A copula-based parametric maximum likelihood estimation for dependently left-truncated data

Presenter: Chi-Hung Pan

Advisor: Takeshi Emura

Graduate Institute of Statistics, National Central University

Abstract: Traditional statistical methods for left-truncated lifetime data rely on the independence assumption regarding the truncation variable. However, the dependence between a lifetime variable X of interest and its left-truncation variable L usually occurs in many real data from reliability and biomedical analysis. In this paper, we propose a copula-based dependence model between L and X with the marginal distributions specified by parametric models. Then we consider the maximum likelihood estimator (MLE) under the copula-based dependence model. To calculate the MLE, explicit formulas of the inclusion probability $c(\boldsymbol{\theta}) = \Pr(L \le X)$ and its partial derivatives are obtained under the Clayton copula and Weibull marginal model, which are new results in this paper. Then we derive explicit expression for the randomized-Newton-Raphson algorithm for maximizing the log-likelihood. We perform simulations to verify the correctness of the proposed. We illustrate our method by real data from a field reliability study on the lifetimes of brake pads.

1. Introduction

Figure 1 shows an example of left-truncated data that appears in a field reliability study on the lifetimes of brake pads of automobiles (Kalbfleisch and Lawless 1992). The brake pads have a measure of failure time, which is the number of kilometers driven before the pads are broken.

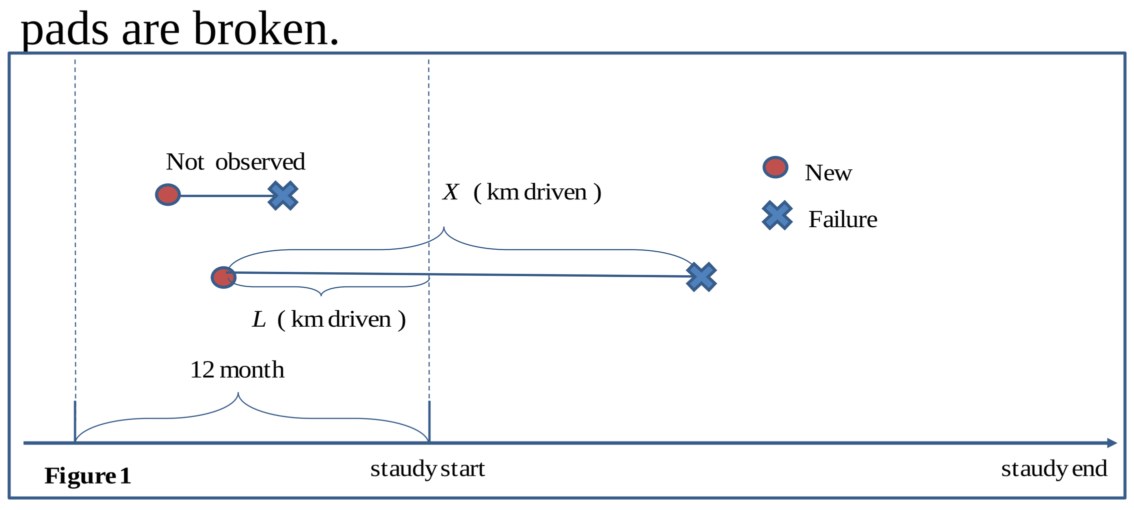


Figure 1

Traditionally, most literature on the truncated data consider statistical estimation by assuming that L and X are independent. However, the dependence usually occurs in many real data.

2. Copula models

A copula is a bivariate function $C_\alpha : [0,1]^2 \mapsto [0,1]$ with uniform margins, where α is a parameter control the degree of dependence. Let $F_L(l;\boldsymbol{\theta}_L)$ and $F_X(l;\boldsymbol{\theta}_X)$ be the marginal distributions of L and X, where $\boldsymbol{\theta}_L$ and $\boldsymbol{\theta}_X$ is a vector of parameters. We propose to model the dependence between L and X by a copula function. Consider the bivariate copula model defined as

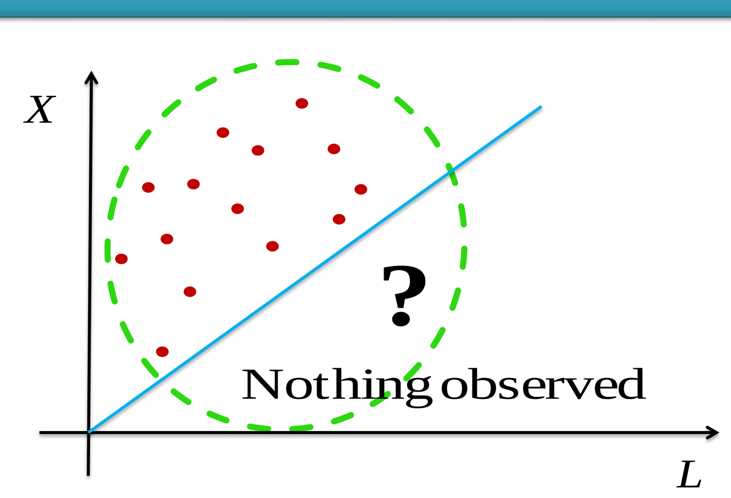
$$\Pr_{\boldsymbol{\theta}}(L \le l, X \le x) = C_\alpha[F_L(l;\boldsymbol{\theta}_L), F_X(x;\boldsymbol{\theta}_X)],$$

Then we can drive the density function of as

$$f_{L,X}(l,x;\boldsymbol{\theta}) = C_\alpha^{[1,1]}[F_L(l;\boldsymbol{\theta}_L), F_X(x;\boldsymbol{\theta}_X)] f_L(l;\boldsymbol{\theta}_L) f_X(x;\boldsymbol{\theta}_X),$$

where $f_L(l;\boldsymbol{\theta}_L)$ and $f_X(l;\boldsymbol{\theta}_X)$ are marginal density of L and X, $C_\alpha^{[1,1]}[U,V] = \partial^2 C_\alpha[U,V]/\partial U \partial V$.

3. Likelihood construction



Given the observed data $\{(L_j, X_j);\ j = 1,2,\ldots,n\}$ subject to $L_j \le X_j$, the likelihood function has the form

$$L_n(\boldsymbol{\theta}) = c(\boldsymbol{\theta})^{-n} \prod_j f_{L,X}(L_j, X_j;\boldsymbol{\theta})$$

where

$$c(\boldsymbol{\theta}) = \Pr(L \le X) = \iint_{l \le x} f_{L,X}(l,x;\boldsymbol{\theta})\,dx\,dl,$$

(inclusion probability).

We hope that there is a simple form of $c(\boldsymbol{\theta})$. Then it is more easy to perform likelihood inference.

3. Theory

Theorem: *Assume that the inverse functions of* $F_L(l;\boldsymbol{\theta}_L)$ *and* $F_X(l;\boldsymbol{\theta}_X)$ *exist. Then*

$$c(\boldsymbol{\theta}) = \Pr(L \le X) = \int_0^1 H(u;\boldsymbol{\theta})\,du,$$

where

$$H(u;\boldsymbol{\theta}) = h_\alpha[F_L\{F_X^{-1}(u;\boldsymbol{\theta}_X);\boldsymbol{\theta}_L\}, u],$$

and where

$$h_\alpha(u_1,u_2) = \Pr(U_1 \le u_1 \mid U_2 = u_2) = \frac{\partial C_\alpha(u_1,u_2)}{\partial u_2},$$

is called h-function. (Schepsmeier and Stöber 2014)

Lemma: *[p.301 of Khuri (2003)]*

Let a_i *and* b_i *be real numbers with* $a_i < b_i,\ i = 1,2,\ldots,p$. *Let* $H : D \to R$, *where* $D = \{(u,\theta_1,\theta_2,\ldots,\theta_p) \mid 0 \le u \le 1, a_i \le \theta_i \le b_i, \text{for } \forall i = 1,\ldots,p\}$. *For fixed* $\theta_j, j \ne i$ *let* $D_i = \{(u,\theta_i) \mid 0 \le u \le 1, a_i \le \theta_i \le b_i, a_i, b_i \in R\}$ *If* H *and* $\partial H/\partial\theta_i$ *are continuous in* D_i, *then*

$$\frac{\partial}{\partial\theta_i}\int_0^1 H(u;\boldsymbol{\theta})\,du = \int_0^1 \frac{\partial H(u;\boldsymbol{\theta})}{\partial\theta_i}du,\ i = 1,2,\ldots,p.$$

4. Full parametric model

We consider the Clayton copula

$$C_\alpha(u_1,u_2) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-\frac{1}{\alpha}}, \qquad \alpha \ge 0.$$

Assume $F_L(l;\lambda_L,\nu_L) = 1 - \exp(-\lambda_L l^{\nu_L})$ and $F_X(x;\lambda_X,\nu_X) = 1 - \exp(-\lambda_X x^{\nu_X})$ (Weibull lifetime). Then the inclusion probability is

$$c(\boldsymbol{\theta}) = \Pr(L \le X) = \int_0^1 H(u;\boldsymbol{\theta})du,$$

where $\boldsymbol{\theta}' = (\alpha,\lambda_L,\lambda_X,\nu_L,\nu_X)$,

$$H(u;\boldsymbol{\theta}) = u^{-\alpha-1}B(u,\boldsymbol{\theta})^{-\frac{1}{\alpha}-1},$$

$$B(u;\boldsymbol{\theta}) = (1 - \exp[-\lambda_L\{-\log(1-u)/\lambda_X\}^{\frac{\nu_L}{\nu_X}}])^{-\alpha} + u^{-\alpha} - 1.$$

5. "Randomize" Newton-Raphson

It is well-known that Newton-Raphson algorithm is sensitive to the initial values. To stabilize the algorithm, we propose the randomized algorithm for Newton-Raphson algorithm (Hu and Emura 2015).

Step 1. Choose initial value $\boldsymbol{\theta}^{(0)} = \{\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_p^{(0)}\}'$.

Step 2. Repeat the following iterations,

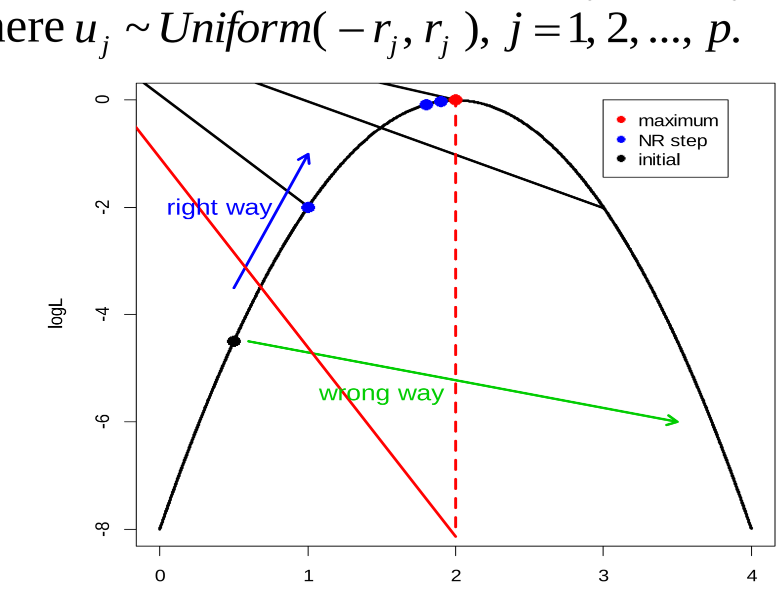
$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - H^{-1}\{\boldsymbol{\theta}^{(k)}\}S\{\boldsymbol{\theta}^{(k)}\},$$

where

$$H\{\boldsymbol{\theta}\} = \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}}\ell_n(\boldsymbol{\theta}) \text{ (Hessian matrix)},$$

$$S\{\boldsymbol{\theta}\} = \frac{\partial}{\partial\boldsymbol{\theta}}\ell_n(\boldsymbol{\theta}) \text{ (Score vector)}.$$

- If $\max|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}| < \varepsilon$, then and hessian matrix is negative define, then stop. $\boldsymbol{\theta}^{(k+1)}$ is the MLE.
- If $\max|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}| > b$ replace $\theta_j^{(k+2)} = \theta_j^{(0)} \times \exp(u_j)$, where $u_j \sim Uniform(-r_j, r_j),\ j = 1,2,\ldots,p$.



6. Simulations

Simulation results under the Clayton copula with Weibull margins based on 1000 repetitions. $\boldsymbol{\theta} = (\alpha,\lambda_L,\lambda_X,\nu_L,\nu_X)$.

		$c(\boldsymbol{\theta})$	n	$E(\hat\alpha)$	$E(\hat\lambda_L)$	$E(\hat\lambda_X)$	$E(\hat\nu_L)$	$E(\hat\nu_X)$	AI
$\alpha=1$ $(\tau=0.33)$	$\lambda_L=2,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.736	100	1.042	2.035	1.031	1.010	1.005	6.6
			200	1.018	2.017	1.017	1.006	1.003	6.4
			300	1.012	2.009	1.014	1.002	1.001	6.4
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.500	100	1.078	1.012	1.049	1.013	1.011	8.8
			200	1.046	1.014	1.030	1.003	1.001	7.8
			300	1.030	1.009	1.023	1.002	1.001	7.8
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=2,\nu_X=1$	0.416	100	1.037	0.976	1.101	2.116	1.038	32.7
			200	1.012	0.983	1.044	2.057	1.019	30.9
			300	1.005	0.987	1.020	2.036	1.018	27.1
$\alpha=2$ $(\tau=0.5)$	$\lambda_L=2,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.804	100	2.065	2.040	1.031	1.007	1.003	6.8
			200	2.037	2.017	1.009	1.003	1.004	6.1
			300	2.029	2.012	1.008	1.001	1.002	6.1
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.500	100	2.170	1.025	1.127	0.993	0.980	65.1
			200	2.086	1.014	1.052	0.993	0.990	8.7
			300	2.057	1.009	1.032	0.996	0.995	8.2
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=2,\nu_X=1$	0.387	100	2.075	0.972	1.164	2.119	1.026	69.4
			200	2.019	0.986	1.051	2.044	1.018	53.6
			300	2.003	0.986	1.019	2.036	1.019	50.0
$\alpha=6$ $(\tau=0.75)$	$\lambda_L=2,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.945	100	6.105	2.041	1.028	1.010	1.008	58.6
			200	6.073	2.014	1.020	1.003	1.004	41.2
			300	6.063	2.011	1.006	1.002	1.003	39.0

		$c(\boldsymbol{\theta})$	n	$MSE(\hat\alpha)$	$MSE(\hat\lambda_L)$	$MSE(\hat\lambda_X)$	$MSE(\hat\nu_L)$	$MSE(\hat\nu_L)$
$\alpha=1$ $(\tau=0.33)$	$\lambda_L=2,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.736	100	0.1045	0.1111	0.0387	0.0078	0.0141
			200	0.0551	0.0494	0.0186	0.0035	0.0067
			300	0.0351	0.0319	0.0104	0.0022	0.0044
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.500	100	0.1470	0.0559	0.1198	0.0087	0.0260
			200	0.0688	0.0277	0.0505	0.0032	0.0138
			300	0.0423	0.0176	0.0338	0.0022	0.0096
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=2,\nu_X=1$	0.416	100	0.4369	0.0522	0.3778	0.1298	0.0676
			200	0.1743	0.0235	0.1305	0.0574	0.0333
			300	0.1083	0.0143	0.0715	0.0317	0.0206
$\alpha=2$ $(\tau=0.5)$	$\lambda_L=2,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.804	100	0.1874	0.0796	0.0306	0.0064	0.0084
			200	0.0941	0.0360	0.0084	0.0029	0.0036
			300	0.0634	0.0253	0.0053	0.0020	0.0023
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.500	100	0.3593	0.0450	0.2509	0.0134	0.0247
			200	0.1281	0.0200	0.0631	0.0042	0.0113
			300	0.0752	0.0129	0.0319	0.0027	0.0072
	$\lambda_L=1,\lambda_X=1,$ $\nu_L=2,\nu_X=1$	0.387	100	1.0228	0.0594	0.6311	0.1830	0.0730
			200	0.3653	0.0218	0.1421	0.0552	0.0301
			300	0.2245	0.0140	0.0874	0.0363	0.0209
$\alpha=6$ $(\tau=0.75)$	$\lambda_L=2,\lambda_X=1,$ $\nu_L=1,\nu_X=1$	0.945	100	0.9622	0.0508	0.0846	0.0064	0.0067
			200	0.5022	0.0216	0.0694	0.0028	0.0039
			300	0.2841	0.0137	0.0033	0.0018	0.0018

7. Data Analysis

We analyze the lifetimes of brake pads data (Kalbfleisch and Lawless 1992). Data size is 98. We consider two models which are the Clayton copula with exponential margins and Weibull margins. To compare the two models, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are adopted for model selection. The preferred model is the one with minimum AIC (BIC) value, given by

$$AIC = -2\log L + 2k(BIC = -2\log L + k\log n).$$

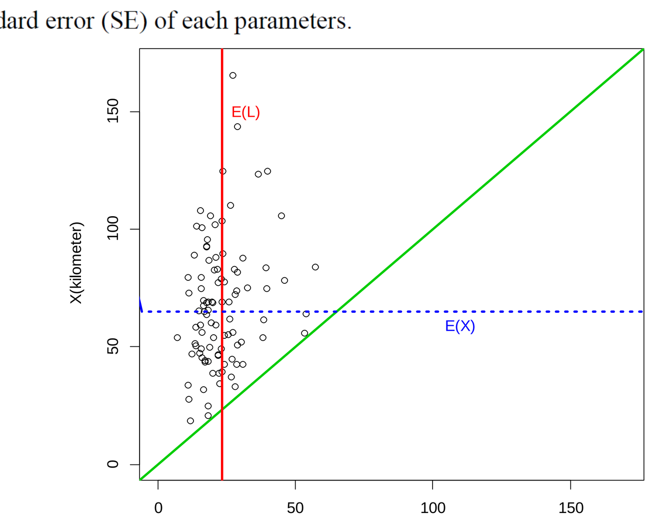
Results are summary in the following table.

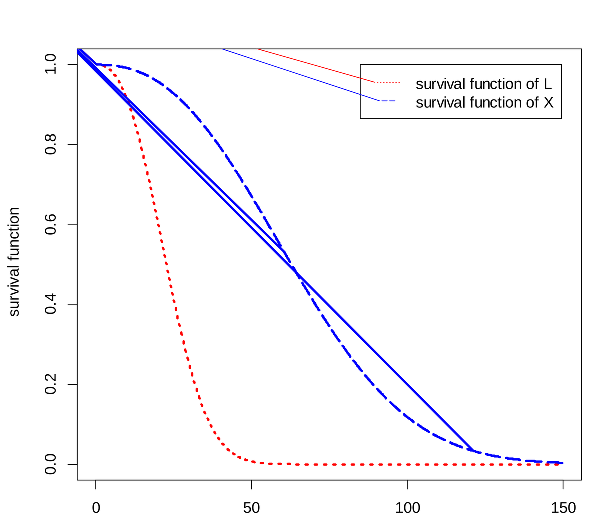
Model	$\hat E(X)$	$\hat\alpha$	$\hat\lambda_L$	$\hat\lambda_X$	$\hat\nu_L$	$\hat\nu_X$	$\log L$	AIC	BIC
M_1	48.31	0.9242	0.0364	0.0207	×	×	-874.31	1754.61	1762.36
	(5.4850)	(0.6634)	(0.0081)	(0.0024)					
M_2	64.85	0.3321	2.89×10^{-4}	3.09×10^{-5}	2.4927	2.4193	-806.62	1623.24	1636.17
	(3.2457)	(0.3580)	(1.86×10^{-4})	(3.15×10^{-5})	(0.1829)	(0.2222)			

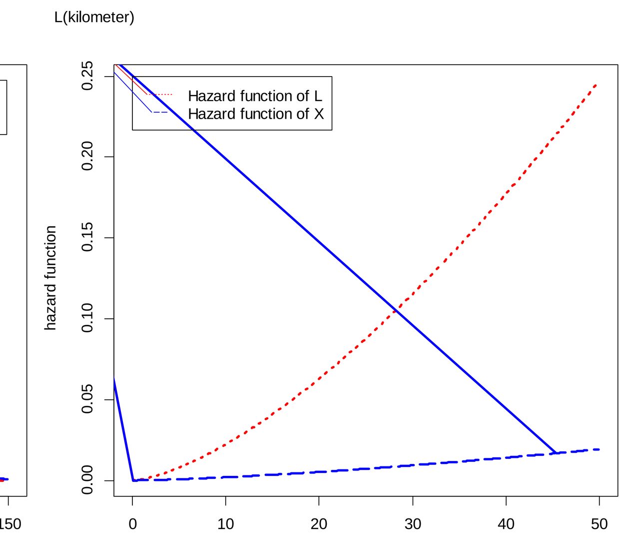
M_1: The Clayton copula with exponential margins.

M_2: The Clayton copula with Weibull margins.

$(\cdot)$: The standard error (SE) of each parameters.







8. Conclusion

- New results for the formulas of the inclusion probability $c(\boldsymbol{\theta})$ is obtained under the copula model.
- we propose the randomized-Newton-Raphson algorithm for maximizing the log-likelihood.
- we found that there is weak positive dependence between two variables. This implies that the results of Kalbfleisch and Lawless (1992) that assumed independence are questionable.

9. Reference

1. Kalbfleisch JD, Lawless JF (1992) Some Useful Statistical Methods for Truncated Data. *Journal of quality technology* 24: 145-152.
2. Schepsmeier U, *Stöber* J (2014) Derivatives and fisher information of bivariate copulas. *Statistical Papers* 55:525-542.
3. Hu YH, Emura T (2015) Maximum likelihood estimation for a special exponential family under random double-truncation. *Computational Statistics* DOI: 10.1007/s00180-015-0564-z.
4. Khuri AI (2003) *Advanced Calculus with Applications in Statistics* (2nd ed.). John Wiley & Sons, Inc., Hoboken, New Jersey.