Stratified logrank test of no randomized treatment effect with missing stratum information

Dupuy Jean-François

Institut de Mathématiques de Toulouse
Université Paul Sabatier, Toulouse

Journées 2007 du GDR Statistique et Santé

Outline of the talk

Introduction
- Notations - Nelson-Aalen estimator
- Comparing K survival distributions - Stratified test

The problem

The proposed test statistic
- The test statistic
- The null asymptotic distribution

A simulation study

Some perspectives

Notations

- T^0: time to failure (positive random variable),
- C: time to censoring (independent of T^0),
- $T = \min(T^0, C)$: observed duration,
- $\Delta = 1(T^0 \leq C)$: censoring indicator,

The hazard function: $\lambda(t) = \lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(T^0 < t + h | T^0 \geq t)$.

The cumulative hazard function: $\Lambda(t) = \int_0^t \lambda(u) \, du$.

The Nelson-Aalen estimator of Λ

Let

- $N(t) = 1(T \leq t, \Delta = 1)$ be the failure counting process,
- $Y(t) = 1(T \geq t)$ be the at-risk process.

$N(t)$ has intensity $Y(t)\lambda(t)$ and compensator $\int_0^t Y(u)\lambda(u)\,du$.

The Nelson-Aalen estimator of $\Lambda(t)$:

$$\widehat{\Lambda}(t) = \int_0^t \frac{\sum_{i=1}^n dN_i(u)}{\sum_{j=1}^n Y_j(u)} = \sum_{T_i \leq t} \frac{\Delta_i}{\sum_{j=1}^n Y_j(T_i)}.$$

Comparing K survival distributions

Test of the null hypothesis $H_0 : \lambda_1 = \ldots = \lambda_K$.

- $\overline{N}_k(t)$ and $\overline{Y}_k(t)$: failure counting and at-risk processes for group k,
- $\overline{N}_\bullet(t)$ and $\overline{Y}_\bullet(t)$: pooled-groups processes.

Logrank statistic: Let

$$Z_k = \int_0^\infty \left\{ d\overline{N}_k(t) - \frac{\overline{Y}_k(t)}{\overline{Y}_\bullet(t)} d\overline{N}_\bullet(t) \right\} \equiv O_k - E_k.$$

Under H_0,

$$LR = (Z_1, \ldots, Z_{K-1}) \widehat{\Sigma}^{-1} (Z_1, \ldots, Z_{K-1})' \xrightarrow{d} \chi^2_{K-1}.$$

Stratified test

- Control for some other potentially important factors.
- Suppose we have L strata, formed from the other factor(s) we want to control for.

Hazard for a subject in the kth treatment and lth stratum: $\lambda_{k,l}$.

Test of:

$$H_0 : \lambda_{1,l} = \ldots = \lambda_{K,l} \quad \text{for all } l = 1, \ldots, L$$

Stratified logrank test (SLR):

- basically compares $O_{k,l}$ to $E_{k,l}$ under H_0,
- combine information by summing across strata l,
- SLR is asymptotically χ^2_{K-1} under H_0.

- Pb: what if the stratum is unknown for some (but not all) patients?
- For each patient, we observe:

$$\mathcal{O} = (T, \Delta, G, R, RS, W) \qquad \text{with}$$

  - $G \in \{1, \ldots, K\}, \quad S \in \{1, \ldots, L\}, \quad R = 1(S \text{ is observed}),$
  - W an auxiliary variable for S.
- Assume a missing-at-random (MAR) mechanism.

Statistical pb: implement SLR from n independent copies of $\mathcal{O}$.

The test statistic

- For $1 \leq i \leq n$, let $D_i^l = R_i 1(S_i = l) + (1 - R_i)\mathbb{P}(S = l|W_i)$.
- For $1 \leq k \leq K$, let

$$Z_k^* = \sum_{i=1}^{n} \int_0^{\infty} \left\{ 1(G_i = k) - \sum_{l=1}^{L} D_i^l \frac{\sum_{j=1}^{n} Y_j(t) 1(G_j = k) D_j^l}{\sum_{j=1}^{n} Y_j(t) D_j^l} \right\} dN_i(t).$$

- Key to the null asymptotic distribution of LR: represent Z_k as a martingale process $\sum_i \int H_i \, (dN_i - d\Lambda_i)$.
- Not possible for Z_k^* due to subjects $\{j : R_j = 0\}$, who contribute to each stratum.

The null asymptotic distribution

- Approximate Z_k^* by a sum of n iid random vectors, so as to use the multivariate CLT:

$$n^{-1/2}Z_k^* = n^{-1/2}\sum_{i=1}^{n} Q_{k,i} + o_p(1), \quad k = 1, \ldots, K.$$

- Innovation theorem. If $N_i(t)$ has intensity

$$Y_i(t)\sum_{k,l}\lambda_{k,l}(t)1(G_i = k)1(S_i = l) := Y_i(t)\lambda_i(t)$$

wrt $\mathcal{F}_t^i = \sigma\{N_i(s), N_i^C(s), G_i, S_i, W_i : s \leq t\}$, it has intensity

$$Y_i(t)\mathbb{E}[\lambda_i(t)|\mathcal{G}_{t-}^i] := Y_i(t)\alpha_i(t)$$

wrt $\mathcal{G}_t^i = \sigma\{N_i(s), N_i^C(s), G_i, W_i : s \leq t\} \subset \mathcal{F}_t^i$.

The null asymptotic distribution (ctd)

- Let $\mathcal{H}_t^i = (\mathcal{F}_t^i)^{R_i}(\mathcal{G}_t^i)^{1-R_i}$. $N_i(t)$ has intensity $Y_i(t)\tilde{\lambda}_i(t)$ with

$$\tilde{\lambda}_i(t) = R_i\lambda_i(t) + (1 - R_i)\alpha_i(t)$$

so that $M_i(t) = N_i(t) - \int_0^t Y_i(s)\tilde{\lambda}_i(s)ds$ is a martingale wrt $\mathcal{H}_t^i$.

- Under H_0, $\mathbb{E}Q_{k,i} = 0$ and thus,

$$n^{-1/2}\mathbf{Z}^* = n^{-1/2}\left(Z_1^*, \ldots, Z_{K-1}^*\right)' \xrightarrow{d} N(0, \mathbf{V}).$$

Finally, under H_0,

$$\mathbf{U} = \mathbf{Z}^{*\prime}\widehat{\mathbf{V}}^{-1}\mathbf{Z}^* \xrightarrow{d} \chi^2_{K-1}.$$

The null asymptotic distribution (ctd)

- The $\mathbb{P}(S = l|W_i)$ will usually be unknown:
  - estimate from $\{j : R_j = 1\}$ and substitute in D_i^l (by MAR),
  - yields an estimated version of $\mathbf{U}$:

$$\widehat{\mathbf{U}} = \widehat{\mathbf{Z}}^{*\prime}\widetilde{\mathbf{V}}^{-1}\widehat{\mathbf{Z}}^{*}$$

- The null asymptotic distribution of $\widehat{\mathbf{U}}$ may be distorted from the χ^2_{K-1}:
  - may affect size of the test,
  - $\Rightarrow$ simulations.

A simulation study

- 2 treatments, 2 strata, half patients in each treatment and stratum,
- S=1, G=1: $T^0 \sim \mathcal{W}(0.75, 1.5)$,
- S=1, G=2: $T^0 \sim \mathcal{W}(0.75, 1.5 \times r_1)$, with r_1 the hazard ratio,
- similarly for stratum 2 with $(0.5, 0.75, r_2)$,
- $C \sim \mathcal{E}(\theta)$, $W \in \mathbb{R}^2$,
- 2000 repetitions,

Parameter	Values	Description
(r_1, r_2)	(1,1), (1.5,1.5), (2,3)	Ratios of hazard rates
c_p	0, 25, 50	Censoring %
sm_p	0, 20, 40, 60	Stratum missingness %
n	60, 100, 140	Sample size

A simulation study: some conclusions

Under the MAR scenario:

- Both SLR_{cc} and $\widehat{\mathbf{U}}$ globally satisfy the prescribed level,
- For (1.5,1.5) and (2,3):
  - the powers of $\widehat{\mathbf{U}}$ are greater than those of SLR_{cc}, for every sm_p and censoring fraction,
  - $\widehat{\mathbf{U}}$ maintains a high power even when $sm_p = 60\%$, while at the same time, the powers of the SLR_{cc} seriously decrease.

Robustness to deviations to the MAR assumption?

Simulation scenario:

- $R \sim \mathcal{B}(p_1)$ in stratum 1 and $R \sim \mathcal{B}(p_2)$ in stratum 2,
- $sm_p = 20\%, 40\%, 60\%,$
- for $sm_p = 40\%$, $(sm_p(1), sm_p(2)) = (50\%, 30\%)$, $(60\%, 20\%)$, $(70\%, 10\%)$.

Under a non-ignorable scenario:

- The power of $\widehat{\mathbf{U}}$ decreases as deviation to MAR increases,
- Similarly for SLR_{cc}, but in a less extent.

Robustness (ctd)

- Heavy stratum missingness
  - power of $\widehat{\mathbf{U}}$ greater than power of SLR_{cc}, even for important deviations to MAR,
- Small / moderate stratum missingness
  - moderate deviation to MAR: power of $\widehat{\mathbf{U}}$ greater than power of SLR_{cc},
  - strong deviation to MAR:
    - power of $\widehat{\mathbf{U}} \geq$ power of SLR_{cc} if heavy censoring,
    - power of $\widehat{\mathbf{U}} \approx$ power of SLR_{cc} if moderate censoring,
    - power of $\widehat{\mathbf{U}} \leq$ power of SLR_{cc} if no censoring.

Some perspectives

- Group-dependent censoring: the IPCW approach,
- Consider the more general problem of estimation and testing in the stratified Cox model with missing stratum information.