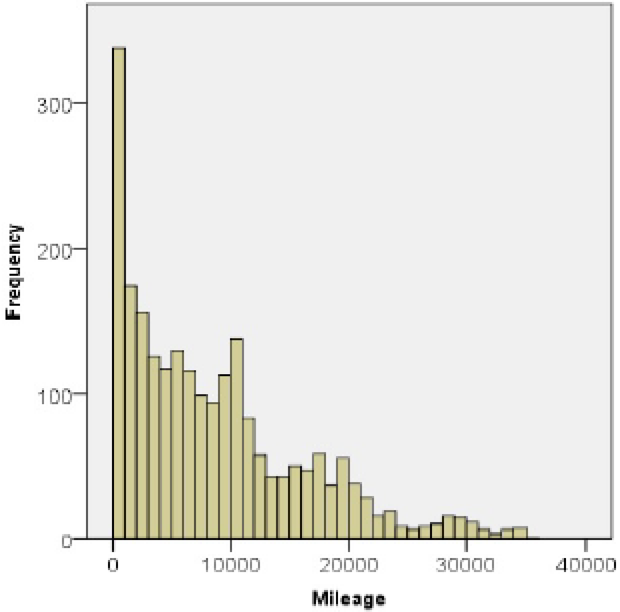
- $C_0(v_1, v_2) = \left(v_1^{-\theta} + v_2^{-\theta} - 1\right)^{-1/\theta}$;
- $C_1(v_1, v_2) = C_0(1, v_2) - C_0(1 - v_1, v_2) = v_2 - \left((1 - v_1)^{-\theta} + v_2^{-\theta} - 1\right)^{-1/\theta}$
- $C(v_1, v_2) = p_0 C_0(v_1, v_2) + p_1 C_1(v_1, v_2);\ p_0 = 0.7, p_1 = 0.3, \theta = 0.8$

Data and histogram

3,466 car warranty claims were collected from a car manufacturer. In those observations, the warranty of 2,289 cars were claimed within 36 months or 30,000 miles and the rest 1,177 cars were not claimed.



(a) Histogram of age



(b) Histogram of Usage (ie., Mileage)

Three models

- Let
  - $C(v_1, v_2) = v_1 + v_2 - 1 + C_0(1 - v_1, 1 - v_2)$ and
  - $\check{C}(v_1, v_2; \theta_2) = C(1, v_2; \theta_1) - C(1 - v_1, v_2; \theta_1)$

- Model 1: proposed model

$$C_1(v_1, v_2) = p_0\, C(v_1, v_2; \theta_1) + p_1 \check{C}(v_1, v_2; \theta_2)$$

- Model 2: a mixture of two Gumbel copulas with different parameters θ_1 & θ_2

$$C_2(v_1, v_2) = p_0\, C(v_1, v_2; \theta_1) + p_1 C(v_1, v_2; \theta_2)$$

- Model 3: the model proposed by Jung & Bai (2007)

$$C_2(v_1, v_2) = C(v_1, v_2; \theta_1)$$

Estimation of copulas

- Copulas can be estimated parametrically, semiparametrically or fully nonparametrically, such as maximum likelihood estimation (MLE), inferences function of margins (IFM), pseudo maximum likelihood estimation (PML)or Canonical maximum likelihood CML, method of moment using Kendall's tau and Spearman's rho, Nonparametric and Bayesian estimation

- Here we use maximum likelihood estimation

$$L_k(\boldsymbol{\theta}) = \sum_{i \in D} \log f_k(x_1, x_2) + \sum_{i \notin D} \log(1 - F_k(A_w, U_w))$$

Model performance on the original data

- AIC (Akaike information criterion): an estimator of prediction error and thereby relative quality of statistical models for a given set of data

$$\mathrm{AIC} = 2k - 2\ln \hat{L}$$

k is the number of parameters in a model

- A model with smaller AIC is favourable

Parameters and performance of the three models.

α_1	β_1	α_2	β_2	θ_1	θ_2	p_0	AIC	Methods
0.78	27.00	0.77	24496.98	3.76	0.48	0.90	30945.39	Model 1
0.73	33.38	0.65	37332.16	4.62	1.86	0.75	31061.48	Model 2
0.81	25.01	0.71	24965.02	3.01			31035.36	Model 3

Training datasets

Basic statistics of the age and the usage of the claimed cars.

Dataset	N	N_r	Age/Usage	Mean	Std. Deviation	Skewness	Kurtosis
D_1	500	324	Age	10.28	8.25	0.70	-0.31
			Usage	8890.70	7953.59	0.99	0.33
D_2	1500	981	Age	9.78	8.33	1.06	0.60
			Usage	8520.89	7475.36	1.13	0.86
D_3	3000	1939	Age	10.21	8.44	0.90	0.13
			Usage	8842.91	7628.02	1.05	0.66

Performance of the method

Wu, S., (2014). Construction of asymmetric copulas and its application in two-dimensional reliability modelling. *European Journal of Operational Research*, *238*(2), pp.476-485.

Table 4
Parameters and performance of the three models.

N	α_1	β_1	α_2	β_2	θ_1	θ_2	p_0	AIC	Methods
500	0.79	26.79	0.79	23656.14	3.77	0.46	0.89	**4469.18**	Model 1
	(0.032)	(2.72)	(0.041)	(2988.96)	(0.28)	(0.027)	(0.029)	(124.89)	
	0.78	26.39	0.74	24071.96.11	2.26	3.50	0.42	4471.50	Model 2
	(0.056)	(2.68)	(0.067)	(2618.457)	(1.53)	(2.10)	(0.27)	(123.53)	
	0.81	25.13	0.72	24268.27	2.88			4486.98	Model 3
	(0.033)	(2.45)	(0.043)	(2381.91)	(0.16)			(126.39)	
1500	0.77	26.68	0.76	24017.97	3.90	0.54	0.92	**13442.77**	Model 1
	(0.018)	(1.61)	(0.036)	(1346.64)	(0.30)	(0.19)	(0.033)	(198.83)	
	0.75	25.15	0.71	23237.66	3.65	2.74	0.58	13455.11	Model 2
	(0.024)	(1.62)	(0.025)	(1798.94)	(1.35)	(1.48)	(0.23)	(191.18)	
	0.79	24.16	0.71	23544.67	3.13			13483.00	Model 3
	(0.022)	(1.63)	(0.026)	(1678.95)	(0.083)			(190.87)	
3000	0.77	30.23	0.75	29308.12	3.78	0.77	0.87	**26919.03**	Model 1
	(0.021)	(3.26)	(0.052)	(5412.91)	(0.38)	(0.66)	(0.21)	(339.52)	
	0.75	30.38	0.69	31128.44	4.62	2.09	0.67	26953.39	Model 2
	(0.024)	(4.30)	(0.030)	(6455.89)	(0.89)	(0.59)	(0.14)	(327.83)	
	0.79	24.69	0.73	22882.41	3.01			26974.92	Model 3
	(0.017)	(1.67)	(0.016)	(2942.32)	(0.10)			(350.35)	

Warranty Policy Optimisation

Objective: to maximise profit or to minimize cost

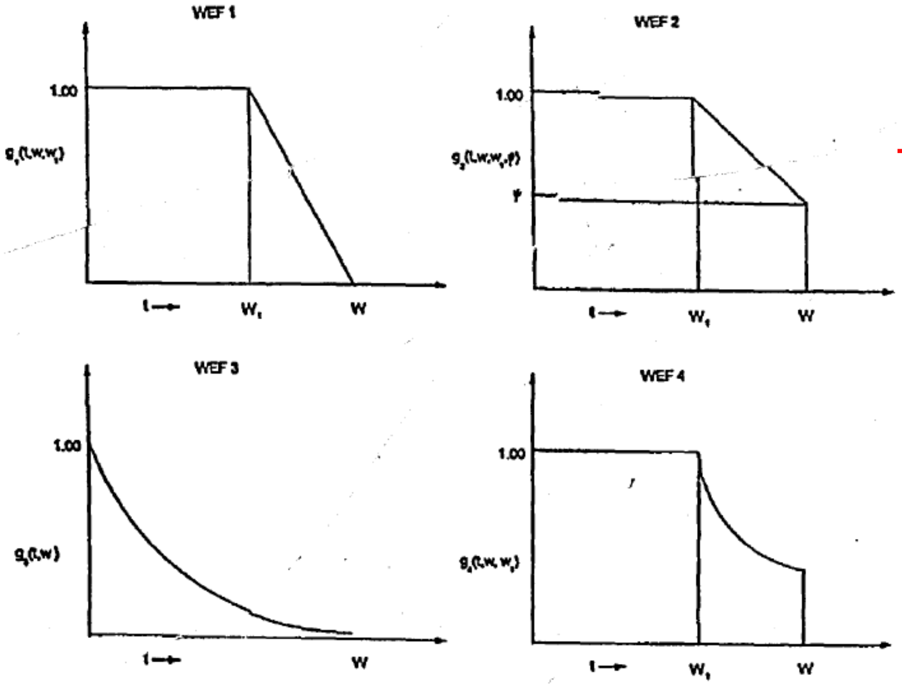
Optimisation variables:

- Warranty price
- Warranty length
- Maintenance policy

Warranty Policy Optimisation

- From a supply chain's perspective
  - Using game theory: comparing different retailers and different warranty policies
- From a reliability engineering perspective
  - To minimise warranty servicing cost, preventive maintenance can be conducted and optimally scheduled
- Most of existing research is done on the basis of
  - Assume that different subsystems (of a system) are independent
  - Individual products
- In an individual system: different subsystems
  - Hardware + software + user (human)
- In a manufacturer: a manufacturer may produce many products
  - Common components are installed in different components

Human factor: Failed but not reported (FBNR)



$$q_{11}(t,w_1,w,\varphi_1)=\begin{cases}1 & 0\le t\le w_1\\ \dfrac{w-\varphi_1 w_1}{w-w_1}-\dfrac{(1-\varphi_1)t}{w-w_1} & w_1\le t\le w,\\ 0 & t>w\end{cases}$$

where $0\le\varphi_1\le 1$ and $0\le w_1<w$, and

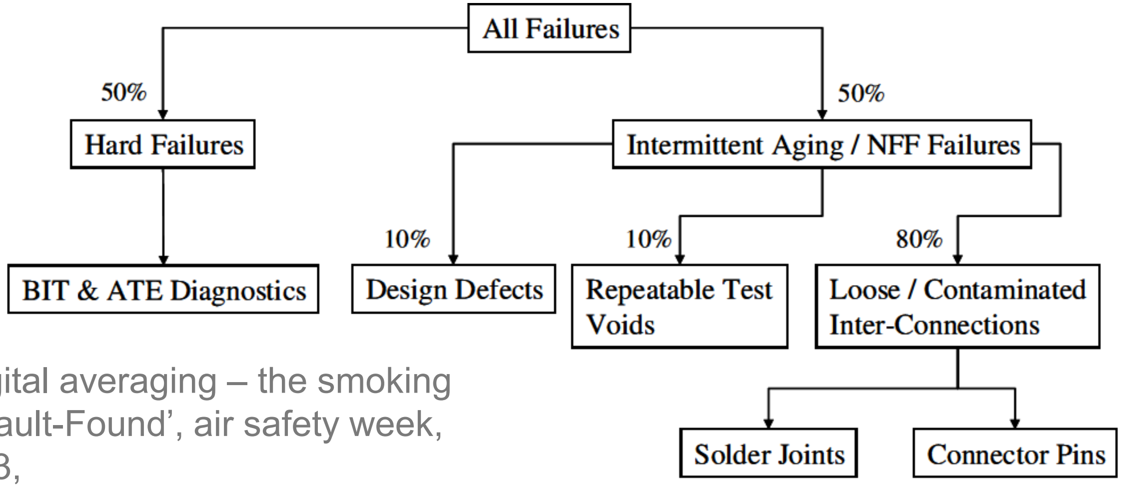
$$q_{12}(t,w_1,w,\varphi_2)=\begin{cases}1 & 0\le t\le w_1\\ e^{-(t-w)/\varphi_2} & w_1\le t\le w,\\ 0 & t>w\end{cases}$$

where $0\le w_1<w$.

*Patankar JG, Mitra A. Warranty and consumer behavior: Warranty execution. Product Warranty Handbook 1996:421-38.

Human factor--Reported but not failed (RBNF)

- Reported but not failed (RBNF)
  - Due to customers
  - Due to manufacturers
- Intermittent failures / NFF failures (No Fault Found)*



* Sorensen B. Digital averaging – the smoking gun behind 'No-Fault-Found', air safety week, February 24; 2003,

Ability to rectify the intermittent failures/NFF

- Manufacturer's ability to rectify RBNF(reported but not failed) "*failures*" is improving
  - with the number of claims*; or
  - with the time since the first RBNF**

$$WC_2(w) = \frac{1}{n}\sum_{k=1}^{\infty}\left([(k-1)c_2+\tilde{c}_2]P_{2k}\frac{e^{-\Lambda_n(w)}[\Lambda_n(w)]^k}{k!}\right)+\frac{1}{n}\sum_{k=1}^{\infty}\sum_{m=k+1}^{\infty}$$

$$\times\left([(m-1)c_2+\tilde{c}_2](1-p_{2m})\frac{e^{-\Lambda_n(w)}[\Lambda_n(w)]^k}{k!}\prod_{j=1}^{m-1}p_{2j}\right) \qquad (2)$$

*Wu, S., Warranty claim analysis considering human factors (2011), Reliability Engineering and System Safety, 96 (1), pp. 131-138
**Wu, S. (2014) Warranty return policies for products with unknown claim causes and their optimisation. International Journal of Production Economics,156. pp. 52-61

Severity of warranty claims

- The total cost of warranty claims of product k is

$$S_k(t) = \sum_{i=1}^{N_k(t)} X_{k,i}$$

where $X_{k,i}$ is the severity of the i-th claim of product k; $N_k(t)$ is the total number of claims

- The relationship among the sales amount M_k, warranty length T_k, and warranty price P_k is assumed

$$M_k = A_k - \beta_k P_k + \eta_k T_k,$$

where A_k, β_k, η_k are positive parameters

Total profit

- The profit of product k, $\omega_k(P_k, T_k)$, is given by

$$\omega_k(P_k, T_k) = M_k[P_k - S_k(T_k) - c_k],$$

where c_k is the fix cost of product k

- Then the total profit for n products in the manufacturer is given by

$$\Omega(\boldsymbol{P}, \boldsymbol{T}) = \sum_{k=1}^{n} (A_k - \beta_k P_k + \eta_k T_k)\omega_k(P_k, T_k)$$

Options– a mean-variance approach

- **Option 1**. to maximise a combination of the profit and the risk of the estimated profit;
  - To maximise $\mathrm{E}[\Omega(\boldsymbol{P},\boldsymbol{T})]-\sqrt{\mathrm{variance}[\Omega(\boldsymbol{P},\boldsymbol{T})]}$
- **Option 2**. to maximise the profit and meanwhile to limit the risk of the estimated profit;
  - To maximise $\mathrm{E}[\Omega(\boldsymbol{P},\boldsymbol{T})]$, subject to $\sqrt{\mathrm{variance}[\Omega(\boldsymbol{P},\boldsymbol{T})]}<\phi_0$
- **Option 3**. to minimise the risk of the estimated profit subject to the constraint that the lower bound of the profit is greater than a pre-specified value
  - To minimise$\sqrt{\mathrm{variance}[\Omega(\boldsymbol{P},\boldsymbol{T})]}$, subject to$\mathrm{E}[\Omega(\boldsymbol{P},\boldsymbol{T})]<\phi_1$

Luo, M., & Wu, S. (2018). A mean-variance optimisation approach to collectively pricing warranty policies. *International Journal of Production Economics*, 196, 101-112.

Hardware, software, users

- Value-at-Risk approach
- Failure of a software subsystem may have two implications:

1) the software needs repairing and installing in its host system; it needs installing/updating in all of the other items of the same product; and
2) the failure of its host hardware system needs repairing, which may have impact on one individual hardware system.

Luo, M., & Wu, S. (2018). A value-at-risk approach to optimisation of warranty policy. *European Journal of Operational Research*, 267(2), 513-522.
Luo, M., & Wu, S. (2019). A comprehensive analysis of warranty claims and optimal policies. *European Journal of Operational Research*, 276(1), 144-159.

Conclusion and future research

- Collecting warranty data with good quality is a challenge;
- Copulas can be applied to warranty data analysis and policy optimisation
- Collectively optimising warranty policies for several products can mitigate risk

- Future research
  - Sensors are installed to monitor the behaviour of items in a system, more data are therefore collected. More sophisticated data analysis methods should be developed for warranty data analysis
  - A product item is normally composed of many components. The reliability and failure process should be properly studied

References

- Blischke, W. R., Karim, M. R., & Murthy, D. P. (2011). *Warranty data collection and analysis*. Springer Science & Business Media.
- Jung, M., & Bai, D. S. (2007). Analysis of field data under two-dimensional warranty. *Reliability Engineering & System Safety*, 92(2), 135-143.
- Luo, M., & Wu, S. (2018). A mean-variance optimisation approach to collectively pricing warranty policies. *International Journal of Production Economics*, 196, 101-112.
- Luo, M., & Wu, S. (2018). A value-at-risk approach to optimisation of warranty policy. *European Journal of Operational Research*, 267(2), 513-522.
- Luo, M., & Wu, S. (2019). A comprehensive analysis of warranty claims and optimal policies. *European Journal of Operational Research*, 276(1), 144-159.
- Sklar, A. (1959), “Fonctions de r´epartition `a n dimensions et leurs marges,” *Publ. Inst. Statist. Univ*. Paris 8, 229-231
- Wu, S. (2011). Warranty claim analysis considering human factors. *Reliability Engineering & System Safety*, 96(1), 131-138.
- Wu, S. (2012). Warranty data analysis: A review. *Quality and Reliability Engineering International*, 28(8), 795-805.
- Wu, S. (2013). A review on coarse warranty data and analysis. *Reliability Engineering & System Safety*, 114, 1-11.
- Wu, S., (2014). Construction of asymmetric copulas and its application in two-dimensional reliability modelling. *European Journal of Operational Research*, 238(2), pp.476-485.
- Wu, S. (2014) Warranty return policies for products with unknown claim causes and their optimisation. *International Journal of Production Economics*,156. pp. 52-61

I would like to thank the organiser of *the 1st International Conference on EMMA-2021 (Engineering, Medicine, Management, Arts and Sciences)* for their invitation

Questions?