Parametric survival analysis with dependent truncation, a copula-based approach

Takeshi Emura

Joint work with Chi-Hung Pan

Graduate Institute of Statistics
National Central University

第二十七屆南區統計研討會
成功大學統計學系
June 29-30, 2018

Outline

- Review
- Estimation (MLE)
- Main theorem (Theorem 1)
- Asymptotic
- Goodness-of-fit (Cramér-von-Mises test)
- R package (*depend.truncation*)
- Simulations
- Data analysis
- Extension (Theorem 3)

Reference:

Emura T, Pan CH (2018-), Parametric maximum likelihood inference and goodness-of-fit tests for dependently left-truncated data, a copula-based approach,

to appear in *Statistical Papers*

Left-truncation:

- (L, X): a pair of random variables
- L: left-truncation time
- X: failure time

If $L \le X$, the sample is available

If $L > X$, nothing is available !

Samples:

$\{(L_j, X_j);\ j = 1, 2, \ldots, n\}$ subject to $L_j \le X_j$

Target of Estimation:

$S_X(x) = P(X > x)$, $E(X) = \int x dF_X(x)$, etc.

Examples of left-truncation

- Channing house data for elderly residents (Hyde, 1980)

  $L =$ Age at entry, $X =$ **Age at death**

- Car brake pads data (Kalbfleisch and Lawless,1992)

  $L =$ kilometers driven at study start

  $X =$ **Kilometers driven at failure**

- Unemployment data for women in Spain (De Uña-álvarez, 2004)

  $L =$ Time at inquiry, $X =$ **Time to finding a job**

- Twin-City study for dementia in France (Rondeau et al. 2015)

  $L =$ Age at entry, $X =$ **Age at diagnosis of dementia**

Target on the distribution of X

Estimation with left-truncated survival data

-Lynden-Bell (1971) **:** Product-limit estimator

$$\hat{S}_X(x) = \prod_{u \le x} \left\{ 1 - \frac{\sum_j \mathbf{I}(L_j \le u, X_j = u)}{\sum_j \mathbf{I}(L_j \le u \le X_j)} \right\}$$

(Woodroofe 1985; Wang et al. 1986; Zhou & Yip, 1999)

-Kalbfleisch and Lawless (1992): Parametric MLE

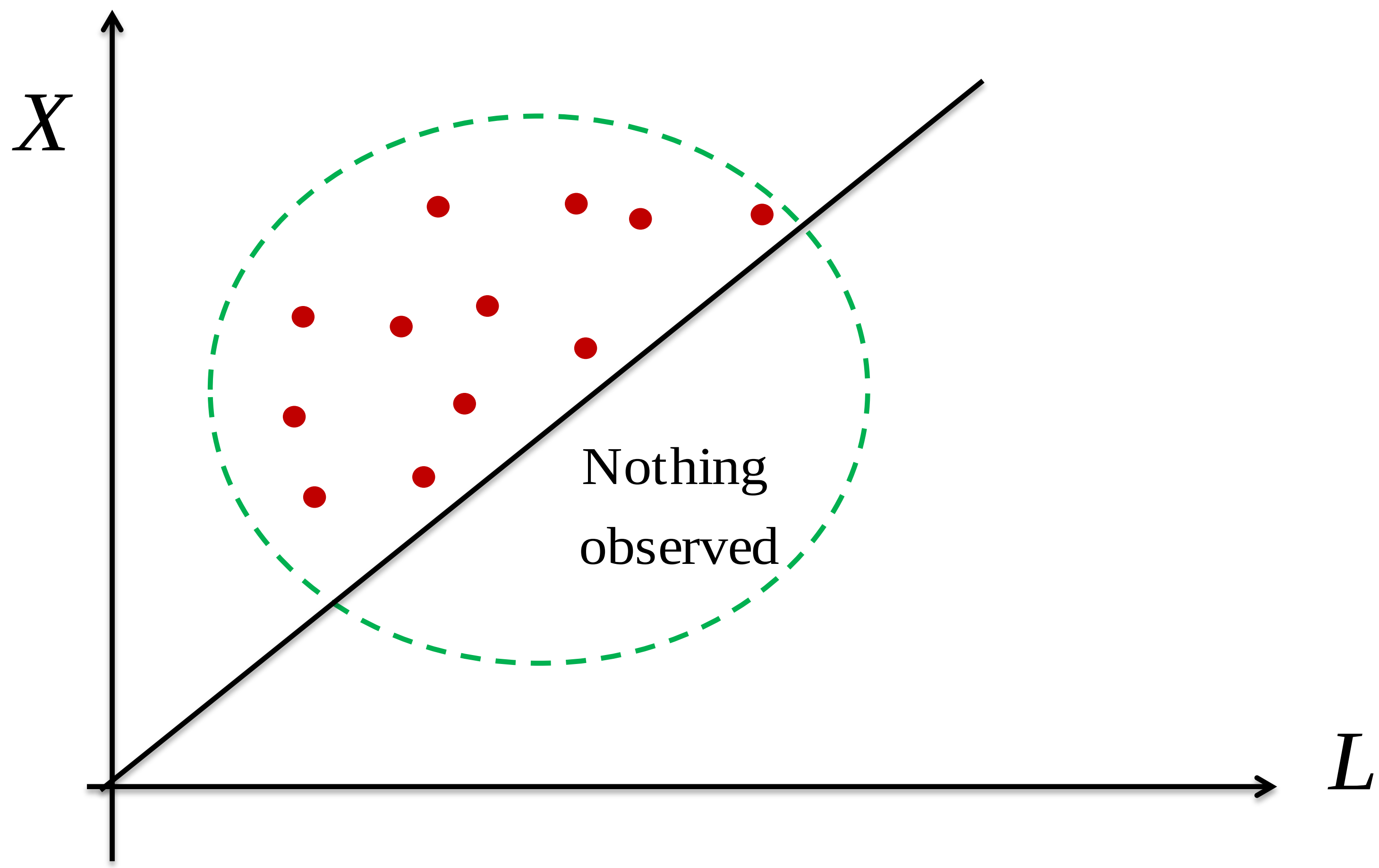
$$\hat{S}_X(x) = 1 - \Phi[\{\log(x) - \hat{\mu}\} / \hat{\sigma}],$$

$$(\hat{\mu}, \hat{\sigma}) = \arg\max_{(\mu, \sigma)} \prod_{i=1}^{n} \left[\frac{f_X(X_i; \mu, \sigma)}{P(L_i \le X; \mu, \sigma)} \right]$$

These estimators are consistent under $L \perp X$:
, **independent truncation assumption**

- Independence assumption $L \perp X$ is testable by left-truncated data (Tsai, 1990)

$$\{ (L_j, X_j); j = 1, 2, \ldots, n \} \text{ subject to } L_j \le X_j$$



Quasi-independence assumption (Tsai, 1990)

$$H_0: \ \Pr(X = x, L = l \mid X \leq Y) \propto dF_X(x)dF_L(l)$$

Available test statistics:

1. **Chen** et al. (1996) - test with the conditional Pearson-correlation
2. Tsai (1990); Martin & Betensky (2005)
   - test with the conditional Kendall's tau
3. Emura & Wang (2010), **Shen** (2011)
   - weighted-logrank test ; - score test under copulas
4. De Uña-Álvarez (2012); Rodríguez Girondo & de Uña-Álvarez (2012)
   - test with the Markov condition
5. Strzalkowska-Kominiak & Stute (2013)
   - test with Kendall's tau or Spearman's rho

- Quasi-independence is rejected in real examples
- Results not robust against dependent truncation (Bakoyannis & Touloumi, 2015)

Methods for dependent truncation

1. Semiparametric model (marginal unspecified)

Chaieb et al.(2006), Emura et al. (2011), Emura and Murotani (2015)

- Estimation under Archimedean copulas

Beaudoin & Lakhal-Chaieb (2008),

- Goodness-of-fit for Archimedean copulas

Strzalkowska-Kominiak & Stute (2013)

- Estimation under general copulas

Emura & Wang (2012)

- Estimation & model selection under general copulas

2. Parametric model

Emura and Konno (2012a) - Bivariate normal model

Emura and Konno (2012b) - Bivariate Poisson, Bernoulli-Poisson

Parametric models are still very limited

$\boldsymbol{\theta} = (\alpha, \boldsymbol{\theta}_L, \boldsymbol{\theta}_X) \in \Theta$

Proposed parametric models

- Joint distribution:

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

- Copula:

$$C_\alpha : [0,1]^2 \mapsto [0,1],$$

$$\alpha \in R: \text{ dependence parameter}$$

- Continuous parametric margins

$$F_L(l; \boldsymbol{\theta}_L) = P_{\boldsymbol{\theta}_L}(L \le l), \quad F_X(x; \boldsymbol{\theta}_X) = P_{\boldsymbol{\theta}_X}(X \le x)$$

- Unknown parameters:

$$\boldsymbol{\theta} = (\alpha, \boldsymbol{\theta}_L, \boldsymbol{\theta}_X) \in \Theta$$

Likelihood under dependent left-truncation

- Joint distribution:

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

- Joint density:

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

$C_\alpha^{[1,1]} = \partial^2 C_\alpha / \partial u \partial v$: copula density

$f_L = dF_L / dl$ and $f_X = dF_X / dx$: marginal densities

- Likelihood under left-truncation criterion: $L_j \le X_j$

$$L_n(\boldsymbol{\theta}) = \prod_{j=1}^{n} \frac{f_{L,X}(L_j, X_j;\boldsymbol{\theta})}{P_{\boldsymbol{\theta}}(L \le X)}$$

Need the form of $c(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(L \le X) = \iint_{l \le x} f_{L,X}(l, x;\boldsymbol{\theta}) dx dl$

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

- Marshal and Olkin (1967) bivarate exponential

$$\Pr(L > l, X > x) = \exp\{-\lambda_L l - \lambda_X x - \lambda_{LX} \max(l, x)\} \quad (l \ge 0, x \ge 0)$$

$$\Rightarrow \quad c(\boldsymbol{\theta}) = (\lambda_X + \lambda_{LX})(\lambda_L + \lambda_X + \lambda_{LX})^{-1}$$

- Bivariate normal, bivariate *t*

$$c(\boldsymbol{\theta}) = \Psi\left(\frac{\mu_X - \mu_L}{\sqrt{\sigma_X^2 + \sigma_L^2 - 2\sigma_{LX}}} ; v \right)$$

- FGM copula with Burr III margin

$$c(\boldsymbol{\theta}) = \frac{\gamma}{\gamma + \beta} + \alpha \frac{\gamma\beta(\gamma - \beta)}{(\gamma + \beta)(2\gamma + \beta)(\gamma + 2\beta)}$$

- Bivariate Poisson

$$c(\boldsymbol{\theta}) = e^{-\lambda_L - \lambda_X} \sum_{l=0}^{\infty} \frac{\lambda_L^l}{l!} S_V(l)$$

- Bernoulli-Poisson

$$c(\boldsymbol{\theta}) = 1 - p e^{-\lambda_X}$$

Ref: Emura and Konno (2012a, b): Domma and Giordano (2013)

- However, a general formula of $c(\boldsymbol{\theta})$ seems unavailable

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

$$= \iint_{l \le x} C_{\alpha}^{[1,1]}[F_L(l; \boldsymbol{\theta}_L), F_X(x; \boldsymbol{\theta}_X)]\, f_L(l; \boldsymbol{\theta}_L)\, f_X(x; \boldsymbol{\theta}_X)\,dl\,dx$$

- Define *h-function* (Schepsmeier and StÖber, 2014)

$$h_{\alpha}(u, v) \equiv \partial C_{\alpha}(u, v) / dv$$

Theorem 1 (new result in this work)

The inclusion probability can be simplify as follows:

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

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

Proposed likelihood estimation

- The log-likelihood function:

$$\ell_n(\boldsymbol{\theta}) = -n\log c(\boldsymbol{\theta}) + \sum_j \log f_L(L_j;\boldsymbol{\theta}_L) + \sum_j \log f_X(X_j;\boldsymbol{\theta}_X) + \sum_j \log C_\alpha^{[1,1]}[F_L(L_j;\boldsymbol{\theta}_L), F_X(X_j;\boldsymbol{\theta}_X)],$$

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

- Maximum likelihood estimator (MLE):

$$\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\theta}}_L, \hat{\boldsymbol{\theta}}_X) = \arg\max_{\boldsymbol{\theta}\in\Theta} \ell_n(\boldsymbol{\theta})$$

- The Newton-Raphson algorithm

$$\boldsymbol{\theta} = \boldsymbol{\theta} - [\partial^2 \ell_n(\boldsymbol{\theta}) / \partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T]^{-1} \partial \ell_n(\boldsymbol{\theta}) / \partial\boldsymbol{\theta}$$

Hessian — Score

Example: Weibull models with the Clayton copula

- Clayton copula: $C_\alpha(u_1, u_2) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha}, \quad \alpha \geq 0$

  ➔ h-function: $h_\alpha(u_1, u_2) = u_2^{-\alpha-1}(u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha-1}$

- Weibull marginals: $F_L(l; \lambda_L, \nu_L) = 1 - \exp(-\lambda_L l^{\nu_L})$

  $F_X(x; \lambda_X, \nu_X) = 1 - \exp(-\lambda_X x^{\nu_X})$

- Unknown parameters: $\boldsymbol{\theta} = (\alpha, \lambda_L, \lambda_X, \nu_L, \nu_X) \in \Theta = (0, \infty)^5$

- Sample inclusion probability:

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

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

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

Example: Weibull models with the Clayton copula

Computation of $c(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(L \le X) = \iint\limits_{l \le x} f_{L,X}(l, x; \boldsymbol{\theta}) dl dx$

- **Proposed method**

$$c_{\alpha}(\boldsymbol{\theta}) \equiv \frac{\partial c(\boldsymbol{\theta})}{\partial \alpha} = \int_0^1 \frac{\partial H(u;\boldsymbol{\theta})}{\partial \alpha} du \qquad H(u;\boldsymbol{\theta}) = u^{-\alpha-1} B(u, \boldsymbol{\theta})^{-1/\alpha-1}$$

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

- **Naïve method (double-integral)**

$$c(\boldsymbol{\theta}) = \lambda_L \lambda_X \nu_L \nu_X (1+\alpha)$$

$$\times \iint\limits_{l \le x} \frac{l^{\nu_L - 1} x^{\nu_X - 1} \exp(-\lambda_L l^{\nu_L}) \exp(-\lambda_X x^{\nu_X})[\{1-\exp(-\lambda_L l^{\nu_L})\}\{1-\exp(-\lambda_X x^{\nu_X})\}]^{-\alpha-1}}{(\{1-\exp(-\lambda_L l^{\nu_L})\}^{-\alpha} + \{1-\exp(-\lambda_X x^{\nu_X})\}^{-\alpha} - 1)^{1/\alpha+2}} dl dx .$$

Two methods produce the same value, but different computing time .

Example:
Weibull models with the Clayton copula

- **Score function** (for α):

$$\partial \ell_n(\boldsymbol{\theta})/\partial\alpha = -nc_\alpha(\boldsymbol{\theta})/c(\boldsymbol{\theta}) + \sum_j \partial \log\{\, f_{L,X}(L_j, X_j; \boldsymbol{\theta})\}/\partial\alpha$$

where

$$c_\alpha(\boldsymbol{\theta}) \equiv \frac{\partial c(\boldsymbol{\theta})}{\partial\alpha} = \int_0^1 \frac{\partial H(u;\boldsymbol{\theta})}{\partial\alpha}\, du$$

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

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

$$\frac{\partial \log f_{L,X}(l,x;\boldsymbol{\theta})}{\partial\alpha} = \frac{1}{1+\alpha} - \log\{1-\exp(-\lambda_L l^{\nu_L})\} - \log\{1-\exp(-\lambda_X x^{\nu_X})\}$$

$$+\frac{\log A(l,x;\boldsymbol{\theta})}{\alpha^2} - \left(\frac{1}{\alpha}+2\right)\frac{A_\alpha(l,x;\boldsymbol{\theta})}{A(l,x;\boldsymbol{\theta})},$$

$$A_\alpha(l,x;\boldsymbol{\theta}) \equiv \frac{\partial A(l,x;\boldsymbol{\theta})}{\partial\alpha} = -\frac{\log\{1-\exp(-\lambda_L l^{\nu_L})\}}{\{1-\exp(-\lambda_L l^{\nu_L})\}^\alpha} - \frac{\log\{1-\exp(-\lambda_X x^{\nu_X})\}}{\{1-\exp(-\lambda_X x^{\nu_X})\}^\alpha},$$

$\max\{|\alpha^{(k+1)}-\alpha^{(k)}|,|\lambda_L^{(k+1)}-\lambda_L^{(k)}|,|\lambda_X^{(k+1)}-\lambda_X^{(k)}|,|\nu_L^{(k+1)}-\nu_L^{(k)}|,|\nu_X^{(k+1)}-\nu_X^{(k)}|\}<\varepsilon$

Newton-Raphson algorithm

Step 1: Set the initial values

$$\alpha^{(0)}=2\hat{\tau}/(1-\hat{\tau}),\ \lambda_L^{(0)}=1/\bar{L},\ \lambda_X^{(0)}=1/\bar{X},\ \nu_L^{(0)}=1 \text{ and } \nu_X^{(0)}=1$$

Step 2: Repeat

$$\begin{bmatrix}\alpha^{(k+1)}\\ \lambda_L^{(k+1)}\\ \lambda_X^{(k+1)}\\ \nu_L^{(k+1)}\\ \nu_X^{(k+1)}\end{bmatrix}=\begin{bmatrix}\alpha^{(k)}\\ \lambda_L^{(k)}\\ \lambda_X^{(k)}\\ \nu_L^{(k)}\\ \nu_X^{(k)}\end{bmatrix}-\left.\begin{bmatrix}\frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\alpha^2} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_X\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_X\partial\alpha}\\ \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L^2} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L\partial\lambda_X} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L\partial\nu_L} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L\partial\nu_X}\\ \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_X\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_L\partial\lambda_X} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\lambda_X^2} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\lambda_X} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_X\partial\lambda_X}\\ \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\lambda_L} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\lambda_X} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L^2} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\nu_X}\\ \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_X\partial\alpha} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_X\partial\lambda_L} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_X\partial\lambda_X} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_L\partial\nu_X} & \frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial\nu_X^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial\ell_n(\boldsymbol{\theta})}{\partial\alpha}\\ \frac{\partial\ell_n(\boldsymbol{\theta})}{\partial\lambda_L}\\ \frac{\partial\ell_n(\boldsymbol{\theta})}{\partial\lambda_X}\\ \frac{\partial\ell_n(\boldsymbol{\theta})}{\partial\nu_L}\\ \frac{\partial\ell_n(\boldsymbol{\theta})}{\partial\nu_X}\end{bmatrix}\right|_{\substack{\alpha=\alpha^{(k)}\\ \lambda_L=\lambda_L^{(k)},\ \lambda_X=\lambda_X^{(k)}\\ \nu_L=\nu_L^{(k)},\ \nu_X=\nu_X^{(k)}}}$$

If $\|\boldsymbol{\theta}^{(k+1)}-\boldsymbol{\theta}^{(k)}\|<\varepsilon$, (e.g., $\varepsilon=10^{-4}$)

then, $\boldsymbol{\theta}^{(k+1)}=(\alpha^{(k+1)},\lambda_L^{(k+1)},\lambda_X^{(k+1)},\nu_L^{(k+1)},\nu_X^{(k+1)})$ is the MLE

Asymptotic inference

- Target : $g(\boldsymbol{\theta})$
- Standard error (SE)

Last step of
the Newton-Raphson

$$SE\{g(\hat{\boldsymbol{\theta}})\} = \sqrt{\left\{\frac{\partial}{\partial \boldsymbol{\theta}} g(\hat{\boldsymbol{\theta}})\right\}^{\mathrm{T}} \times \left\{-\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\mathrm{T}}} \ell_n(\hat{\boldsymbol{\theta}})\right\}^{-1} \times \frac{\partial}{\partial \boldsymbol{\theta}} g(\hat{\boldsymbol{\theta}})}$$

- $(1-\alpha)100\%$ confidence interval

$$[\, g(\hat{\boldsymbol{\theta}}) - Z_{\alpha/2} \cdot SE\{g(\hat{\boldsymbol{\theta}})\},\quad g(\hat{\boldsymbol{\theta}}) + Z_{\alpha/2} \cdot SE\{g(\hat{\boldsymbol{\theta}})\}\,]\,.$$

- Example: Mean failure time

$$g(\boldsymbol{\theta}) = E(X) = \int x f_X(x; \boldsymbol{\theta}_X)\,dx$$

Goodness-of-fit test

- Cramér-von Mises type or K-S statistic

$$CM = \iint_{l \le x} n\{ \hat{F}_{L,X}(l,x) - F_{L\le X}(l,x;\hat{\boldsymbol{\theta}}) / c(\hat{\boldsymbol{\theta}}) \}^2 d\hat{F}(l,x)$$

$$KS = \sup_{x,y} | \hat{F}_{L,X}(l,x) - F_{L\le X}(l,x;\hat{\boldsymbol{\theta}}) / c(\hat{\boldsymbol{\theta}}) |$$

1) Nonparametric estimate

$$\hat{F}_{L,X}(l,x) = \sum_j \mathbf{I}(L_j \le l,\ X_j \le x) / n$$

2) Parametric estimate

$$F_{L\le X}(l,x;\hat{\boldsymbol{\theta}}) / c(\hat{\boldsymbol{\theta}})$$

Theorem 2: *The truncated distribution function is expressed as the univariate integrals*

$$F_{L\le X}(l,x;\boldsymbol{\theta}) = \Pr(L \le l,\ X \le x, L \le X) = \int_0^{F_X(l)} H(v;\boldsymbol{\theta})dv + \int_{F_X(l)}^{F_X(x)} h_\alpha\{ F_L(l), v\}dv .$$

Goodness-of-fit test

- Null hypothesis
  $H_0 : F_{\boldsymbol{\theta}}(l, x) = C_{\alpha}[\, F_L(l; \boldsymbol{\theta}_L), F_X(x; \boldsymbol{\theta}_X)\,]$ $\exists \alpha, \boldsymbol{\theta}_L, \boldsymbol{\theta}_X$
- A parametric Bootstrap test similar to
  Emura and Konno (2012a, b)

Step1. Compute CM (Cramér-von-Mises) or K-S type statistics).

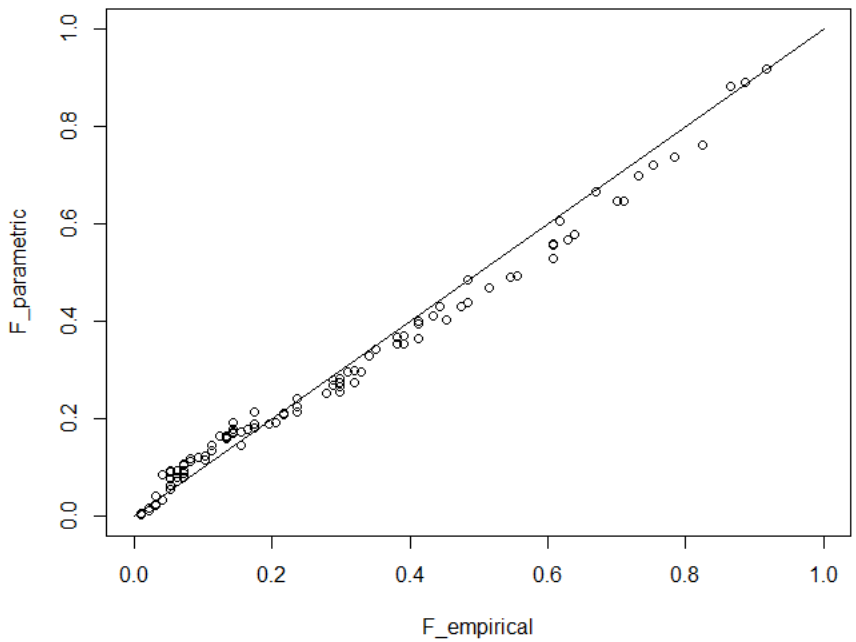
Step2. $(L_k^{(b)}, X_k^{(b)}) \sim C_{\hat{\alpha}}[\, F_L(l; \hat{\boldsymbol{\theta}}_L), F_X(x; \hat{\boldsymbol{\theta}}_X)\,]/c(\hat{\boldsymbol{\theta}})$, for $b = 1, 2, \ldots, B$, $k = 1, 2, \ldots, n$.

Step3. $CM^{*(b)} = \sum_k \{\hat{F}_e^{(b)}(L_k^{(b)}, X_k^{(b)}) - C_{\hat{\alpha}^{(b)}}[\, F_L(L_k^{(b)}; \hat{\boldsymbol{\theta}}_L^{(b)}), F_X(X_k^{(b)}; \hat{\boldsymbol{\theta}}_X^{(b)})\,]/c(\hat{\boldsymbol{\theta}}^{(b)})\}^2$

Step4. Approximate p-values is $\sum_b (CM^{*(b)} \geq CM)/B$.

- Approximate p-value < 0.05 ➔ Reject

R package *depend.truncation*




```
> PMLE.Clayton.Weibull(l.trunc,x.trunc)
$alpha
  Estimate           SE
0.2417720         0.1968614
$lambda_L
      Estimate              SE
0.019245463         0.006887897
$lambda_X
        Estimate               SE
0.0001256100         0.0001038773
$nu_L
  Estimate           SE
1.4567607          0.1160334
$nu_X
  Estimate           SE
2.1628590            0.1860137
$mean_X
  Estimate           SE
56.345513          2.937671
$CM       0.09483959
$KS        0.07892345
```

Simulation setting

- Weibull model with the Clayton copula

$$\Pr(L \le l, X \le x)$$

$$= [\{1 - \exp(-\lambda_L l^{\nu_L})\}^{-\alpha} + \{1 - \exp(-\lambda_X x^{\nu_X})\}^{-\alpha} - 1]^{-1/\alpha}$$

where $\alpha = 2$ s.t. $(\tau = 0.5)$

- Parameters $(\lambda_L, \lambda_X, \nu_L, \nu_X)$ chosen to be

$c(\boldsymbol{\theta}) > 0.5$, $c(\boldsymbol{\theta}) = 0.5$, or $c(\boldsymbol{\theta}) < 0.5$

- Target: $(\alpha, \lambda_L, \lambda_X, \nu_L, \nu_X)$

$$\mu_X = E(X) = \Gamma(1 + 1/\nu_X)/(\lambda_X^{1/\nu_X})$$

- Generatè $\{(L_j, X_j); j = 1, 2, \ldots, n\}$ subject to $L_j \le X_j$

$v_0 = u_0$

Simulation results under Weibull models with the Clayton copula based on 1000 replications

True: $\alpha = 2$ $(\tau = 0.5)$

$c(\boldsymbol{\theta})$	n	$E(\hat{\alpha})$	$E(\hat{\lambda}_L)$	$E(\hat{\lambda}_X)$	$E(\hat{\nu}_L)$	$E(\hat{\nu}_X)$	AI
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
0.500	100	2.167	1.023	1.130	0.992	0.979	65.2
	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
0.387	100	2.064	0.971	1.169	2.123	1.025	69.9
	200	2.020	0.986	1.050	2.044	1.018	54.2
	300	2.003	0.986	1.019	2.036	1.019	50.0

AI = The average number of Newton-Raphson iterations until convergence

$\lambda_X = \nu_X = 1$ (True)

$v_0 - u_0$

Simulation results under Weibull models with the Clayton copula based on 1000 replications

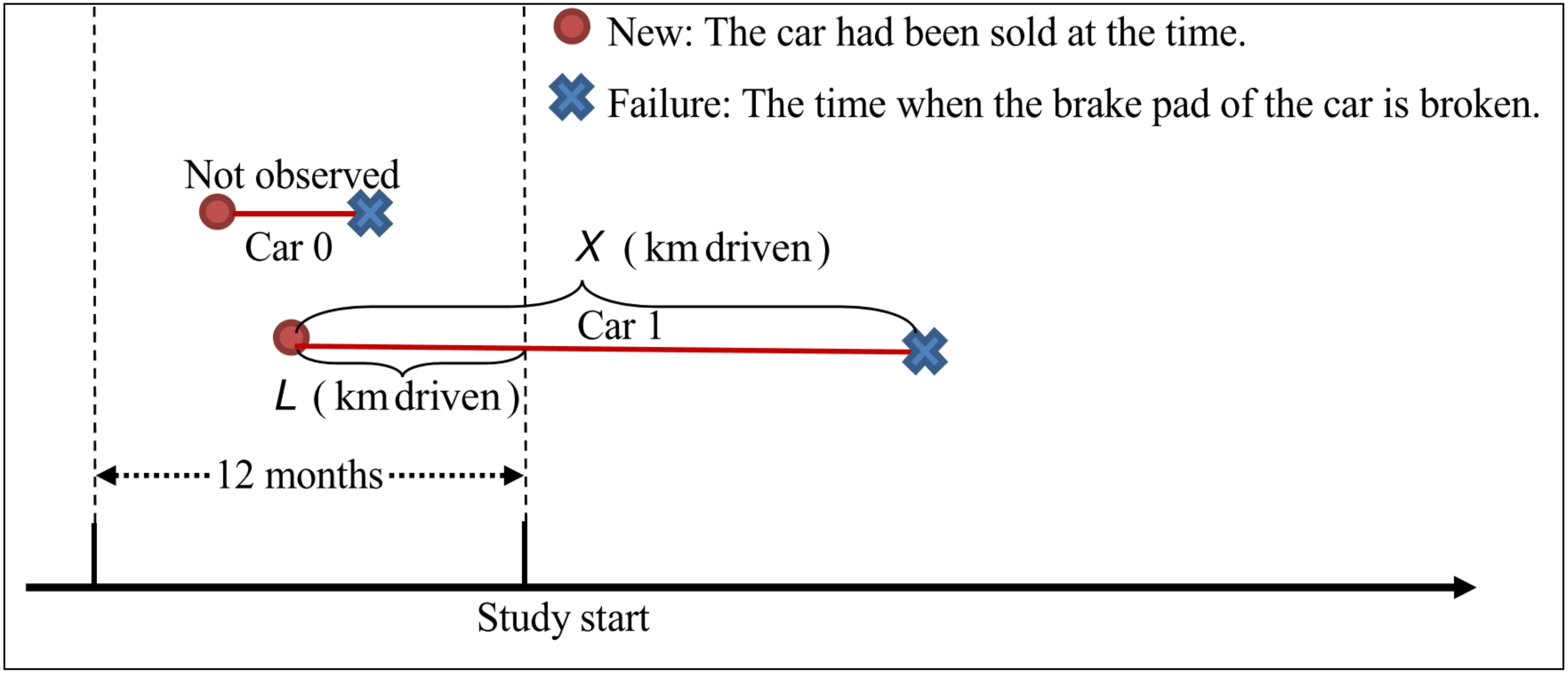
$c(\boldsymbol{\theta})$	n	$SD(\hat{\alpha})$	$E\{SE(\hat{\alpha})\}$	95%Cov	$SD(\hat{\mu}_X)$	$E\{SE(\hat{\mu}_X)\}$	95%Cov
0.804	100	0.428	0.448	0.954	0.117	0.112	0.947
	200	0.305	0.312	0.951	0.077	0.076	0.948
	300	0.250	0.254	0.949	0.062	0.062	0.960
0.500	100	0.576	0.552	0.944	0.232	0.198	0.916
	200	0.348	0.351	0.946	0.157	0.145	0.928
	300	0.268	0.279	0.948	0.123	0.118	0.941
0.387	100	1.009	0.894	0.921	0.344	0.308	0.919
	200	0.607	0.585	0.940	0.232	0.230	0.952
	300	0.474	0.456	0.934	0.190	0.187	0.958

95%Cov = Coverage rates for the 95% confidence intervals.

$$\mu_X = E(X) = \Gamma(1+1/\nu_X)/(\lambda_X^{1/\nu_X})$$

Car brake pads data
(Kalbfleisch and Lawless,1992 JQT)

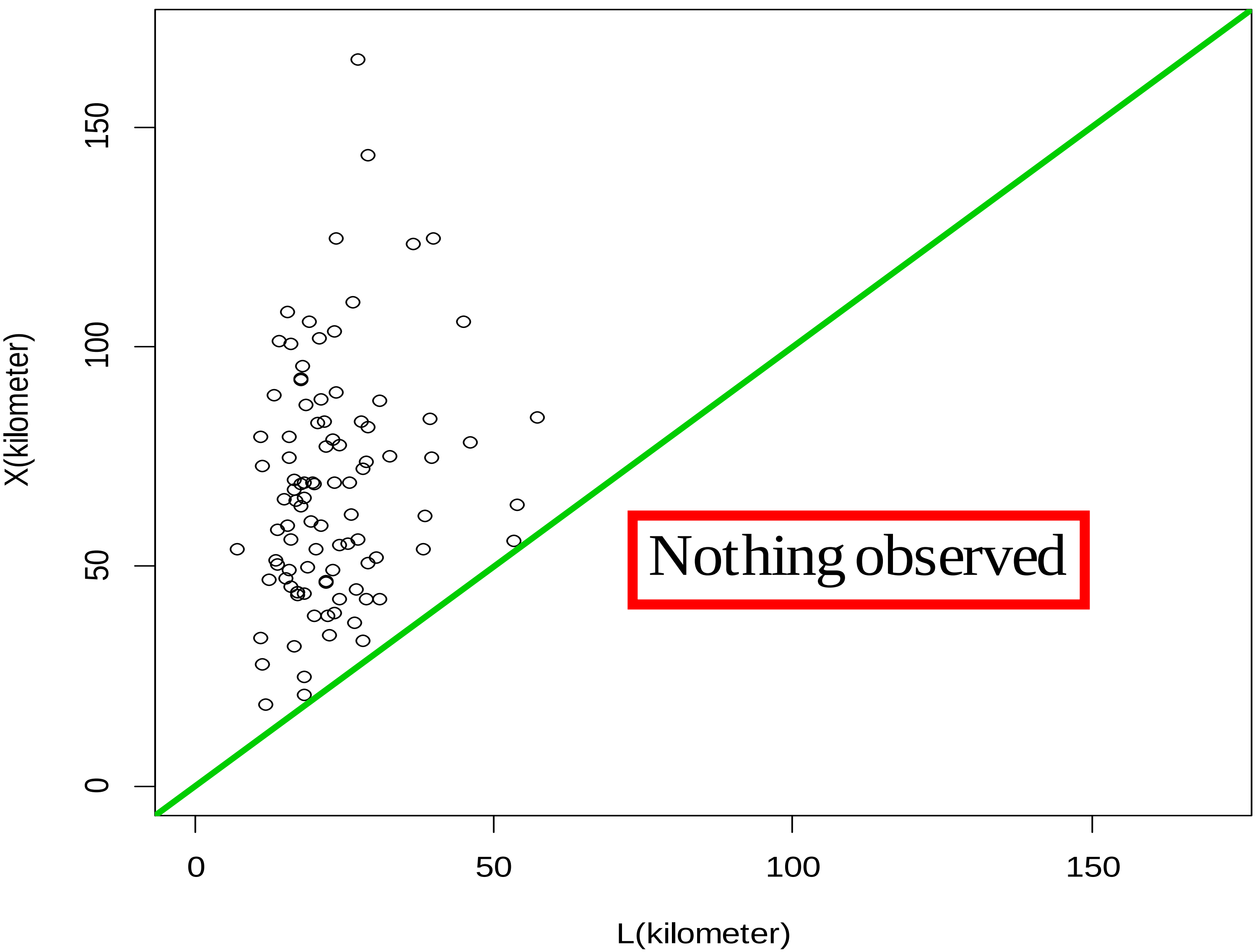
- X : The number of kilometers driven until failure
- L : The number of kilometers driven until the study start



Observations : $\{ (L_j, X_j); j = 1, 2, \ldots, n \}$ subject to $L_j \le X_j$

Scatter plot of the brake pads data

$$\{ (L_j, X_j); \; j = 1, 2, \ldots, n \}, n = 98, \quad \text{subject to } L_j \le X_j$$



	Weibull lifetime Clayton copula	Exponential lifetime Clayton copula	Lognormal lifetime Gaussian copula
Copula parameter	$\hat{\alpha}$ =0.242 (SE=0.197) 95%CI: [0.049, 1.193]	$\hat{\alpha}$ =0.000 (SE=0.006) 95%CI: Not available	$\hat{\rho}$ =0.209 (SE=0.112) 95%CI: [-0.011, 0.429]
Parameters for L	$\hat{\lambda}_L$ =0.019 (SE=0.007) $\hat{\nu}_L$ =1.457 (SE=0.116)	$\hat{\lambda}_L$ =0.054 (SE=0.008) ν_L =1 (fixed)	$\hat{\mu}_L$ = 2.38 (SE=0.097) σ_L^2 =0.722 (SE= 0.124)
Parameters for X	$\hat{\lambda}_X$ =0.00013 (SE=0.0001) $\hat{\nu}_X$ =2.163 (SE=0.186) Accept at 10%	$\hat{\lambda}_X$ =0.022 (SE=0.002) ν_X =1 (fixed)	$\hat{\mu}_X$ =3.92 (SE=0.054) σ_X^2 =0.266 (SE=0.039)
KS test CvM test	K = 0.079 (P= 0.200) C = 0.095 (P= 0.113)	K = 0.189 (P= 0.000) C = 1.066 (P= 0.000)	K = 0.066 (P= 0.494) C = 0.099 (P= 0.094)

SE=Standard Error; KS test=Kolmogorov-Smirnov test for goodness-of-fit; CvM test=Cramér-von Mises test for goodness-of-fit; P=P-value based on the parametric bootstrap; 95%CI=95% Confidence Interval (it is not available if $\hat{\alpha}$ =0)

Weibull is the best model

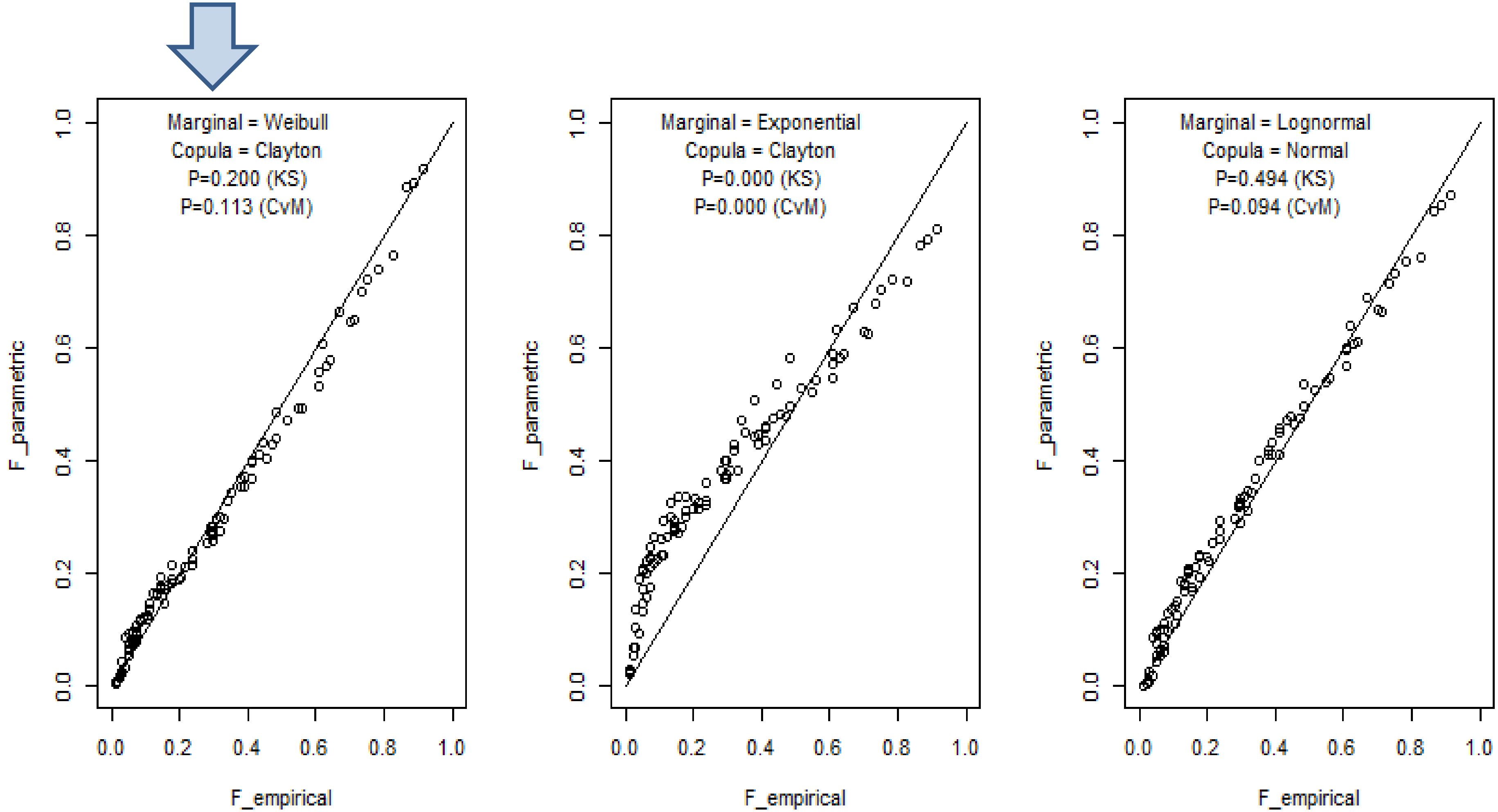
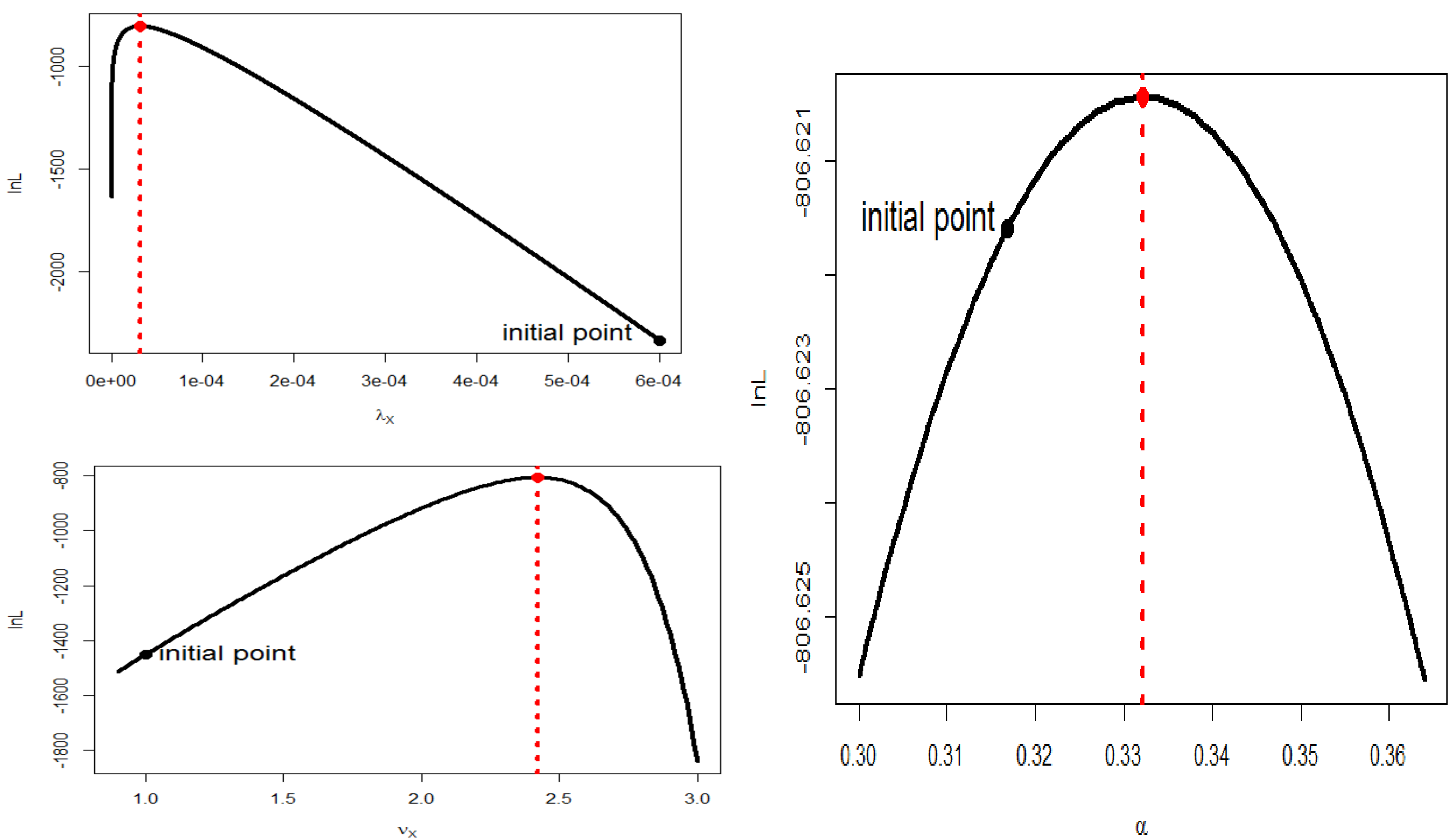


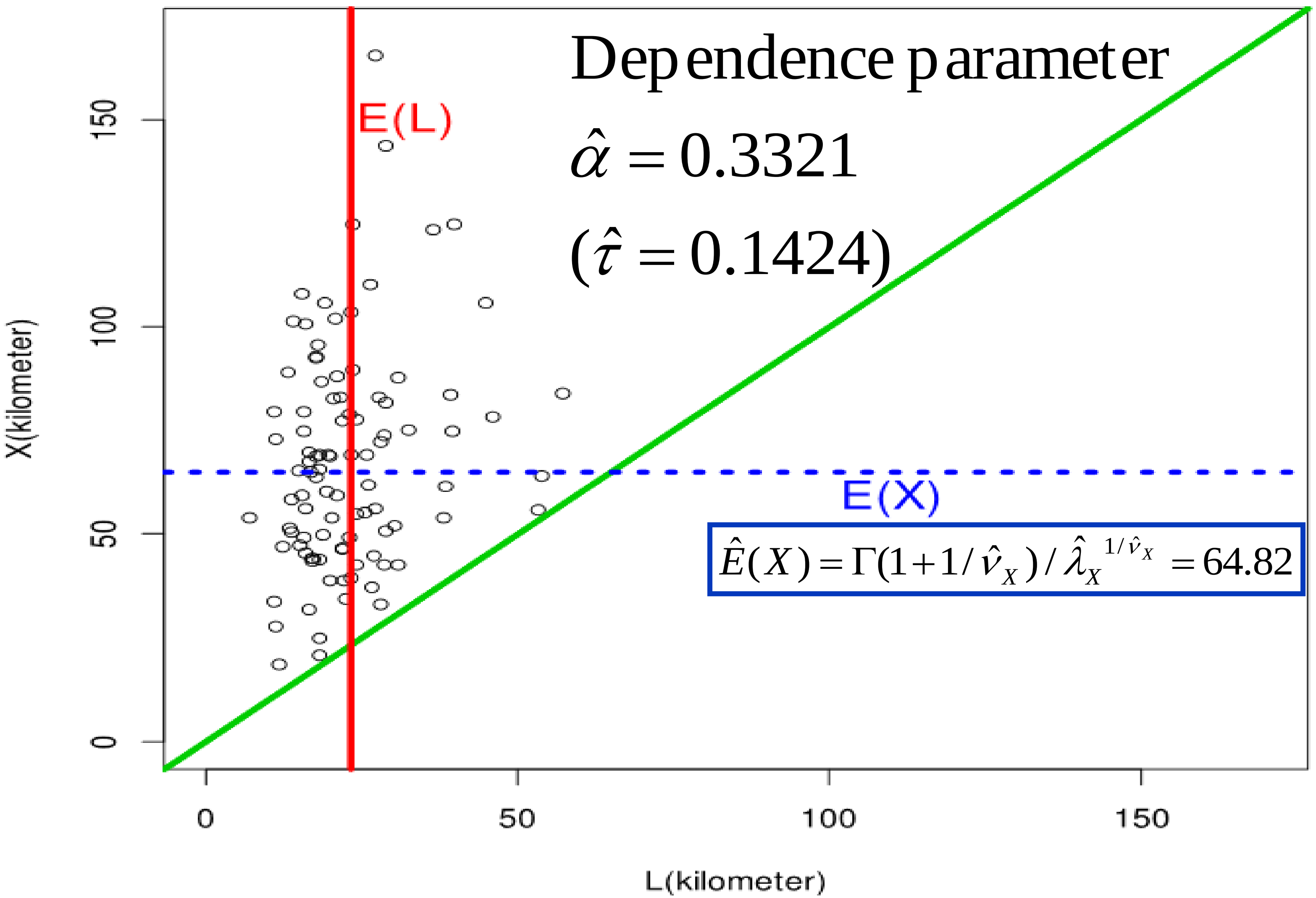
Figure 4. Goodness-of-fit tests for the brake pad lifetime data (Kalbfleisch and Lawless 1992). KS=Kolmogorov-Smirnov test; CvM=Cramér-von Mises test; P=P-value based on the parametric bootstrap.

The log-likelihood under the Weibull

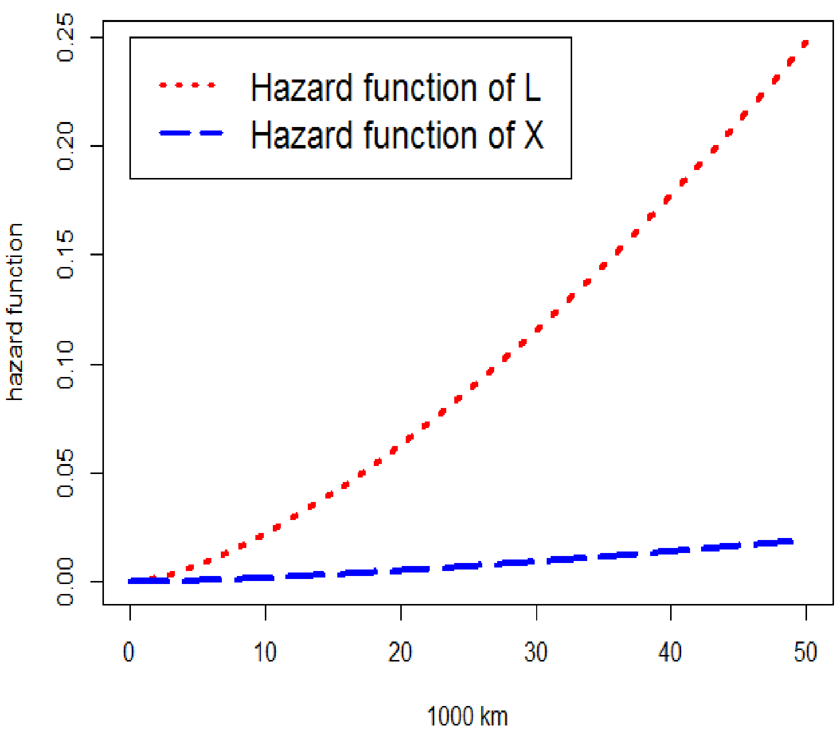


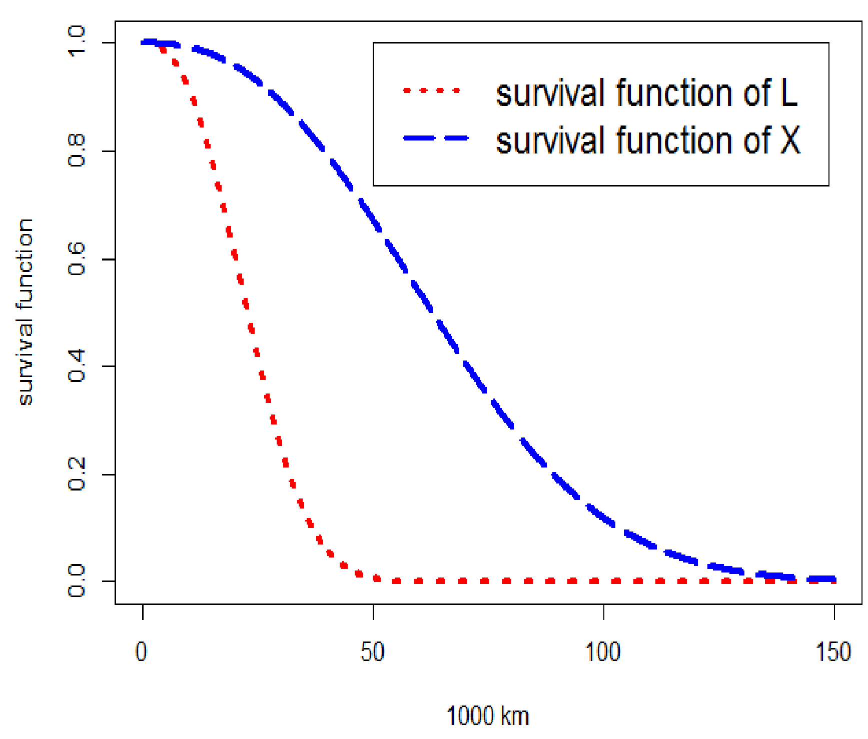
Parameter estimates under the Weibull model

$$\hat{E}(L) = \Gamma(1+1/\hat{\nu}_L)/\hat{\lambda}_L^{1/\hat{\nu}_L} = 23.33$$



Estimated survival functions for marginal distributions: **Weibull models**





Estimates: $\hat{\lambda}_L = 2.89 \times 10^{-4}$, $\hat{\nu}_L = 2.4927$; $\hat{\lambda}_X = 3.09 \times 10^{-5}$, and $\hat{\nu}_X = 2.4193$.

Extension to doubly truncation

- (L, X, R) : a pair of random variables
- L: left-truncation time
- X: failure time
- R: right-truncation time

Sample inclusion criterion:

If $L \le X \le R$, the sample is available

Doubly-truncated samples

$\{(L_j, X_j, R_j); j = 1, 2, \ldots, n\}$ subject to $L_j \le X_j \le R_j$

Existing inference methods are developed under $(L, R) \perp X$

Efron and Petrosian (1999), Shen (2010, 2011, 2017), Emura and Hu (2015), Moreira and Álvarez (2010, 2012), Moreira and Van Keilegom (2013), Dörre (2017), Frank and Dörre (2017)

$\boldsymbol{\theta} = (\alpha, \boldsymbol{\theta}_L, \boldsymbol{\theta}_X) \in \Theta$

Extension to doubly truncation

- Joint distribution:

$$\Pr(L \le l,\ X \le x,\ R \le r) = C_{\boldsymbol{\alpha}}[\,F_L(l), F_X(x), F_R(r)\,]$$

- Copula: $C_{\boldsymbol{\alpha}} : [0,1]^3 \mapsto [0,1]$,

  $\boldsymbol{\alpha} \in R$: dependence parameters

- Continuous parametric margins

$$F_L(l) = P(L \le l), \quad F_X(x) = P(X \le x), \quad F_R(r) = P(R \le r)$$

Theorem 3: Let $h(u, v, w) \equiv \partial C_{\boldsymbol{\alpha}}(u, v, w) / v$

$\Pr(L \le X \le R) =$

$$\int_0^1 h[\,F_L\{F_X^{-1}(v), v, 1\,]dv - \int_0^1 h[\,F_L\{F_X^{-1}(v), v, F_R\{F_X^{-1}(v)\,]dv\,.$$

More developments for *dependent double-truncation* are necessary and remain to be done

Thanks for listening