On competing risks and problem of identification

Petr Volf, ÚTIA AV ČR

O U T L I N E:

1. Introduction, problem of competing risks
2. Observed data, cumulative incidence function
3. Problem of model identifiability
4. Copula as a tool expressing multivariate distributions,
   use in competing risks setting
5. Competing risks with regression, again question of identifiability
6. Examples using gaussian copula, ML estimation
7. How good is approximation with gaussian copula?

1 Competing risks and incidence

Situation: Certain event (e. g. a failure of a device) can be caused by several reasons

– there is K (possibly dependent) random variables $T_j, j = 1, ..., K$,

(we do not consider any censoring)

Denote $\overline{F}_K(t_1, ..., t_K) = P(T_1 > t_1, ..., T_K > t_K)$ their joint survival function.

We observe $Z = \min(T_1, ..., T_K)$ and indicator $\delta = j$ if $Z = T_j$.

In general, from data (Z_i, δ_i), $i = 1, \ldots, N$ it is not possible to identify joint distribution of (T_j),

while the **incidence** of T_j, i.e. the distributions of $(T_j, \delta = j)$ can be estimated, consistently.

2 Cumulative Incidence Function

Estimable: distribution of $Z = \min(T_1, ..., T_K)$

e.g. $S(t) = P(Z > t) = \overline{F}_K(t, ..., t)$ – its survival function

and **"Incidence densities"**

$$f_j^*(t) = P(Z = t, \delta = j) = -\frac{\partial \overline{F}_K(t_1, ..., t_K)}{\partial t_j} |(t_1 = ... = t_K = t),$$

so that also their integrals, "Cumulative incidence functions"

$$F_j^*(t) = \int_0^t f_j^*(s)\, ds = P(Z \leq t, \delta = j).$$

Notice that $\lim F_j^*(t) = P(\delta = j) < 1$ if $t \to \infty$, $\quad S(t) = 1 - \sum_{j=1}^K F_j^*(t)$.

A more practical form of incidence function:

Estimable: Cause–specific hazard functions for events $j = 1, 2, \ldots, K$:

$$h_j^*(t) = \lim_{d \to 0} \frac{P(t \le Z < t + d,\, \delta = j \,|\, Z \ge t)}{d},$$

overall hazard rate for $Z = \min(T_1, ..., T_K)$:

$$h^*(t) = \lim_{d \to 0} \frac{P(t \le Z < t + d \,|\, Z \ge t)}{d} = \sum_{j=1}^{K} h_j^*(t),$$

integrals = cumulated hazard rates $H_j^*(t)$, $H^*(t)$,

overall survival function $S(t) = P(Z > t) = \exp(-H^*(t))$.

Then $$f_j^*(t) = h_j^*(t) \cdot S(t)$$

and **cumulative incidence functions** are:

$$F_j^*(t) = P(Z \le t, \delta = j) = \int_0^t S(s) \cdot h_j^*(s)\,\mathrm{d}s.$$

3 Non-identifiability

A. Tsiatis (1975) has shown that for arbitrary joint model we can find a model with independent components having the same incidences,

i.e. we cannot distinguish the models.

Remark: Even if model is parametric and MLE yields consistent estimates, we don't know parameters of which model are estimated.

Namely, this 'independent' model is given by cause-specific hazard functions $h_j^*(t)$.

Example (Tsiatis 1975)

Consider just $K = 2$ random variables S, T with exponential marginal and joint survival functions

$$\overline{F}_S(s) = e^{-\lambda s}, \;\; \overline{F}_T(t) = e^{-\mu t}, \;\; \overline{F}_2(s,t) = e^{-\lambda s - \mu t - \theta s t}.$$

Hence, $S(t) = \overline{F}_2(t,t) = exp(-\lambda t - \mu t - \theta t^2)$.

Corresponding cause-specific hazard rates and their integrals are

$$h_S^*(t) = (\lambda + \theta t), \;\; h_T^*(t) = (\mu + \theta t), \;\; H_S^*(t) = (\lambda t + \frac{\theta}{2} t^2), \;\; H_T^*(t) = (\mu t + \frac{\theta}{2} t^2),$$

and $S(t) = exp(-H_S^*(t) + H_T^*(t))$ is the same as above.

It means that independent random variables with marginal survival functions

$$\overline{G}_S(s) = e^{-\lambda s - \frac{\theta}{2} s^2}, \;\; \overline{G}_T(t) = e^{-\mu t - \frac{\theta}{2} t^2}$$

yield the same competing risk scheme.

4 Competing risk and copula

In the sequel we shall also consider just 2 competing events, i.e. random variables S, T and data $Z_i = min(S_i, T_i), \delta_i = 1, 2$.

Copula offers a way how to model their joint distribution function:

$$F_2(s,t) = C(F_S(s),\ F_T(t)), \tag{1}$$

where F_S, F_T are marginal distribution functions of variables S, T.

Zheng and Klein (1995) proved that when the copula is known, the marginal distributions are estimable consistently from 'competing risk' data (and then also joint distribution, from (1)).

They dealt with non-parametric (so that quite general) case.

It is obvious that the "knowledge" of copula is still an unrealistic supposition.

Nevertheless, we can try to use certain sufficiently flexible class of copulas, for approximation.

So called **Survival copula** ties survival functions:

As

$$\overline{F}_2(s,t) = 1 - F_S(s) - F_T(t) + F_2(s,t) = \overline{F}_S(s) + \overline{F}_T(t) - 1 + C(F_S(s),\ F_T(t)),$$

then

$$\overline{F}_2(s,t) = \overline{C}(\overline{F}_S(s), \overline{F}_T(t)),$$

where $\overline{C}$ is also copula, namely

$$\overline{C}(u,v) = u + v - 1 + C(1-u, 1-v).$$

5 Competing risks and regression models

let us first return to Tsiatis' example

and add assumption that both variables follow Cox model with a covariate x:

$$\overline{F}_S(s) = e^{-a(x)s},\ \overline{F}_T(t) = e^{-b(x)t},\ \overline{F}(s,t) = e^{-a(x)s-b(x)t-\theta st},$$

(θ can also depend on x).

Then in 'equivalent' independent model it should hold that

$$\overline{F}_S(s) = e^{-a(x)s-\theta s^2/2},\ \overline{F}_T(t) = e^{-b(x)t-\theta t^2/2},$$

– it is not a form of Cox model, unless $a(x) = c_1\theta(x) = c_2 b(x)$.

Identifiability in Cox model case, Heckman and Honoré (1990):

Let $H_S(s;x) = H_{0S}(s) \cdot a(x)$, $H_T(t;x) = H_{0T}(t) \cdot b(x)$, $H_{0S}(0) = H_{0T}(0) = 0$,

$\overline{F}(s,t;x) = C(\overline{F_S}(s;x), \overline{F}_T(t;x)) = C(exp(-H_{0S}(s) \cdot a(x)), exp(-H_{0T(t)} \cdot b(x)))$.

Then

$$f_S^* = -\frac{\partial \overline{F}(s,t;x)}{\partial s} |(s=t) =$$

$$= C^{(1)}(exp(-H_{0S}(t) \cdot a(x)), exp(-H_{0T}(t) \cdot b(x))) \cdot exp(-H_{0S}(t) \cdot a(x)) \cdot h_{0S}(t) \cdot a(x).$$

Step 1: Imagine, in neighborhood of each x, lim $t \to 0$,

we can 'estimate' proportions $a(x_1)/a(x_2)$, so that $a(x)$ up to a multipl. constant.

(– it is seen that $H_{S0}(t) \cdot a(x)$ are determined up to a multipl. constant)

Step 2: Consider now

$\overline{F}(t,t;x) = C(exp(-H_{0S}(t) \cdot a(x)), exp(-H_{0T}(t) \cdot b(x)))$

at such t where both $H_{0S}(t) = H_{0T}(t) = 1$, at different x.

We can 'estimate' $C(exp(-a(x)), exp(-b(x)))$, so that copula C.

Step 3: We already 'know' $C(u,v)$, $a(x)$, $b(x)$, it remains to estimate marginal distributions.

Here we can utilize result of Zheng, Klein (1995) to estimation

$\overline{F}_S(s;x)$ and $\overline{F}_T(t;x)$, at fixed x,

and then baseline marginal distributions $\overline{F}_{0S}(s)$ and $\overline{F}_{0T}(t)$.

Remark 1: H+H solve question whether it is possible to identify model, it is not the procedure of practical estimation.

Estimation should be based e.g. on the ML method.

Remark 2: Copula C does not depend on x, if the Cox model is taken as a case of 'transformation' model:

Let r.v. T_x fulfills Cox model with regression function $a(x)$,

hence $H_T(T_x) = H_O(T_x) \cdot a(x) \sim Exp(1)$ distribution.

Transformation model then means that T_x rises from a r.v. $e_0 \sim Exp(1)$:

$e_0 = H_O(T_x) \cdot a(x),$ for each x

and copula C then actually ties two $Exp(1)$ random variables.

Case of the **AFT** (Accelerated Failure Time) model

Here marginal survival functions are

$$\overline{F}_S(s;x) = \overline{F}_{0S}(s \cdot a(x)), \;\; \overline{F}_T(t;x) = \overline{F}_{0T}(t \cdot b(x))$$

and joint survival function can be expressed via copula

$$\overline{F}(s,t;x) = C(\overline{F}_{0S}(s \cdot a(x)), \overline{F}_{0T}(t \cdot b(x))).$$

Then

$$f_S^* = -\frac{\partial \overline{F}(s,t;x)}{\partial s}|(s=t) \;=\; C^{(1)}(\overline{F}_{0S}(t \cdot a(x)), \overline{F}_{0T}(t \cdot b(x))) \cdot f_{0S}(t \cdot a(x)) \cdot a(x),$$

'proof' of identifiability uses similar steps as in the Cox model case (again Heckman and Honoré, 1990).

6 Special cases, Gaussian copula

Let X, Y be standard normal r. v.-s with $\rho = \rho(X, Y)$, joint density

$$\varphi_2(x, y, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{1}{2}\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}\} \tag{2}$$

where $\boldsymbol{x} = (x, y)'$ and $\Sigma =$ covariance matrix $[1, \rho; \rho, 1]$.

Take $U = \phi(X),\ V = \phi(Y)$, we obtain copula

$$C(u, v) = \phi_2(\phi^{-1}(u),\ \phi^{-1}(v), \rho). \tag{3}$$

Naturally, $\rho(U, V) \neq \rho(X, Y)$

Spearman's correlations coincide, $\rho_{\mathrm{SP}}(X, Y) = \rho_{\mathrm{SP}}(U, V) = \rho(U, V)$.

Corresponding density is:

$$c(u, v) = \frac{\varphi_2(x, y, \rho)}{\varphi(x) \cdot \varphi(y)},$$

again with $u = \phi(x),\ v = \phi(y)$, φ, ϕ are $N(0, 1)$ density and distr. function.

Approximation model for distribution of competing variables S, T:

Let $C(u,v)$ of (1) be a gaussian copula (3). Then

$$F_2(s,t) = \phi_2(\phi^{-1}(F_S(s)),\, \phi^{-1}(F_T(t)), \rho), \tag{4}$$

$$S = F_S^{-1}(\phi(X)),\ \ T = F_T^{-1}(\phi(Y)).$$

Further, $\rho_{\mathrm{SP}}(S,T) = \rho_{\mathrm{SP}}(U,V)$,

"initial" $\rho = \rho(X,Y)$ is the only one parameter describing the dependence of S and T.

All values $\rho(S,T)$ can be achieved by convenient choice of $\rho(X,Y)$.

7 ML Estimation

When parameter ρ is known, copula (3) is fully defined and it follows that the distribution of (S, T) is estimable, consistently (again Zheng, Klein,1995).

On the other hand, when marginal distributions F_S, F_T are known then $\rho = \rho(X, Y)$ is estimable, then the joint distribution $F_2(s, t)$ is, too.

Data are $(Z_i, \delta_i),\ i = 1, \ldots, N,$ the likelihood function:

$$L = \prod_{i=1}^{N} \left\{ \frac{-\partial}{\partial s} \overline{F}_2(s,t) \right\}^{I[\delta_i=1]} \cdot \left\{ \frac{-\partial}{\partial t} \overline{F}_2(s,t) \right\}^{I[\delta_i=2]} \cdot \overline{F}_2(s,t)^{I[\delta_i=0]},$$

evaluated at $s = t = Z_i$, with

$\overline{F}_2(s, t) = P(S > s,\, T > t) = 1 - F_S(s) - F_T(t) + F_2(s, t)$

and $F_2(s, t) = \phi_2(x, y, \rho)$ with $x = \phi^{-1}(F_S(s))$, $y = \phi^{-1}(F_T(t))$.

After some computation – integration of 2-dimensional Gauss density $\varphi_2(x, y, \rho)$,

$$L = \prod_{i=1}^{N} \left\{ f_S(Z_i) \left[1 - \phi_1(Y_i; \rho X_i, 1-\rho^2) \right] \right\}^{I[\delta_i=1]} \cdot$$

$$\cdot \left\{ f_T(Z_i) \left[1 - \phi_1(X_i; \rho Y_i, 1-\rho^2) \right] \right\}^{I[\delta_i=2]} \cdot \{1 - F_S(Z_i) - F_T(Z_i) + \phi_2(X_i, Y_i, \rho)\}^{I[\delta_i=0]}, \tag{5}$$

where $\phi_1(x; \mu, \sigma^2)$ denotes the c.d.f. of $N(\mu, \sigma^2)$, evaluated at x,
$X_i = \phi^{-1}(F_S(Z_i))$, $Y_i = \phi^{-1}(F_T(Z_i))$.

The problem of maximization has to be solved by a search procedure.

Parameter ρ is hidden in ϕ_1 and in ϕ_2. Distributions of S and T are present both explicitly and implicitly, in X_i, Y_i.

Following examples show that both problems (estimate of F_S, F_T for given ρ, estimate of ρ for given F_S, F_T) are solvable and seems to have unique solution.

8 How flexible is gaussian copula?

A reverse problem: Let us consider a 2-variate (continuous type) distribution of random couple (S, T), with a c.d.f $F_2(s,t)$.

How close we can approach with distribution constructed from a Gauss copula, having the same marginals $F_S(s), F_T(t)$ of S, T?

i.e. we construct

$$F_\rho^*(s,t) = \phi_2(\phi^{-1}(F_S(s)), \phi^{-1}(F_T(t)), \rho),$$

where parameter $\rho \in [-1, 1]$ should be optimized in order to achieve $\min_\rho$ of some distance between F_2 and F_ρ^*.

We shall use $\sup_{s,t} |F_2(s,t) - F_\rho^*(s,t)|$ – and show here just two examples.

8.1 Bivariate exponential distribution

One of possible construction of 2-variate exponential distribution (Marshall and Olkin, 1967):

Let X_1, X_2, X_3 be independent exp. r.v. with parameters λ_j, $j = 1, 2, 3$.

Then (S, T), with $S = min(X_1, X_3)$ and $T = min(X_2, X_3)$, have the following bivariate distribution:

Joint distribution:

$$P(S > s, T > t) = e^{-\lambda_1 s} \cdot e^{-\lambda_2 t} \cdot e^{-\lambda_3 max(s,t)},$$

Marginals are

$$S \sim Exp(\lambda_1 + \lambda_3), \quad T \sim Exp(\lambda_2 + \lambda_3), \quad corr(S, T) = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}.$$

From such a definition, an interesting feature follows – **"singularity"**,

$$P(S = T) = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}.$$

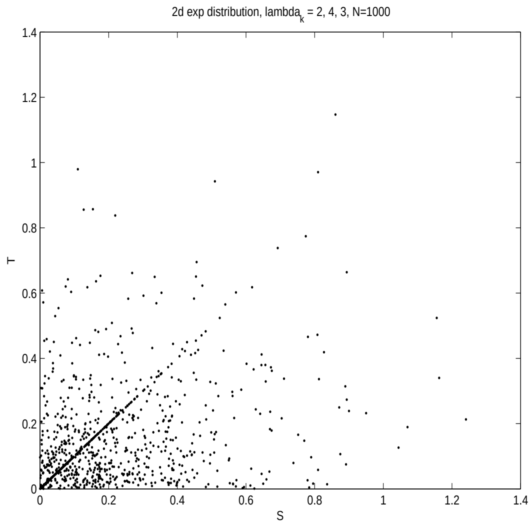

Figure 1: Generated example of 2-variate exponent. data, $\lambda_1, \lambda_2, \lambda_3 \ = \ 2, 4, 3, \ N = 1000$.

Comparison

It is clear that such a distribution cannot be fitted precisely by any $F^*_\rho(s,t)$ constructed from gaussian copula, nevertheless, we can approach rather close.

In Figure 2, the 1-st subplot displays $F_2(s,t)$,

again for the case $\lambda_1, \lambda_2, \lambda_3 \; = \; 2, 4, 3$.

The last subplot shows dependence of $\sup_{s,t} |F_2(s,t) - F^*_\rho(s,t)|$ on ρ.

It is seen that optimal $\rho \sim 0.53$ and maximal difference is ~ 0.0267.

Subplot 2 shows $F^*_\rho(s,t)$ and 3-rd subplot differences $F_2(s,t) - F^*_\rho(s,t)$, for optimal $\rho = 0.53$.

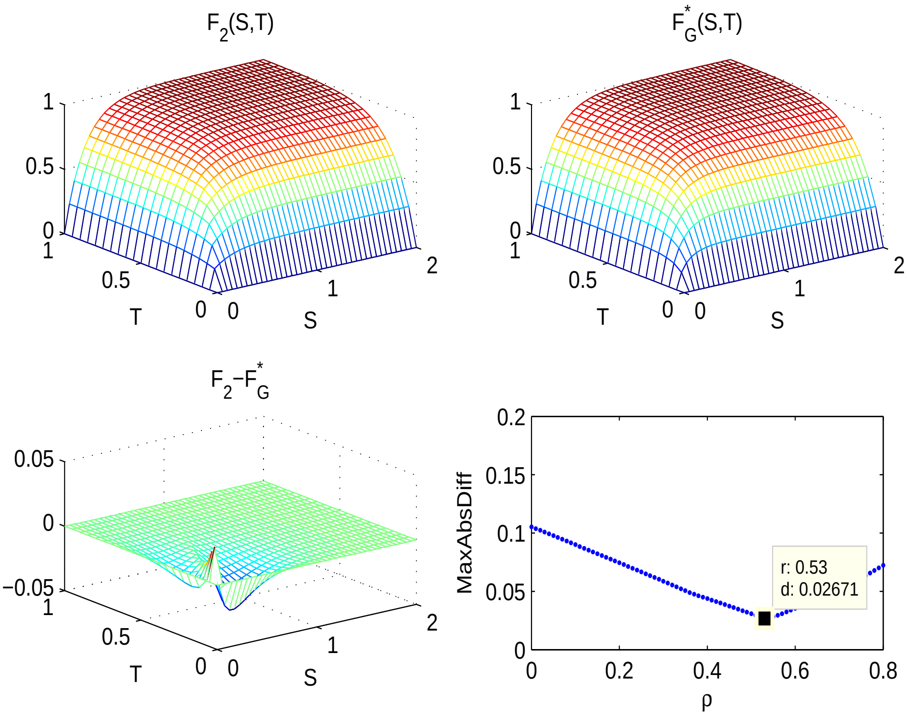

Figure 2: Comparison of both distribution functions, $F_2(s,t)$ and $F^*_\rho(s,t)$.

8.2 Comparison with Morgenstern's copula

In this example we just consider a copula of certain form, namely

$$C_a(u,v) = uv[1 + a(1-u)(1-v)],$$

with parameter $|a| \leq 1$, $u, v \in [0,1]^2$ (see also in Marshall,Olkin, 1967, p.30).

The copula is symmetric, $\mathrm{corr}(U,V) = a/3$,

$a = 0$ means independence of U, V.

We shall again try to find ρ yielding the closest gaussian copula

$$C^*_\rho(u,v) = \phi_2(\phi^{-1}(u), \phi^{-1}(v), \rho),$$

i.e. minimizing $\sup_{u,v} |C_a(u,v) - C^*_\rho(u,v)|$ (when parameter a is given).

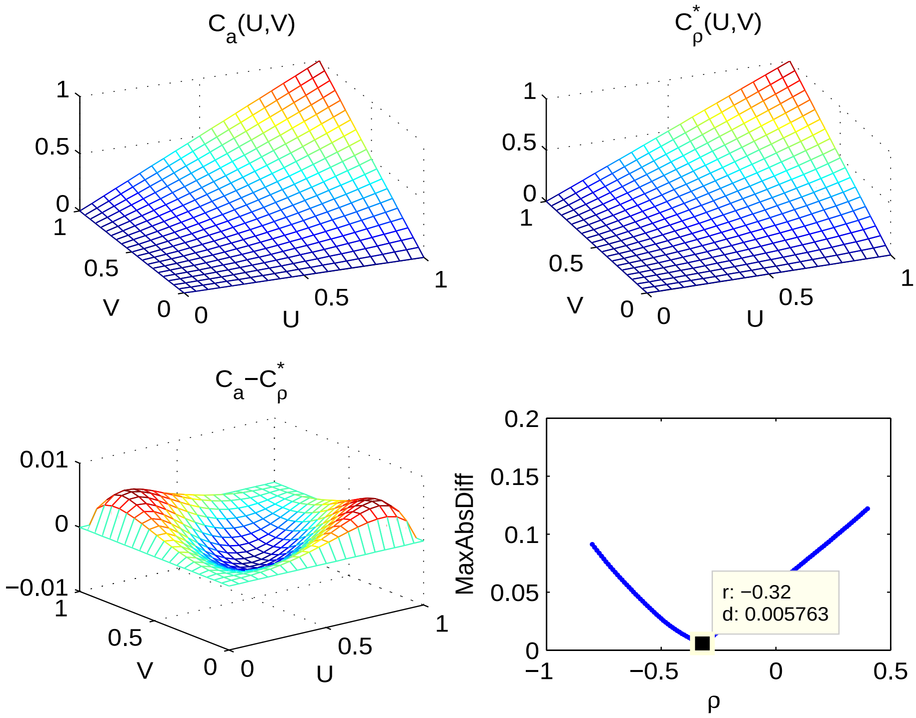

Figure 3: Comparison of $C_a(u,v)$ and $C^*_\rho(u,v)$, for $a = -0,9$, i.e. $\rho(U,V) = -1/3$.

Just for illustration, Figure shows 3d and contour plots of density of copula $C_a(u, v)$, for $a = -0.9$ (left) and $a = 0.5$ (right).

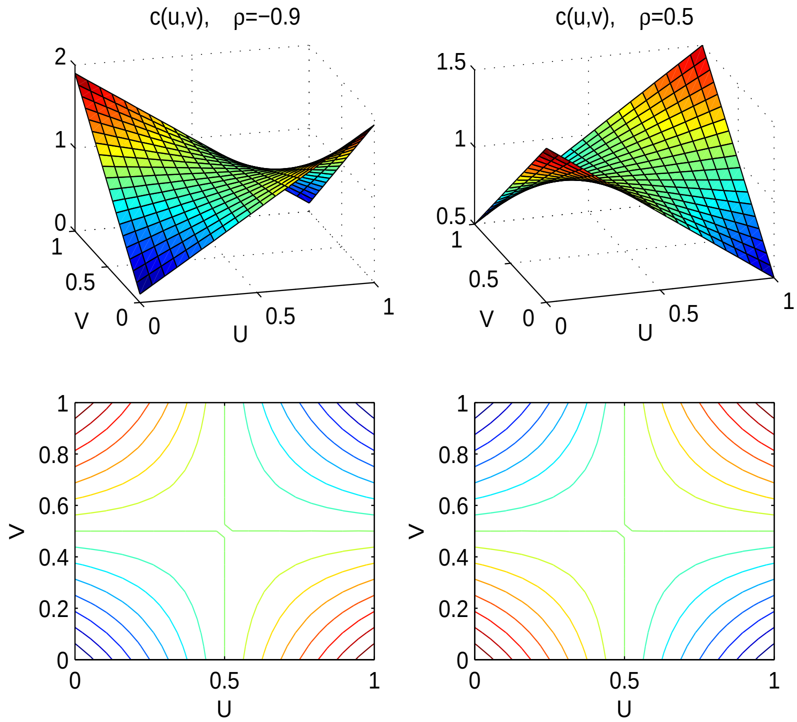

Figure 4: Density $c_a(u, v)$ of copula C_a, for $a = -0.9$ (left) and $a = 0.5$ (right).

References

[1] Heckman, J.J., Honoré, B.E.: The identifiability of the competing risks model, *Biometrika* **76**, 325–330, 1989.

[2] Marshall, A.W. and Olkin, I.: A multivariate exponential distribution, *J.A.S.A.* **62**, 30–44, 1967.

[3] Lin, D.Y.: Non-parametric inference for cumulative incidence functions in competing risks studies, *Statistics in Medicine* **16**, 901–910, 1997.

[4] Scheike, T.H., Zhang, M.: Flexible competing risks regression modelling and goodness-of-fit, *Lifetime Data Analysis* **14**, 464–483, 2008.

[5] Tsiatis, A.: A nonidentifiability aspects of the problem of competing risks, *Proc. Nat. Acad. Sci. USA* **72**, 20–22, 1975.

[6] Zheng, M., Klein, J.P.: Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika* **82**, 127–138, 1995.

9 Examples

In all examples I fixed: $\rho = 0.5$ or $\rho = -0.7$,

$S \sim$ Weibull $(a_s = 100,\ b_s = 1.2)$, $T \sim$ Weibull $(a_t = 130,\ b_t = 3)$, $C \sim |\text{Normal}(\mu = 150,\ \sigma = 50)|$.

The rate of censoring was among $10-20\,\%$. Weibull distribution function was taken in form $F(s) = 1 - \exp\left(-\left(\frac{s}{a}\right)^b\right),\ s > 0$.

Example 1 shows how normal variables (X, Y) transformed to (U, V) by (3) and then to (S, T) by Gauss copula (4),

a) 'empirically' i.e. with the aid of generated data,

b) numerically.

Numerically computed correlations yield

$\rho(U, V) = 0.432$, $\rho(S, T) = 0.376$ in the case $\rho(X, Y) = 0.5$,

$\rho(U, V) = -0.685$, $\rho(S, T) = -0.625$ in the case $\rho(X, Y) = -0.7$.

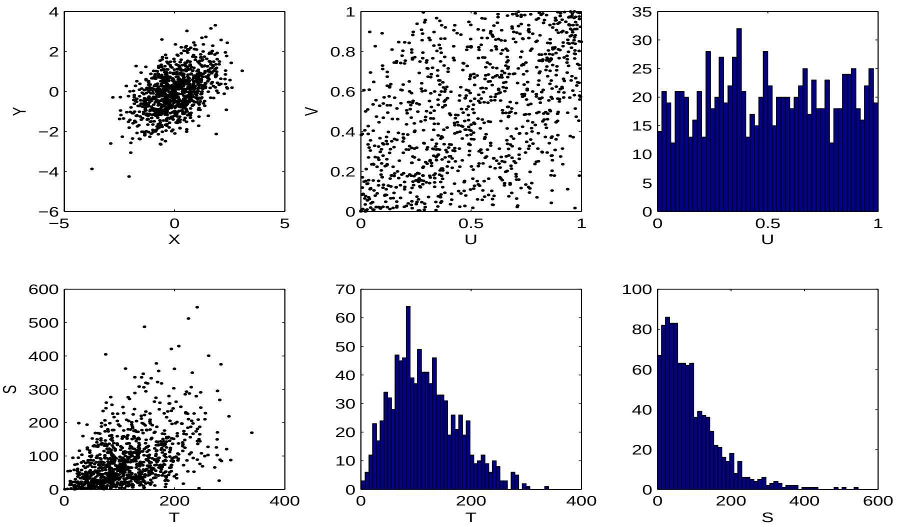

Figure 5: Scatter-plots and histograms of generated representation of X, Y, then transformed to U, V and S, T, the case with $\rho = 0.5$, $N = 1000$.

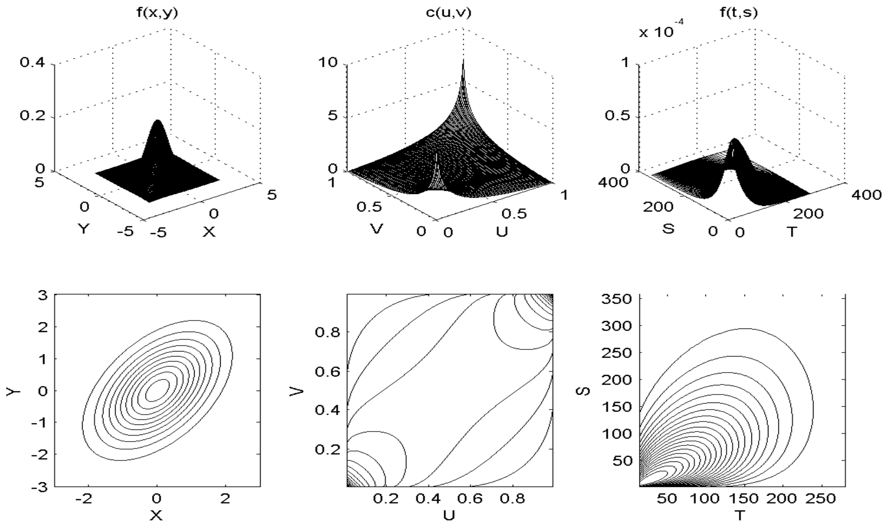

Figure 6: 3d plots and contours of density functions of joint distributions (X, Y), (U, V), (S, T), the case with $\rho = 0.5$.

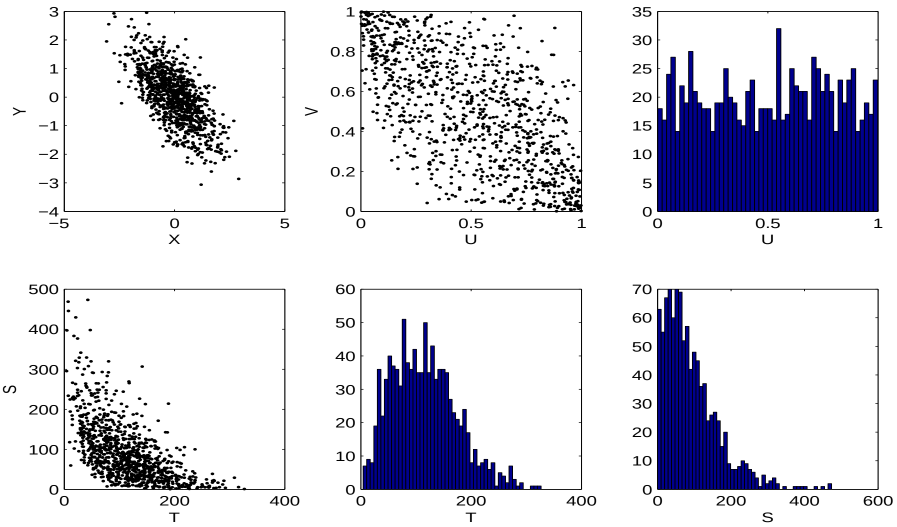

Figure 7: Scatter-plots and histograms of generated representation of X, Y, then transformed to U, V and S, T, the case with $\rho = -0.7$, $N = 1000$.

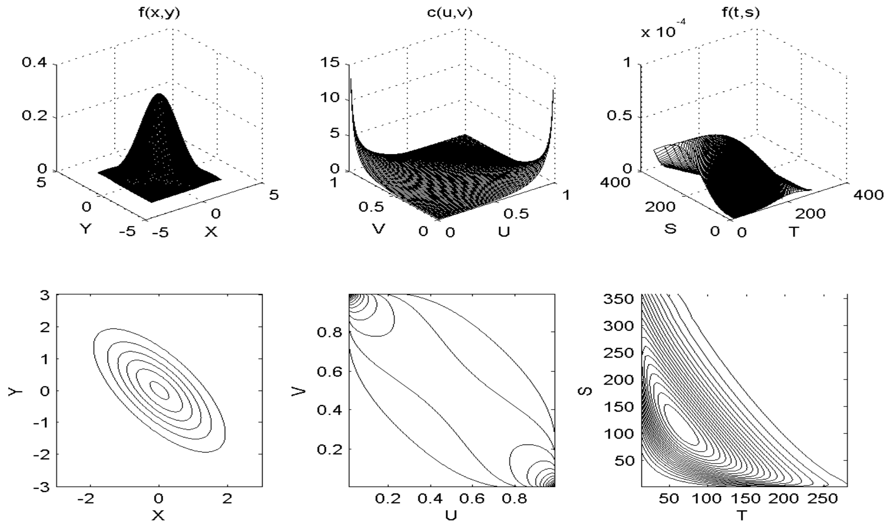

Figure 8: 3d plots and contours of density functions of joint distributions (X, Y), (U, V), (S, T), the case with $\rho = -0.7$.

Example 2: Estimation of ρ when F_S, F_T are known,

Randomized estimation:

ρ randomly proposed from $(-1, 1)$, randomly accepted along likelihood proportion (∼Metropolis algorithm).

Figure 5 shows data when $\rho = 0.5$.

Figure 6 shows 500 last accepted values ρ from 10 000 steps.

Figures 7 and 8 displays the same in the case $\rho = -0.7$.

In both cases size of data was $N = 200$.

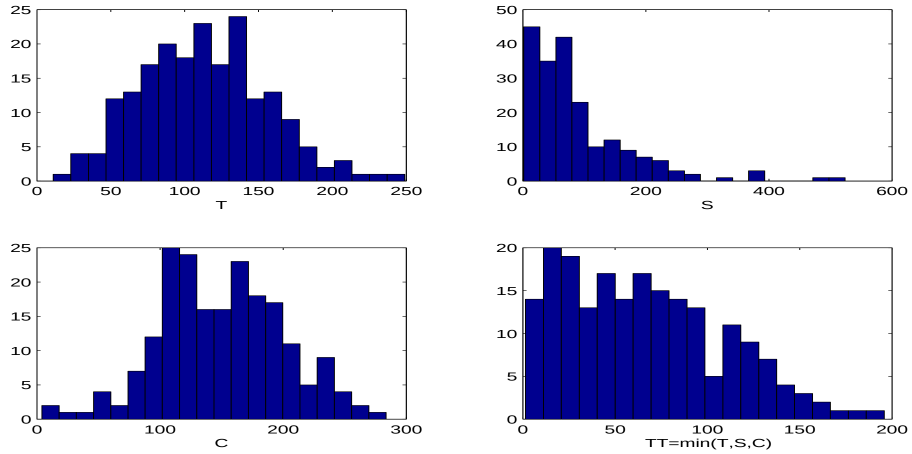

Figure 9: Generated data representing T, S, C and $Z = \min(T, S, C)$, with $\rho = 0.5$.

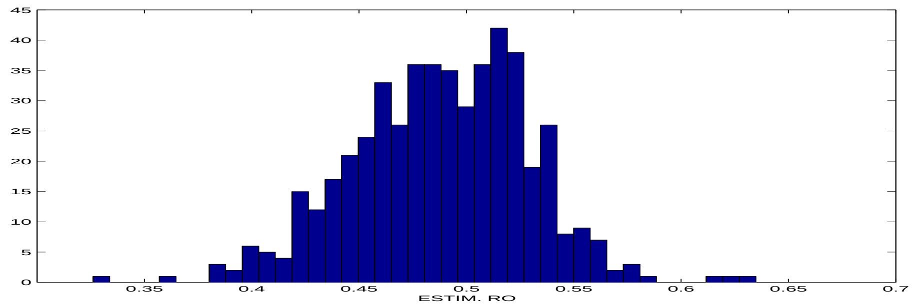

Figure 10: Sample of 500 last accepted estimates of ρ, when $\rho = 0.5$.

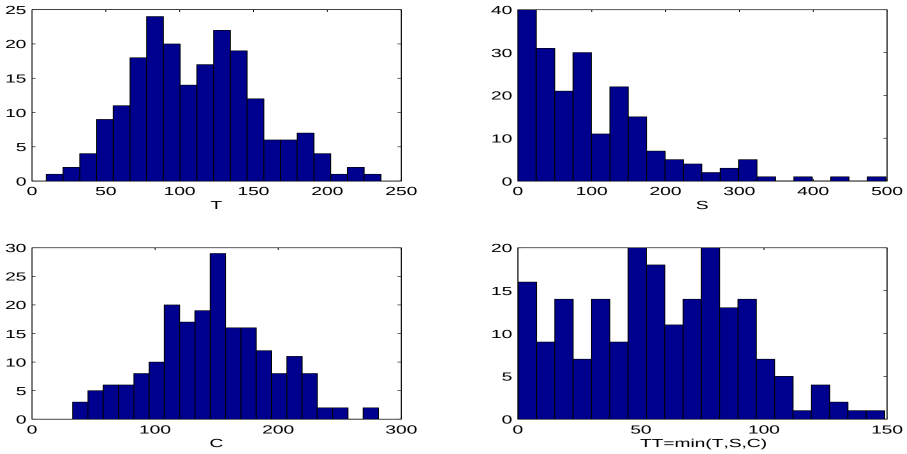

Figure 11: Generated data representing T, S, C and $Z = \min(T, S, C)$, with $\rho = -0.7$.

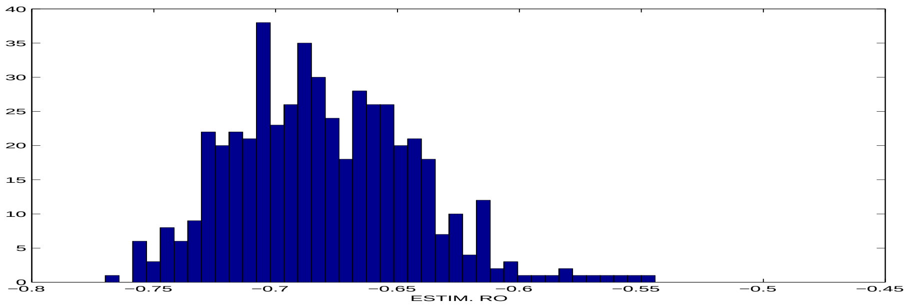

Figure 12: Sample of 500 last accepted estimates of ρ, when $\rho = -0.7$.

Example 3: It was assumed that ρ was known,

4 parameters of two Weibull distributions of S and T were estimated.

Again, a randomized search was used, via Metropolis–Hastings algorithm.

Figure 9 shows last 500 values from 10000 generated, of all four estimated parameters a_s, b_s of S, a_t, b_t of T, in the case when $\rho = 0.5$.

Figure 10 displays the results in the case with known $\rho = -0.7$.

Again, size of data was $N = 200$.

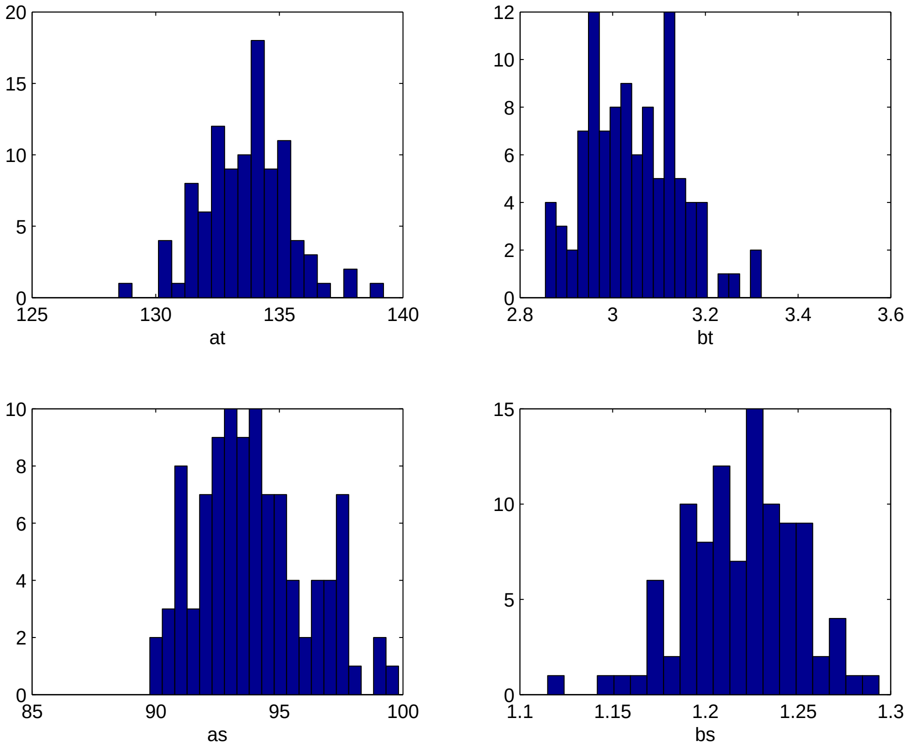

Figure 13: Sample of 500 last accepted estimates of Weibull parameters a_t, b_t, a_s, b_s, when $\rho = 0.5$.

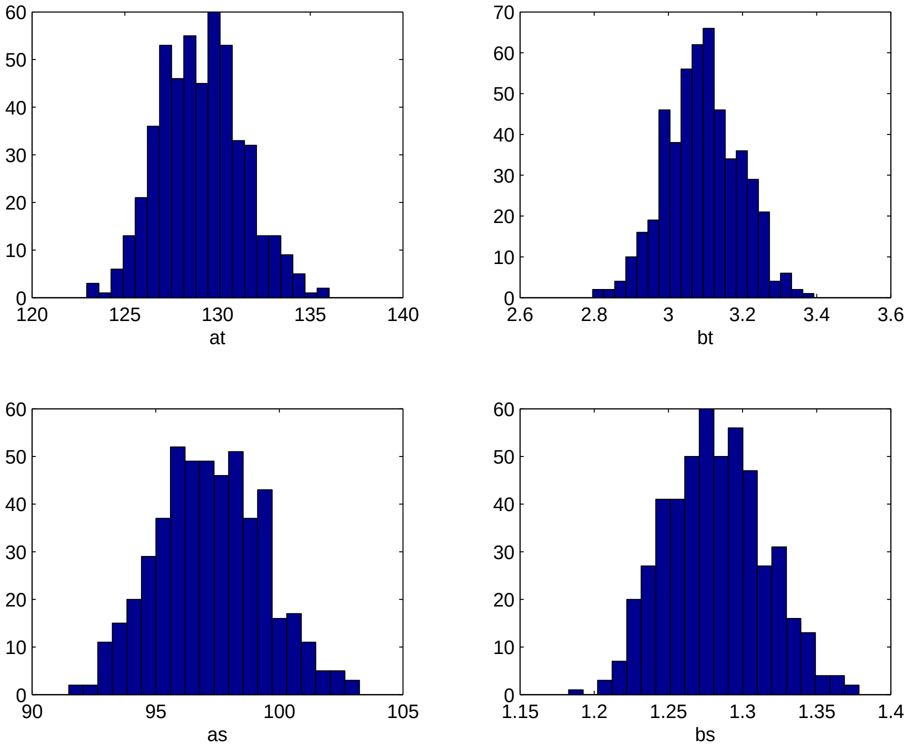

Figure 14: Sample of 500 last accepted estimates of Weibull parameters a_t, b_t, a_s, b_s, when $\rho = -0.7$.

Example 4: Non-parametric estimate of unknown distribution functions F_S and F_T assuming that ρ is known,.

– as a mixture of Gauss c.d.f-s.

After some experiments, both mixtures were composed from 7 components,

their centers and variances were adapted during the estimation.

It was again iterative, with acceptance criterion of Metropolis–Hastings algorithm,

Figures show just the best result, i.e. with maximal achieved value of likelihood, from 10000 iterations.

Now the length of data was $N = 500$.

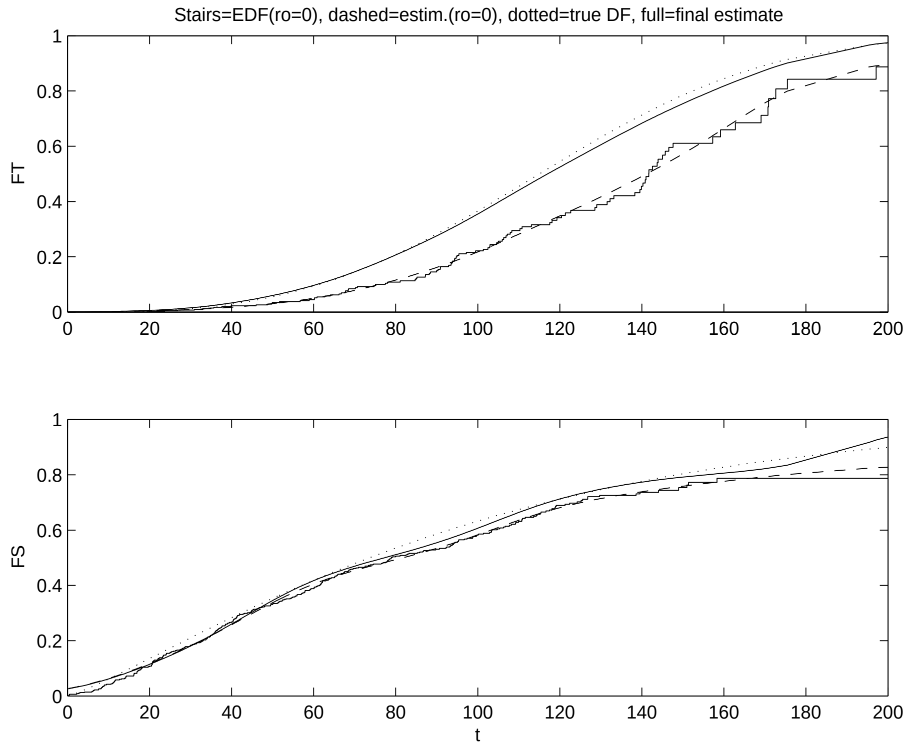

Figure 15: Estimated distribution functions: Initial estimates PLE (stepwise) and smoothed (dashed), then final estimate (smooth full) and 'true' distribution function (dotted). F_T above, F_S below, the case with $\rho = 0.5$.

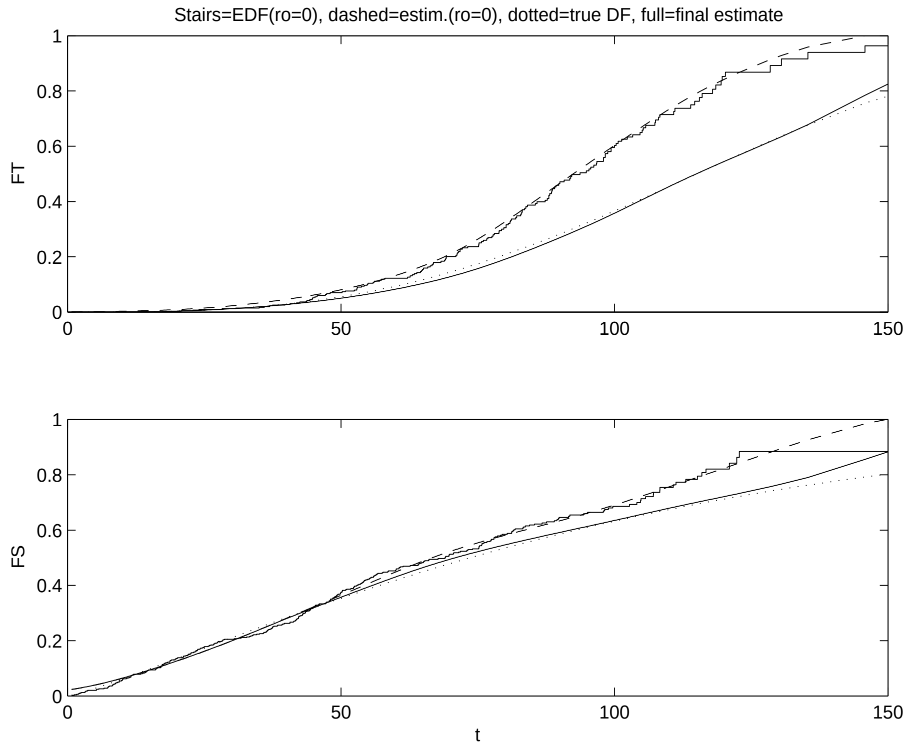

Figure 16: The same as in Figure 11, the case with $\rho = -0.7$.

Example of incidence function: – estimate of cumul. inc. fction

The same type of data as in previous examples was generated.

We display here just the case of $\rho = 0.5, N = 200$.

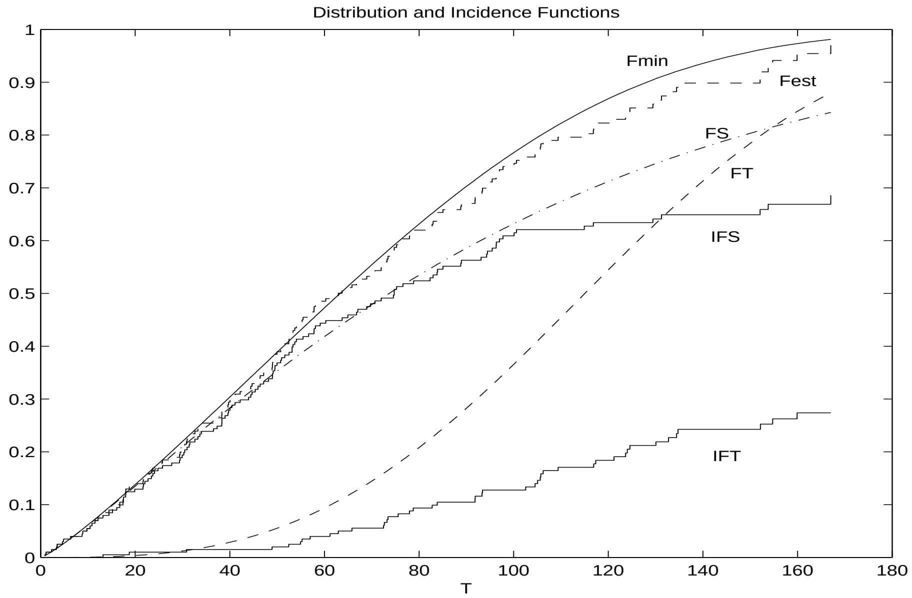

Figure 17: 'True' distribution functions F_S, F_T (dashed), F_{min} of $\min(S,T)$, its PLE F_{est}, estimated cumulative incidence functions IF_S, IF_T.

'True' cumulative incidence functions can be obtained by (numerical) integration of expressions corresponding to the 1-st and 2-nd part of the likelihood,

$$dIF_S(t) = f_S(t)\left[1 - \phi_1(y; \rho x, 1-\rho^2)\right], \;\; dIF_T(t) = f_T(t)\left[1 - \phi_1(x; \rho y, 1-\rho^2)\right],$$

where again $x = \phi^{-1}(F_S(t))$, $y = \phi^{-1}(F_T(t))$.

Figure 14 compares them with their estimates:

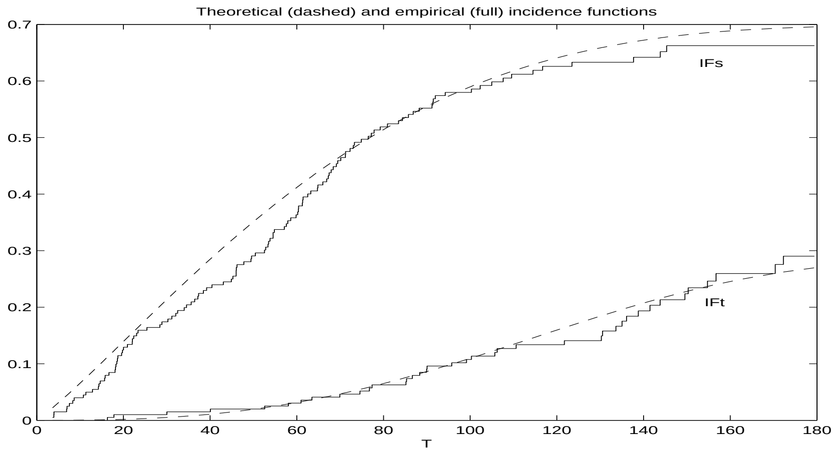

Figure 18: 'True' and estimated (stepwise) cumulative incidence functions IF_S, IF_T.

References

[1] Heckman, J.J., Honoré, B.E.: The identifiability of the competing risks model, *Biometrika* **76**, 325–330, 1989.

[2] Marshall, A.W. and Olkin, I.: A multivariate exponential distribution, *J.A.S.A.* **62**, 30–44, 1967.

[3] Lin, D.Y.: Non-parametric inference for cumulative incidence fumctions in competing risks studies, *Statistics in Medicine* **16**, 901–910, 1997.

[4] Scheike, T.H., Zhang, M.: Flexible competing risks regression modelling and goodness-of-fit, *Lifetime Data Analysis* **14**, 464–483, 2008.

[5] Tsiatis, A.: A nonidentifiability aspects of the problem of competing risks, *Proc. Nat. Acad. Sci. USA* **72**, 20–22, 1975.

[6] Zheng, M., Klein, J.P.: Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika* **82**, 127–138, 1995.