Introduction
Methods
Simulations
Design
Results
Discussion

Adjusted survival curves by using inverse probability of treatment weighting: the comparison of three adapted log-rank tests

Florent Le Borgne [1,2] & Yohann Foucher [1]

[1] *EA4275-SPHERE - ITUN, University of Nantes, France*

[2] *IDBC/A2com, Pace, France*

florent.le-borgne@etu.univ-nantes.fr

27 August 2014

Introduction

Introduction
Methods
Simulations
Design
Results
Discussion

Context : Observational study in presence of survival data.

- The causality evaluation between the exposure and the time-to-event requires adjustment.

  $\Rightarrow$ Kaplan-Meier estimator inadequate

  $\Rightarrow$ Multivariate (Cox) model suitable but loss of information :

  Result summarized in a single HR : no graphical representation of a possible evolution over time of the HR

Solution

Adjusted survival curves using the method IPTW (Inverse Probability of Treatment Weighting) based on propensity scores

Introduction

Introduction
Methods
Simulations
Design
Results
Discussion

- The log-rank test = standard test for comparing two survival curves.
- Three versions adapted to the adjusted Kaplan-Meier estimator.
- Other methods based on propensity scores exist (stratification, matching, IPTW).

Objectives of our simulation study

- Evaluate the performances of the adjusted log-rank test compared to the Cox model in terms of type I and II error rates
  $\Rightarrow$ Is it necessary to use a multivariate Cox model ?
- Choose the most powerful among the three

Notations

Introduction
Methods
Simulations
Design
Results
Discussion

- $n =$ sample size
- $T_i =$ participating time $(i = 1, \dots, n)$
- $\delta_i =$ censoring indicator ($\delta_i = 0$ if T_i is a right censoring and $\delta_i = 1$ otherwise)
- $X_i =$ explanatory variable representing the interest exposure factor composed of K groups
- $D_k =$ number of different times for which events are observed in the group k, we then have at time t_j $(j = 1, \dots, D_k)$:
    - $d_{jk} = \sum_{i:t_i=t_j} \delta_i I(X_i = k)$: number of subjects in group k undergoing the event in time t_j
    - $Y_{jk} = \sum_{i:t_i \geq t_j} I(X_i = k)$: number of subjects in group k at risk at time t_j
    - Let $d_j = \sum_{k=1}^{K} d_{jk}$ and $Y_j = \sum_{k=1}^{K} Y_{jk}$

The IPTW weighting

Introduction
Methods
Simulations
Design
Results
Discussion

- The IPTW method proposes to correct the contribution of each individual by a weight $w_{ik} = 1/p_{ik}$.

  where $p_{ik} = P(X_i = k|Z_i)$

  and Z_i the vector of potential confounding factors

- The weighted number of events and individuals at risk can then be obtained :
  - $d_{jk}^w = \sum_{i:t_i=t_j} w_{ik}\delta_i I(X_i = k)$
  - $Y_{jk}^w = \sum_{i:t_i \geq t_j} w_{ik} I(X_i = k)$
  - $d_j^w = \sum_{k=1}^K d_{jk}^w$ et $Y_j^w = \sum_{k=1}^K Y_{jk}^w$

- Let us consider now only two groups, noted $\mathrm{X} = 0$ et $\mathrm{X} = 1$.

Adjusted log-rank proposed by Xu and al. (2012)

Introduction
Methods
Simulations
Design
Results
Discussion

1. Xu and al. proposed an adjusted log-rank test equivalent to the standard one by simply replacing :
   - The observed numbers of events by the weighted ones
   - The numbers of individuals at risk by the weighted ones

- The resulting statistic is : $G^w/\sqrt{\text{Var}(G^w)}$ where :
  - D is the number of different times for which events are observed regardless of group
  - $G^w = \sum_{j=1}^{D} d_{j1}^w - Y_{j1}^w\left(\frac{d_j^w}{Y_j^w}\right)$
  - $\text{Var}(G^w) = \sum_{j=1}^{D} \left\{ \frac{Y_{j0}^w Y_{j1}^w d_j^w (Y_j^w - d_j^w)}{(Y_j^w)^2 (Y_j^w - 1)} \right\}$

Xu S. and al. Extension of kaplan-meier methods in observational studies with time-varying treatment. *Value in Health.* (2012)

Adjusted log-rank proposed by Sugihara (2010)

Introduction
Methods
Simulations
Design
Results
Discussion

2. A second variant of the adjusted log-rank test is given by Sugihara.

- Differs from the first by the formula of the variance used.

$$G^w = \sum_{j=1}^{D} d_{j1}^w - Y_{j1}^w \left(\frac{d_j^w}{Y_j^w}\right)$$

$$\mathrm{Var}(G^{w'}) = \sum_{j=1}^{D} \left\{ \frac{d_j(Y_j - d_j)}{Y_j(Y_j - 1)} \sum_{i=1}^{Y_j} \left[\left(\frac{Y_{j0}^w}{Y_j^w}\right)^2 {w_i}^2 X_i + \left(\frac{Y_{j1}^w}{Y_j^w}\right)^2 {w_i}^2 (1 - X_i) \right] \right\}$$

Sugihara M. Survival analysis using inverse probability of treatment weighted methods based on the generalized propensity score. *Pharmaceutical statistics.* (2010)

Adjusted log-rank proposed by Xie and Liu (2005)

3. Xie and Liu proposed another adaptation of the log-rank test by adjusting the weights of each individual over the time.

- At time t_j ($j = 1, \ldots, D_k$), the weight for an individual i in the group k is reassigned as :

$$w'_{ijk} = w_{ik} . Y_{jk} / Y^{w}_{jk}$$

- The weighted number of events and at risk subjects becomes :

$$d^{w'}_{jk} = \sum_{i:t_i = t_j} w'_{ijk} \delta_i I(X_i = k)$$
$$\text{and } Y^{w'}_{jk} = \sum_{i:t_i \geq t_j} w'_{ijk} I(X_i = k)$$

Xie J. and Liu C. Adjusted kaplan-meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statistics in medicine.* (2005)

Adjusted log-rank proposed by Xie and Liu (2005)

Introduction
Methods
Simulations
Design
Results
Discussion

- Same formulas as those proposed by Sugihara but with different weights :

$$G^{w'} = \sum_{j=1}^{D} d_{j1}^{w'} - Y_{j1}^{w'} \left(\frac{d_j^{w'}}{Y_j^{w'}}\right)$$

$$\mathrm{Var}(G^{w'}) = \sum_{j=1}^{D} \left\{ \frac{d_j(Y_j - d_j)}{Y_j(Y_j-1)} \sum_{i=1}^{Y_j} \left[\left(\frac{Y_{j0}^{w'}}{Y_j^{w'}}\right)^2 w_{ij}^{'2} X_i + \left(\frac{Y_{j1}^{w'}}{Y_j^{w'}}\right)^2 w_{ij}^{'2} (1 - X_i) \right] \right\}$$

Xie J. and Liu C. Adjusted kaplan-meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statistics in medicine.* (2005)

Two other models based on propensity scores

Introduction
Methods
Simulations
Design
Results
Discussion

- Weighted univariate Cox model proposed by Cole et Hernán (2004)
  - Exposure : only variable in the model
  - Weighted by the weights w_{ik}

Cole S.R. and Hernán M. Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine* (2004)

- Matching on the logit of the propensity score
  - Matching 1:1 without replacement with the nearest neighbor
  - Caliper equal to 0.2 of the standard deviation
  - Stratified log-rank test

Rosenbaum PR. and Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* (1985)

Simulation study

- Simulations limited to 5 variables :
  - 1 binary exposure
  - 4 confounders
- Performances of the different models were compared for different :
  - Right-censoring rates (0.30 and 0.68)
  - Sample sizes (100, 250, 500 and 1500)
  - Percentages of exposed subjects (5%, 20% and 40%)
  - Coefficient β_X associated with the exposure variable under interest (0, 0.250, 0.365, 0.500)
- When $\beta_X = 0$ we calculated the percentage of rejection of the null hypothesis (type I error rate).
- When $\beta_X \neq 0$ we calculated the percentage of non-rejection of the null hypothesis (type II error rate).

Simulation Results

Introduction
Methods
Simulations
Design
Results
Discussion

β_X	n	Censoring rate $\approx$ 0.68 Multivariate Cox	Xu	Sugihara	Xie	Weighted Cox	Caliper
(a) 0.000	100	5.69	19.53	6.94	6.12	10.92	4.70
	250	5.31	23.42	6.58	6.27	8.80	4.81
	500	4.59	25.39	5.73	5.46	7.10	5.29
	1500	4.85	27.58	5.58	5.43	6.38	4.93
(b) 0.250	100	90.89	74.44	89.04	89.80	84.04	94.75
	250	86.05	65.78	86.20	86.37	82.57	92.28
	500	77.57	56.41	82.59	82.54	79.11	89.64
	1500	43.78	31.99	65.14	64.74	60.86	75.99
(b) 0.365	100	86.90	70.03	85.35	85.87	79.50	93.10
	250	74.66	55.24	79.20	79.30	73.37	88.20
	500	56.71	40.35	69.68	69.26	64.21	81.67
	1500	13.30	13.79	40.95	40.37	36.97	54.16
(b) 0.500	100	80.13	61.30	79.06	79.79	71.88	90.89
	250	58.80	41.57	68.06	67.87	60.14	81.77
	500	32.03	24.40	54.94	54.21	47.34	69.80
	1500	0.99	2.95	16.81	16.15	14.76	27.00

TABLE 1: Error rates obtained from data with 40% of exposed subjects and 68% of censure. (a) Type I errors rate in percentages. (b) Type II errors rate in percentages. 10 000 samples simulated for each scenario.

Conclusion of simulation study

Introduction
Methods
Simulations
Design
Results
Discussion

- Best performances obtained by the multivariate Cox model.
- Matched model : loss of power.
- Univariate weighted Cox model : more important type I error rate.
- Among the three versions of the adjusted log-rank tests :
  - The one proposed by Xu and al. does not respect the type I error rate
  - The two others show type I and type II error rates slightly higher than those of the multivariate Cox model
  - Slightly better type I error rates for the one proposed by Xie and Liu

Discussion (1)

Introduction
Methods
Simulations
Design
Results
Discussion

- Two limitations appear in our study :
  - We have only considered the case of a binary exposure.
    - Adjusted survival curves can be generalized to more than two groups (multinomial logistic regression)
    - Adjusted log-rank test requires further developments
  - Only the context in which the PH assumption holds true was simulated.

Discussion (2)

Introduction
Methods
Simulations
Design
Results
Discussion

- In conclusion, we retain two good methods :
  - The multivariate Cox model
  - The adjusted survival curves with the log-rank test proposed by Xie et Liu
- Multivariate Cox model : the most efficient, requires verification of assumptions, summarizes the results in one HR.
- Adjusted survival curves : illustrate more precisely the differences in survivals between groups, lower performances of the associated adjusted log-rank test.

Références

Introduction
Methods
Simulations
Design
Results
Discussion

- Xu S. and al. Extension of kaplan-meier methods in observational studies with time-varying treatment. *Value in Health.* (2012)
- Sugihara M. Survival analysis using inverse probability of treatment weighted methods based on the generalized propensity score. *Pharmaceutical statistics.* (2010)
- Xie J. and Liu C. Adjusted kaplan-meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statistics in medicine.* (2005)
- Cole S.R. and Hernán M. Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine* (2004)
- Rosenbaum PR. and Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*