NONPARAMETRIC ESTIMATION OF CONDITIONAL RESIDUAL QUANTILES

Noël Veraverbeke (PART 1)
Paul Janssen (PART 2)

Hasselt University - CENSTAT
B-3590 DIEPENBEEK - BELGIUM

June 2022

Congratulations to **INGRID VAN KEILEGOM**

15 May 1998

RESIDUAL LIFETIME

T_1 : lifetime of patient or device
Suppose that we know that $T_1 > t_1$, then $T_1 - t_1$ is called the **residual lifetime** of T_1 and

$$P(T_1 - t_1 \leq y \mid T_1 > t_1) \tag{1}$$

is called the **residual lifetime distribution**.
Many authors studied

- the **mean** residual lifetime:

$$E(T_1 - t_1 \mid T_1 > t_1)$$

- the **quantiles** of the residual lifetime (e.g. median) i.e. the inverse of the residual lifetime distribution.

CONDITIONAL RESIDUAL LIFETIME

T_1 : lifetime as before
T_2 : some other variable (containing extra information on T_1)
We generalize (1) by adding an **extra conditioning** of the form $\{T_2 \leq t_2\}$ or $\{T_2 > t_2\}$:

$$P(T_1 - t_1 \leq y \mid T_1 > t_1, T_2 \leq t_2\} \quad (2)$$

$$P(T_1 - t_1 \leq y \mid T_1 > t_1, T_2 > t_2\} \quad (3)$$

are **conditional residual lifetime distributions.**
In this talk we discuss **quantiles** of (2) and (3):
For $0 < p < 1$:

$$\widetilde{Q}(p \mid t_1, t_2) = p\text{-th quantile of (2)}$$

$$\widetilde{\widetilde{Q}}(p \mid t_1, t_2) = p\text{-th quantile of (3)}$$

OUR MOTIVATION

Comes from our interest in good **risk ratios** for comparison of risks between 2 groups.
See recent papers of **Abrams, Janssen, Veraverbeke**
Such risk ratios are mostly defined in terms of **conditional hazard rate functions**. For example

- $CR(t_1, t_2) = \frac{\lambda(t_1|T_2=t_2)}{\lambda(t_1|T_2>t_2)}$ **cross ratio** of **Clayton (1978)**
- $RR(t_1, t_2) = \frac{\lambda(t_1|T_2\geq t_2)}{\lambda(t_1|T_2<t_2)}$ **risk ratio**
- or also $\frac{\lambda(t_1|t_{21}<T_2\leq t_{22})}{\lambda(t_1|t_{23}<T_2\leq t_{24})}$

The λ's are conditional hazard rate functions of T_1 at t_1, given that $T_2 = t_2$ (or $T_2 \geq t_2$, or $T_1 < t_2, \ldots$).

We want to replace the conditional hazard rate functions by quantile functions of the conditional residual lifetime:

$$\frac{\widetilde{\widetilde{Q}}(p \mid t_1, t_2)}{\widetilde{Q}(p \mid t_1, t_2)}$$

Reason: ratios of quantiles of conditional residual lifetime are easier to interpret than ratios of conditional hazard rates.

T_1 censored, T_2 not censored

We study nonparametric estimators for $\widetilde{Q}(p \mid t_1, t_2)$ and $\widetilde{\widetilde{Q}}(p \mid t_1, t_2)$ and establish asymptotic normality results.
We allow that the lifetime T_1 is subject to **random right censoring** by a **censoring variable** C. The variable T_2 is always observed.
So the available data are

$$(Z_i, \delta_i, T_{2i}) \quad i = 1, \ldots, n$$

a random sample from (Z, δ, T_2) where

$$Z = T_1 \wedge C \quad \delta = I(T_1 \leq C)$$

Notation: joint distribution function of (T_1, T_2):

$$F(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2)$$

and its margins

$$F_1(t_1) = P(T_1 \leq t_1) \quad F_2(t_2) = P(T_2 \leq t_2)$$

QUANTILE FUNCTIONS $\widetilde{Q}$ and $\widetilde{\widetilde{Q}}$

Conditional distribution of T_1, given $T_2 \le t_2$:

$$\begin{aligned} F(t_1 \mid T_2 \le t_2) &= P(T_1 \le t_1 \mid T_2 \le t_2) \\ &= \frac{F(t_1, t_2)}{F_2(t_2)} \equiv \widetilde{F}_{t_2}(t_1) \end{aligned}$$

Conditional residual lifetime distribution of T_1 at t_1 given that $T_2 \le t_2$:

$$P(T_1 - t_1 \le y \mid T_1 > t_1, T_2 \le t_2)$$

$$= \frac{P(t_1 < T_1 \le t_1 + y, T_2 \le t_2)}{P(T_1 > t_1, T_2 \le t_2)}$$

$$= \frac{\widetilde{F}_{t_2}(t_1 + y) - \widetilde{F}_{t_2}(t_1)}{1 - \widetilde{F}_{t_2}(t_1)}$$

p-th quantile of this distribution function:

$$\begin{aligned}\widetilde{Q}(p \mid t_1, t_2) &= \inf\left\{y : \frac{\widetilde{F}_{t_2}(t_1 + y) - \widetilde{F}_{t_2}(t_1)}{1 - \widetilde{F}_{t_2}(t_1)} \geq p\right\} \\ &= \inf\{y : \widetilde{F}_{t_2}(t_1 + y) \geq \widetilde{F}_{t_2}(t_1) + p(1 - \widetilde{F}_{t_2}(t_1))\} \\ &= -t_1 + \widetilde{F}_{t_2}^{-1}(p + (1 - p)\widetilde{F}_{t_2}(t_1))\end{aligned}$$

where $\widetilde{F}_{t_2}(p) = \inf\{y : \widetilde{F}_{t_2}(y) \geq p\}$ is the inverse of $\widetilde{F}_{t_2}$

A similar expression holds for $\widetilde{\widetilde{Q}}(p \mid t_1, t_2)$ starting from

$$\begin{aligned}F(t_1 \mid T_2 > t_2) &= P(T_1 \leq t_1 \mid T_2 > t_2) \\ &= \frac{F_1(t_1) - F(t_1, t_2)}{1 - F_2(t_2)} \equiv \widetilde{\widetilde{F}}_{t_2}(t_1)\end{aligned}$$

ESTIMATION OF THE JOINT DISTRIBUTION $F(t_1, t_2)$

Start from the relation

$$F(t_1, t_2) = \int_0^{t_2} F(t_1 \mid t) dF_2(t)$$

and plug in estimators

- $F_n(t_1 \mid t)$ for $F(t_1 \mid t)$
- $F_{2n}(t_2) = \frac{1}{n} \sum_{i=1}^{n} I(T_{2i} \leq t_2)$ for $F_2(t_2)$

This idea has been worked out in **Akritas (1994)** and **Akritas and Van Keilegom (2003)**.

The estimator for $F(t_1 \mid t)$ is the **Beran** estimator (or conditional Kaplan-Meier estimator):

$$F_n(t_1 \mid t) = 1 - \prod_{Z_{(i)} \le t_1} \left(1 - \frac{w_{n(i)}(t, h_n)}{\sum_{j=1}^{n} w_{n(j)}(t, h_n) I(Z_j \ge Z_i)} \right)^{\delta_{(i)}}$$

where $Z_{(1)} \le Z_{(2)} \le \ldots \le Z_{(n)}$ are the ordered Z_j-values and $\delta_{(j)}$ is the censoring indicator for $Z_{(j)}$.

The weights are Nadaraya-Watson weights

$$w_{ni}(t, h_n) = \frac{K\left(\frac{t - T_{2i}}{h_n}\right)}{\sum_{j=1}^{n} K\left(\frac{t - T_{2j}}{h_n}\right)}$$

(if the weights are equal to $\frac{1}{n}$ then Beran = Kaplan-Meier).

REPRESENTATION FOR $F_n(t_1, t_2)$

$$F_n(t_1, t_2) = \int_0^{t_1} F_n(t_1 \mid t) dF_{2n}(t)$$

As **Lo and Singh (1986)** did for the Kaplan-Meier estimator, an almost sure asymptotic representation has been proved for the Beran estimator:

$$F_n(t_1 \mid T_{2i}) - F(t_1 \mid T_{2i}) \approx \sum_{j=1}^{n} w_{nj}(T_{2i}, h_n)\xi(t_1, Z_j, \delta_j, T_{2i})$$

where

$$\xi(t, Z, \delta, t) = (1 - F(t_1 \mid t)) \left\{ - \int_0^{Z \wedge t_1} \frac{dH^u(s \mid t)}{(1 - H(s \mid t))^2} + \frac{I(Z \leq t_1, \delta = 1)}{1 - H(Z \mid t)} \right\}$$

Here

$$
\begin{aligned}
H(z \mid t) &= P(Z \leq z \mid T_2 = t) \\
H^u(z \mid t) &= P(Z \leq z, \delta = 1 \mid T_2 = t)
\end{aligned}
$$

We have the well known relations:
(if T_1 and C are independent, given T_2)

$$
\begin{aligned}
1 - H(z \mid t) &= (1 - F(z \mid t))(1 - G(z \mid t)) \\
H^u(z \mid t) &= \int_0^z (1 - G(s- \mid t)) dF(s \mid t)
\end{aligned}
$$

where

$$
\begin{aligned}
F(z \mid t) &= P(T_1 \leq z \mid T_2 = t) \\
G(z \mid t) &= P(C \leq z \mid T_2 = t)
\end{aligned}
$$

The representation is valid for $(t_1, t_2) \in \Omega$, a domain which essentially says that we have to stay away from the upper endpoints of the supports of $H(z \mid t)$ and $F_2(t)$.

REGULARITY ASSUMPTIONS

(A1) $\frac{\log n}{nh_n} \to 0$, $nh_n^4 \to 0$
K is a probability density function with support $[-1, 1]$, twice differentiable, $\int uK(u)du = 0$.

(A2) $F_2(t_2)$ is three times continuously differentiable w.r.t. t_2;
$H(z \mid t_2)$ and $H^u(z \mid t_2)$ are twice continuously differentiable w.r.t. z and t_2 and for $(z, t_2) \in \Omega$, all derivatives are uniformly bounded

Theorem
Under (A1) and (A2) we have

$$F_n(t_1, t_2) = F(t_1, t_2) + \frac{1}{n} \sum_{i=1}^{n} \psi(t_1, Z_i, \delta_i, T_{2i}) + r_n(t_1, t_2)$$

with

$$\psi(t_1, Z_i, \delta_i, T_{2i}) = [F(t_1 \mid T_{2i}) I(T_{2i} \leq t_2) - F(t_1, t_2)] \\ + \xi(t_1, Z_i, \delta_i, T_{2i}) I(T_{2i} \leq t_2)$$

and

$$\sup_{(t_1, t_2) \in \Omega} \mid r_n(t_1, t_2) \mid = o_P(n^{-1/2})$$

Also, as $n \to \infty$,

$$n^{1/2}(F_n(t_1,t_2) - F(t_1,t_2) \xrightarrow{d} N(0, \sigma^2(t_1,t_2))$$

where

$$\begin{aligned}\sigma^2(t_1,t_2) &= \int_0^{t_2} F^2(t_1 \mid t) dF_2(t) - F^2(t_1,t_2) \\ &\quad + \int_0^{t_2} (1 - F(t_1 \mid t))^2 \left\{ \int_0^{t_1} \frac{dH^u(s \mid t)}{(1 - H(s \mid t))^2} \right\} dF_2(t)\end{aligned}$$

ESTIMATION OF $\widetilde{F}_{t_2}(t_1)$ and $\widetilde{\widetilde{F}}_{t_2}(t_1)$

$$\widetilde{F}_{t_2,n}(t_1) = \frac{F_n(t_1,t_2)}{F_{2n}(t_2)}$$

$$\widetilde{\widetilde{F}}_{t_2,n}(t_1) = \frac{F_n(t_1,+\infty) - F_n(t_1,t_2)}{1 - F_{2n}(t_2)}$$

Linearization + Slutsky's theorem leads to asymptotic representations

$$\widetilde{F}_{t_2,n}(t_1) - \widetilde{F}_{t_2}(t_1) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\psi}(t_1, Z_i, \delta_i, T_{2i}) + rem$$

$$\widetilde{\widetilde{F}}_{t_2,n}(t_1) - \widetilde{\widetilde{F}}_{t_2}(t_1) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\widetilde{\psi}}(t_1, Z_i, \delta_i, T_{2i}) + rem$$

Direct (but long) calculations give the covariance functions of

$$E(\widetilde{\psi}(t_1, \ldots)\widetilde{\psi}(t_1', \ldots))$$
$$E(\widetilde{\widetilde{\psi}}(t_1, \ldots)\widetilde{\widetilde{\psi}}(t_1, \ldots))$$

ESTIMATION OF THE QUANTILE FUNCTIONS

Natural plug-in estimators

$$\widetilde{Q}_n(p \mid t_1, t_2) = -t_1 + \widetilde{F}_{t_2,n}^{-1}(p + (1-p)\widetilde{F}_{t_2,n}(t_1))$$

$$\widetilde{\widetilde{Q}}_n(p \mid t_1, t_2) = -t_1 + \widetilde{\widetilde{F}}_{t_2,n}^{-1}(p + (1-p)\widetilde{\widetilde{F}}_{t_2,n}(t_1))$$

Also here we obtain asymptotic representations via **Bahadur** type theorems.
(extra complication: random arguments)

The final result is asymptotic normality for $\widetilde{Q}_n$ and $\widetilde{\widetilde{Q}}_n$ with **explicit expressions** for the asymptotic variances:

$$n^{1/2}(\widetilde{Q}_n(p \mid t_1, t_2) - \widetilde{Q}(p \mid t_1, t_2)) \quad \overset{d}{\to} \quad N(0; \widetilde{\sigma}_p^2(t_1, t_2))$$

$$n^{1/2}(\widetilde{\widetilde{Q}}_n(p \mid t_1, t_2) - \widetilde{\widetilde{Q}}(p \mid t_1, t_2)) \quad \overset{d}{\to} \quad N(0; \widetilde{\widetilde{\sigma}}_p^2(t_1, t_2))$$

For the precise expressions for $\widetilde{\sigma}_p^2$ and $\widetilde{\widetilde{\sigma}}_p^2$, we refer to **Abrams, Janssen, Veraverbeke (2021, Statistics)**

Paul Janssen (PART 2)

CALCULATING THE RESIDUAL QUANTILE

Remember

$$\widetilde{Q}(p \mid t_1, t_2) = \inf\left\{ y : \frac{\widetilde{F}_{t_2}(t_1 + y) - \widetilde{F}_{t_2}(t_1)}{1 - \widetilde{F}_{t_2}(t_1)} \geq p \right\}$$

with

$$\widetilde{F}_{t_2}(t_1) = P(T_1 \leq t_1 \mid T_2 \leq t_2) = \frac{F(t_1, t_2)}{F_2(t_2)}.$$

To solve

$$\widetilde{F}_{t_2}(t_1 + y) - \widetilde{F}_{t_2}(t_1) = p(1 - \widetilde{F}_{t_2}(t_1))$$
$$\Leftrightarrow$$
$$F(t_1 + y, t_2) = pF_2(t_2) + (1 - p)F(t_1, t_2)$$
$$\Leftrightarrow \text{(Sklar)}$$
$$C(u(y), v) = pv + (1 - p)C(u(0), v)$$

with $v = F_2(t_2)$, $u(y) = F_1(t_1 + y)$.

SIMULATIONS

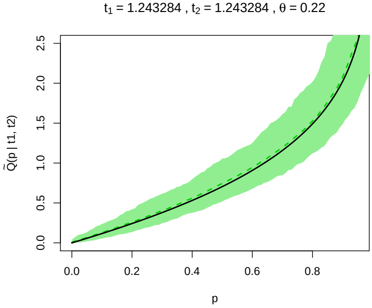
	κ	λ_L	λ_U	$\kappa(\theta)$
Clayton	$\frac{\theta}{\theta+2}$	$2^{-1/\theta}$	0	0.10 (0.22)
				0.20 (0.50)
				0.50 (2.00)
Gumbel	$\frac{\theta-1}{\theta}$	0	$2-2^{1/\theta}$	0.10 (1.11)
				0.20 (1.25)
				0.50 (2.00)
FGM*	$\frac{4\theta}{18}$	0	0	0.10 (0.45)
				0.20 (0.90)

$*$ For $0 \leq \theta \leq 1$, $\kappa \in [0, 2/9]$.

$$\widetilde{Q}(p \mid t_1, t_2) = -t_1 + F_1^{-1}(g(p \mid t_1, t_2, \theta))$$

and $g(p \mid t_1, t_2, \theta)$ a function determined by the specific copula under consideration

- $F_j(t_j) = 1 - \exp(-d_j t_j^{s_j})$, $j = 1, 2$
  $s_j = 1.5$ (shape), $d_j = 0.5$ (decay), scale: d_j^{-1/s_j}
  median $= d_j^{-1/s_j}(\ln 2)^{1/s_j} \approx 1.243$
- $G(t) = 1 - \text{exp}(-d_c t^{s_c})$
  $s_c = 1.5$, $d_c = 0.15$ (23% censoring)
  0.85 (63% censoring)


t_1 = 1.243284 , t_2 = 1.243284 , θ = 0.22
$\tilde{Q}$(p | t1, t2)
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
p
t_1 = 1.243284 , t_2 = 1.243284 , θ = 0.22
$\tilde{Q}$(p | t1, t2)
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
p

DATA ON PRIMARY BILIARY LIVER CIRRHOSIS

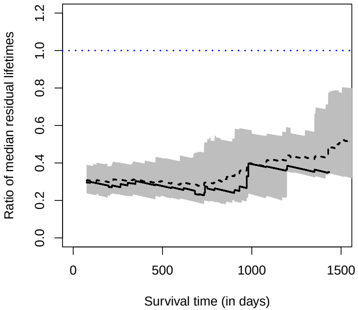
PBC-data (Mayo Clinic)
patients: 424 (55% censoring)
primary endpoint: survival time

two groups based on serum bilirubin level
(related to functioning of liver, high level is bad)

group-low : level $\leq$ 3.4 mg/dl (75%)
group-high: level $>$ 3.4 mg/dl (25%)

The figure below gives the ratio of the medians of the conditional lifetime distributions as a function of t_1.

$$\frac{\widetilde{\widetilde{Q}}(1/2 \mid T_1 > t_1, T_2 > 3.4)}{\widetilde{Q}(1/2 \mid T_1 > t_1, T_2 \leq 3.4)}$$



Note. The 95% pointwise confidence limits are bootstrap based (100 resamples).

CONDITIONAL RESIDUAL LIFETIME - EXTENSION

Remember

$$P(T_1 - t_1 \leq y \mid T_1 > t_1, T_2 \leq t_2)$$

$$P(T_1 - t_1 \leq y \mid T_1 > t_1, T_2 > t_2)$$

Extension

$$P(T_1 - t_1 \leq y \mid T_1 > t_1, t_{21} < T_2 \leq t_{22})$$

Motivation 1 (bilirubin level)

It allows (based on the value of T_2) a more flexible partitioning of the population. Hence, more flexible risk ratios can be considered (think about partitioning of the population based on the bilirubin level).

Motivation 2 (diabetic retinopathy)

Diabetic retinopathy: an eye disease that is caused by high blood sugar and high blood pressure. Over time, these conditions damage the blood vessels in the back of the eye.

Consider the subpopulation: $t_{21} < T_2 \leq t_{22}$, then $t_1 = t_{22}$ is an interesting (medical) choice. Assume that for eye 2 we know that "time to blindness" is in the interval $]t_{21}, t_{22}]$ and that, at time t_{22}, eye 1 is not yet blind, we then look at the conditional distribution

$$P(T_1 - t_{22} \leq y \mid T_1 > t_{22}, t_{21} < T_2 \leq t_{22})$$

It is the distribution of the remaining lifetime for eye 1 starting from the time that eye 2 is blind. Interesting conditional residual quantiles to look at are de deciles, i.e., $p = 0.1, 0.2, ..., 0.9$ (as far as meaningful given the presence of censoring).

$$P(T_1 - t_1 \leq y \mid T_1 > t_1, t_{21} < T_2 \leq T_{22}) = \frac{\widetilde{F}(t_1 + y \mid t_{21}, t_{22}) - \widetilde{F}(t_1 \mid t_{21}, t_{22})}{1 - \widetilde{F}(t_1 \mid t_{21}, t_{22})}$$

with

$$\widetilde{F}(t_1 \mid t_{21}, t_{22}) = F(t_1 \mid t_{21} < T_2 \leq t_{22}) = \frac{F(t_1, t_{22}) - F(t_1, t_{21})}{F_2(t_{22}) - F_2(t_{21})} = \frac{F_V(t_1 \mid t_{21}, t_{22})}{F_{2V}(t_{21}, t_{22})}$$

$$\widetilde{Q}(p \mid t_1, t_{21}, t_{22}) = \inf \left\{ y : \frac{\widetilde{F}(t_1 + y \mid t_{21}, t_{22}) - \widetilde{F}(t_1 \mid t_{21}, t_{22})}{1 - \widetilde{F}(t_1 \mid t_{21}, t_{22})} \geq p \right\}$$

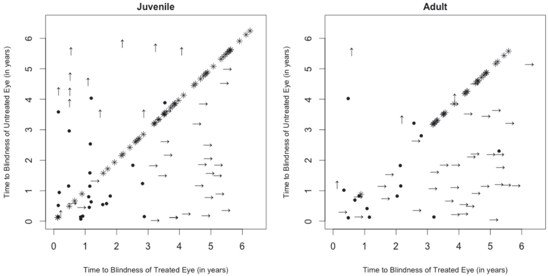
UNIVARIATE CENSORING-EXTENSION

$$D_i = (Z_{1i}, Z_{2i}, \delta_{1i}, \delta_{2i}), i = 1, \ldots, n$$

$$Z_{1i} = T_{1i} \wedge C_i, \delta_{1i} = \mathbf{1}(T_{1i} \leq C_i)$$

$$Z_{2i} = T_{2i} \wedge C_i, \delta_{2i} = \mathbf{1}(T_{2i} \leq C_i)$$

e.g. patients dropping out in a clinical study (diabetic retinopathy)



ESTIMATION OF THE JOINT DISTRIBUTION-REVISITED

Burke (1988) - case of univariate censoring

$$\begin{aligned} H(t_1, t_2) &= P(Z_1 \leq t_1, Z_2 \leq t_2, \delta_1 = 1, \delta_2 = 1) \\ &= P(T_1 \leq t_1, T_2 \leq t_2, T_1 \vee T_2 \leq C) \\ &= \int_0^{t_1} \int_0^{t_2} (1 - G[(z_1 \vee z_2)-]) F(dz_1, dz_2) \end{aligned}$$

$$
\begin{aligned}
F(t_1, t_2) &= \int_0^{t_1}\int_0^{t_2} \frac{1}{1 - G[(z_1 \vee z_2)-]} H(dz_1, dz_2) \\
\widehat{F}_B(t_1, t_2) &= \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_{1i}\delta_{2i}}{1 - \widehat{G}[(Z_{1i} \vee Z_{2i})-]} \mathbf{1}(Z_{1i} \le t_1, Z_{2i} \le t_2)
\end{aligned}
$$

with $\widehat{G}$ the KM based on $(T_{1i} \vee T_{2i}) \wedge C_i$, $\delta_i = 1 - \delta_{1i}\delta_{2i}$. $\widehat{F}_2$ is the Kaplan-Meier estimator based on $T_{2i} \wedge C_i$, δ_{2i}.

Lin & Ying (1993) - case of univariate censoring

$$P(Z_1 > t_1, Z_2 > t_2) = P(T_1 > t_1, T_2 > t_2, C > t_1 \vee t_2)$$
$$\overset{(T_1,T_2) \coprod C}{=} P(T_1 > t_1, T_2 > t_2) P(C > t_1 \vee t_2)$$

$$S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2) = \frac{P(Z_1 > t_1, Z_2 > t_2)}{1 - G(t_1 \vee t_2)}$$

$$F_V(t_1 \mid t_{21}, t_{22}) = F(t_1, t_{22}) - F(t_1, t_2)$$
$$= (S(t_1, t_{22}) - S(t_1, t_{21})) - (S_2(t_{22}) - S_2(t_{21}))$$

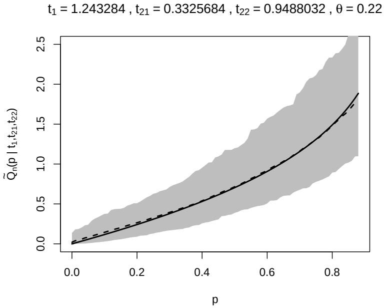
$$\widehat{S}(t_1, t_2) = \frac{1}{n} \frac{1}{1 - \widehat{G}(t_1 \vee t_2)} \sum \mathbf{1}(Z_{1i} > t_1, Z_{2i} > t_2)$$

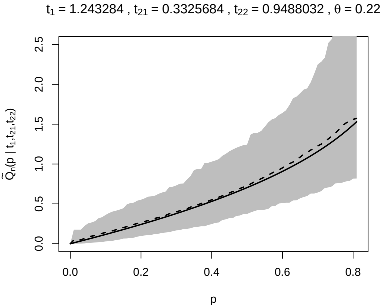
To estimate $S_2(t_{22}) - S_2(t_{21})$, use the KM estimator for S_2 (as before).

$$\widehat{F}_{V,LY}(t_1 \mid t_{21}, t_{22}) = \dots$$

$$
\begin{aligned}
\widehat{F}_{V,LY}(t_1 \mid t_{21}, t_{22}) = \tfrac{1}{n} \Bigg\{ & \tfrac{1}{1-\widehat{G}(t_1 \vee t_{22})} \sum_{i=1}^{n} \mathbf{1}(Z_{1i} > t_1, Z_{2i} > t_{22}) \\
& - \tfrac{1}{1-\widehat{G}(t_1 \vee t_{21})} \sum_{i=1}^{n} \mathbf{1}(Z_{1i} > t_1, Z_{2i} > t_{21}) \Bigg\} \\
- \tfrac{1}{n} \Bigg\{ & \tfrac{1}{1-\widehat{G}(t_1 \vee t_{22})} \sum_{i=1}^{n} \mathbf{1}(Z_{1i} > t_1, Z_{2i} > t_{22}) \\
& - \tfrac{1}{1-\widehat{G}(t_1 \vee t_{21})} \sum_{i=1}^{n} \mathbf{1}(Z_{1i} > t_1, Z_{2i} > t_{21}) \Bigg\}
\end{aligned}
$$

SIMULATIONS (continued)





DATA APPLICATIONS

- PBC-data (3 groups)
  $0.3 < T_2 \leq 1.2$
  $1.2 < T_2 \leq 2.5$
  $T_2 > 2.5$
- Diabetic rethinopathy data (4 groups) based on quantiles