Extrapolation of internal pilot estimates for sample size re-assessment with recurrent event data in the presence of non-constant hazards

Katharina Ingel   Harald Binder   Antje Jahn-Eimermacher

Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI)
University Medical Center of the Johannes-Gutenberg-University Mainz, Germany

June 26, 2015

JG|U UNIVERSITÄTSmedizin. MAINZ

The POET Study

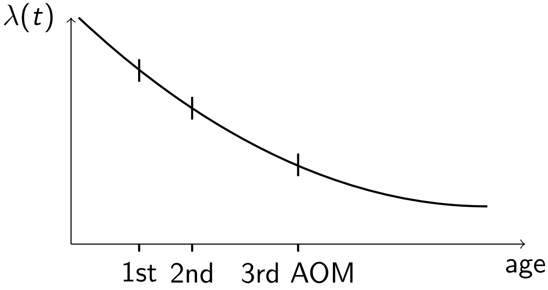
PRYMULA et al. (2006):

- randomised, double-blind efficacy study
- 4968 children aged between 6 weeks and 5 months
- four doses of either protein D conjugate vaccine or hepatitis A vaccine
- follow-up until 24–27 months of age
- aim: reduce risk for acute otitis media (AOM)

The POET Study

PRYMULA et al. (2006):

- randomised, double-blind efficacy study
- 4968 children aged between 6 weeks and 5 months
- four doses of either protein D conjugate vaccine or hepatitis A vaccine
- follow-up until 24–27 months of age
- aim: reduce risk for acute otitis media (AOM)



Andersen-Gill Model

individual hazard:

$$\lambda_i(t) = \lambda_0(t) \cdot Z_i \cdot \exp(\beta^t X_i)$$

with

- $\lambda_0(t)$ baseline hazard
- Z_i frailty variable with $E(Z_i) = 1$ and $Var(Z_i) = \theta$
- β vector of coefficients
- X_i covariates

Sample Size Calculation

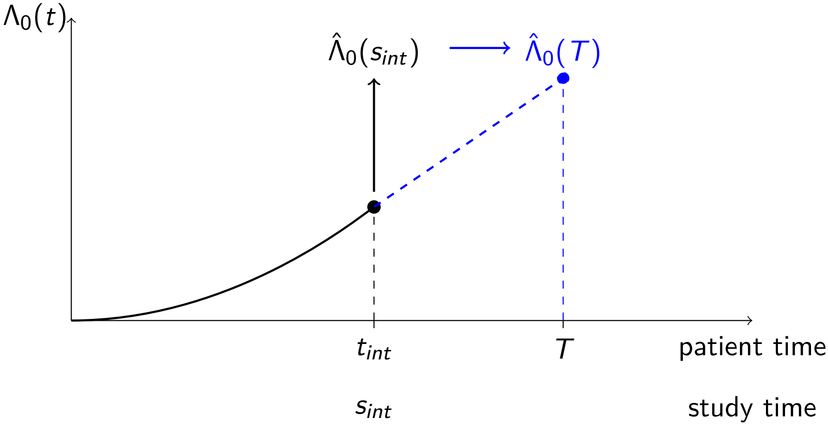
Ingel, Jahn-Eimermacher (2014):

$$L = 4 \cdot \left(1 + \theta \cdot \Lambda_0(T) \cdot \left(\frac{1 + \exp(2\beta_1)}{1 + \exp(\beta_1)}\right)\right) \cdot \left(\frac{z_{1-\frac{\alpha}{2}} + z_{1-\gamma}}{\beta_1}\right)^2$$

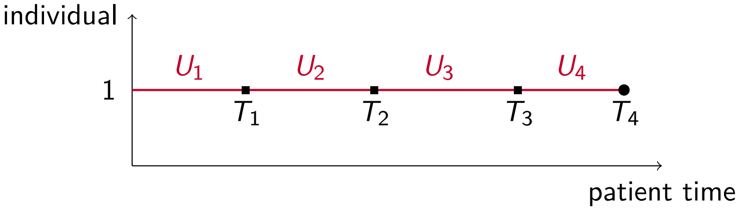
- $[0, T]$ patients' follow-up period
- θ degree of heterogeneity

- $\Lambda_0(T)$ cumulative baseline hazard at T

Internal Pilot Study



Simulation Algorithm



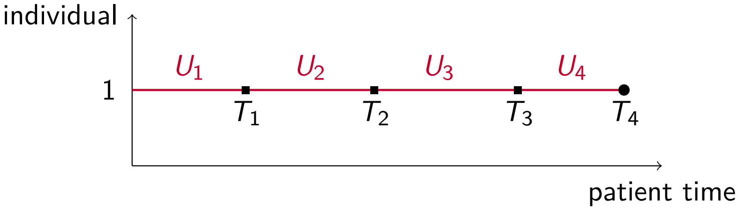
$U_i | T_{i-1} = t$ conditional inter-event time with cumulative hazard

$$\tilde{\Lambda}_t(u) := \tilde{\Lambda}^i(u | T_{i-1} = t) = \Lambda(u + t) - \Lambda(t)$$

simulation via

$$(U_i | T_{i-1} = t) = \tilde{\Lambda}_t^{-1}(-\log(A)) \quad \text{with } A \sim U[0,1]$$

Simulation Algorithm



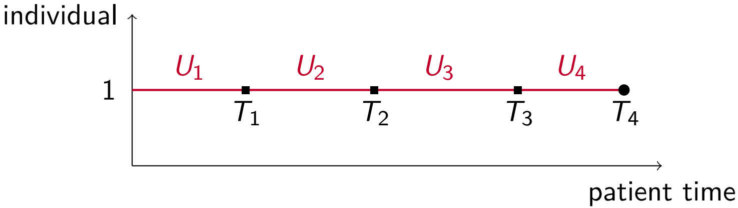
$U_i | T_{i-1} = t$ conditional inter-event time with cumulative hazard

$$\tilde{\Lambda}_t(u) := \tilde{\Lambda}^i(u | T_{i-1} = t) = \Lambda(u + t) - \Lambda(t)$$

simulation via

$$(U_i | T_{i-1} = t) = \tilde{\Lambda}_t^{-1}(-\log(A)) \quad \text{with } A \sim U[0, 1]$$

Simulation Algorithm



$U_i | T_{i-1} = t$ conditional inter-event time with cumulative hazard

$$\tilde{\Lambda}_t(u) := \tilde{\Lambda}^i(u | T_{i-1} = t) = \Lambda(u + t) - \Lambda(t)$$

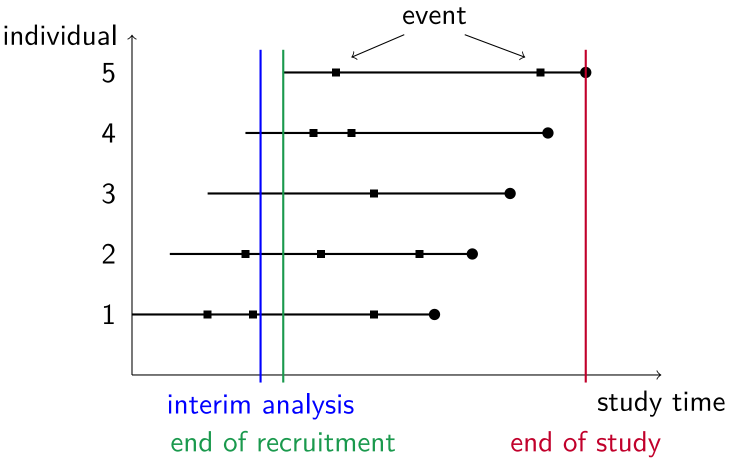
simulation via

$$(U_i | T_{i-1} = t) = \tilde{\Lambda}_t^{-1}(-\log(A)) \quad \text{with } A \sim U[0, 1]$$

Simulation of Patient Data

```
library(simrec)
simdata <- simrec(N = 100,
                  fu.min = 2, fu.max = 2,
                  cens.prob = 0.2,
                  dist.x = "binomial", par.x = 0.5,
                    beta.x = log(0.5),
                  dist.z = "gamma", par.z = 0.5,
                  dist.rec = "weibull",
                    par.rec = c(4,0.585))
```

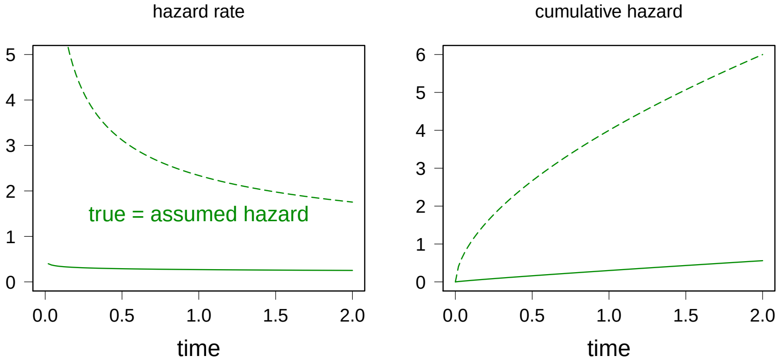
Simulation Study



2 years of follow-up, 20% censoring at random, $HR = 0.5, \theta = 0.5$

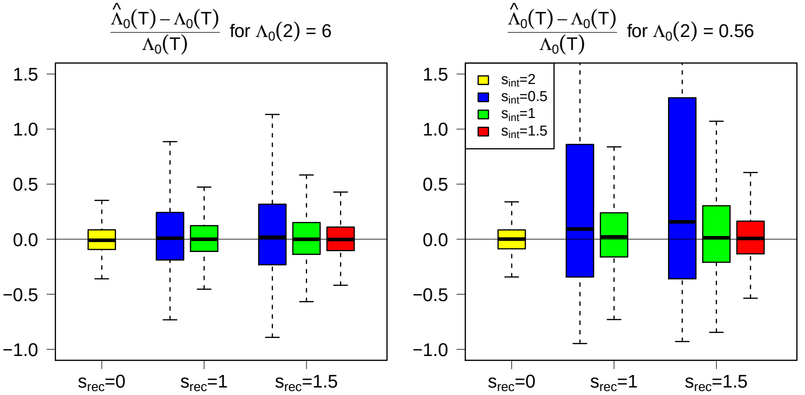
Weibull Distribution

$$\lambda_0(t) = \lambda \cdot \nu \cdot t^{\nu-1}, \quad \Lambda_0(t) = \lambda \cdot t^{\nu}$$

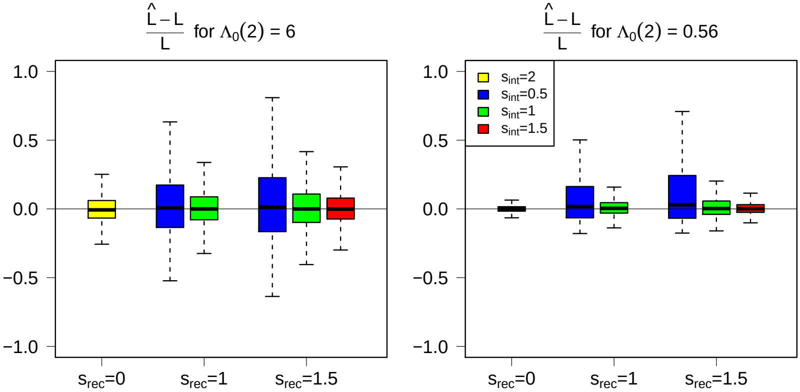


$- - -$ $\lambda = 4, \nu = 0.585$ with $\Lambda_0(2) = 6$,
— $\lambda = 0.3, \nu = 0.9$ with $\Lambda_0(2) = 0.56$

Results - Weibull

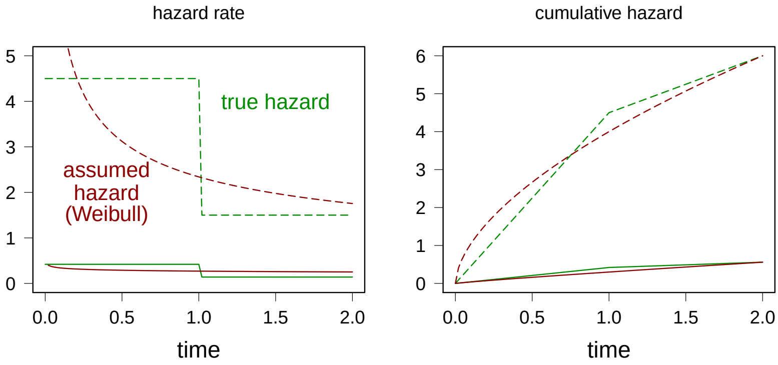


Results - Weibull



Step Function

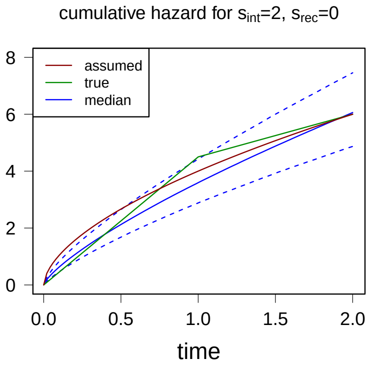
$$\lambda_0(t) = \begin{cases} \lambda_1 \text{ if } t \leq t_1, \\ \lambda_2 \text{ if } t > t_1 \end{cases}$$

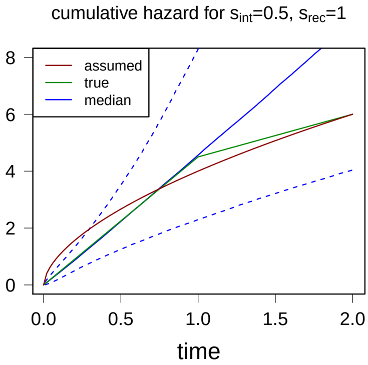


$$- - - \; \Lambda_0(2) = 6, \quad — \; \Lambda_0(2) = 0.56$$

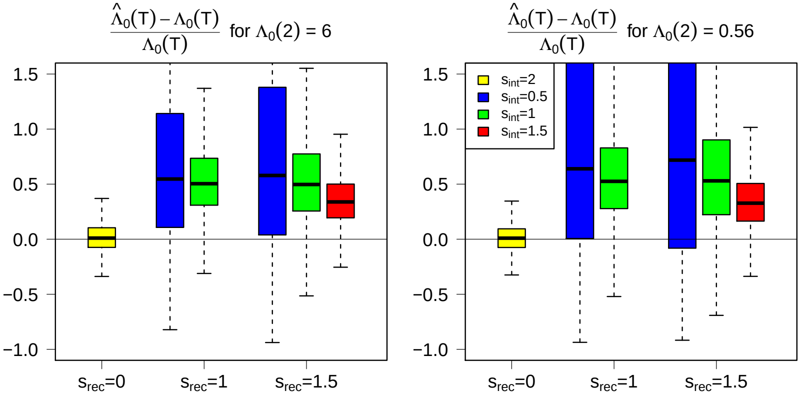
Results - Step Function

$$\Lambda_0(2) = 6$$

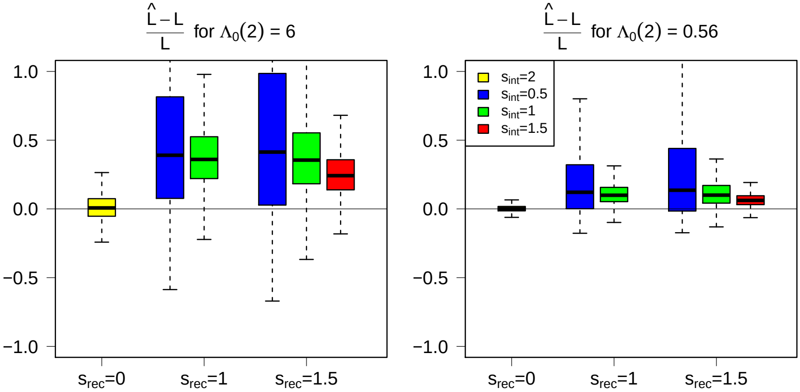




Results - Step Function

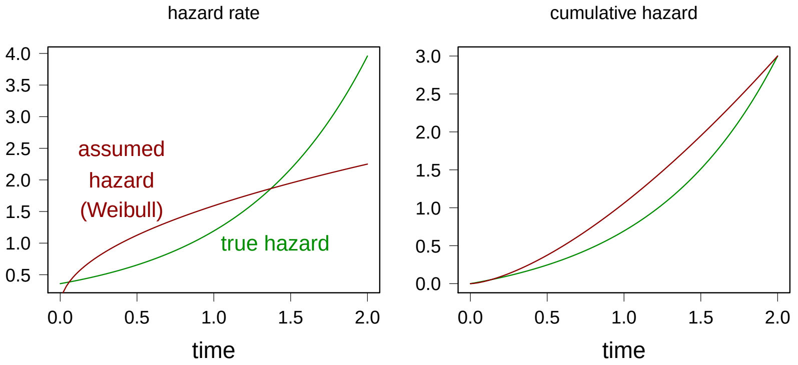


Results - Step Function



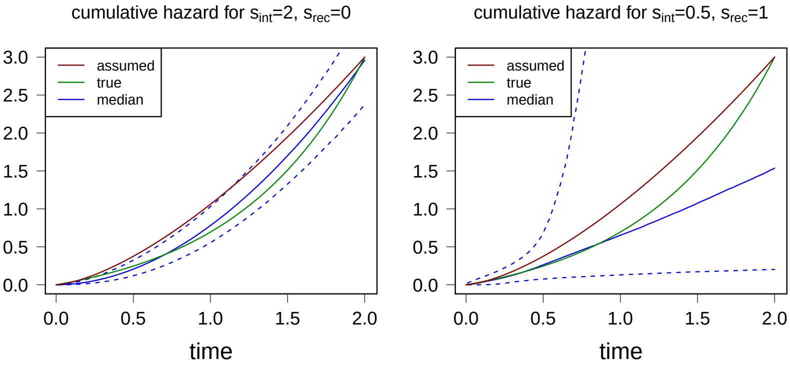
Gompertz Distribution

$$\lambda_0(t) = \lambda \cdot \exp(\alpha t), \quad \Lambda_0(t) = \frac{\lambda}{\alpha}(\exp(\alpha t) - 1)$$

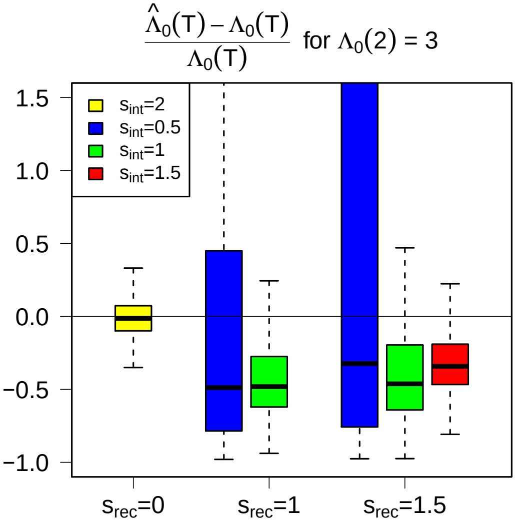


$$\lambda_{GP} = 0.36, \alpha_{GP} = 1.2 \text{ and } \lambda_{WB} = 1.061, \nu_{WB} = 1.5$$
$$\text{with } \Lambda_0(2) = 3$$

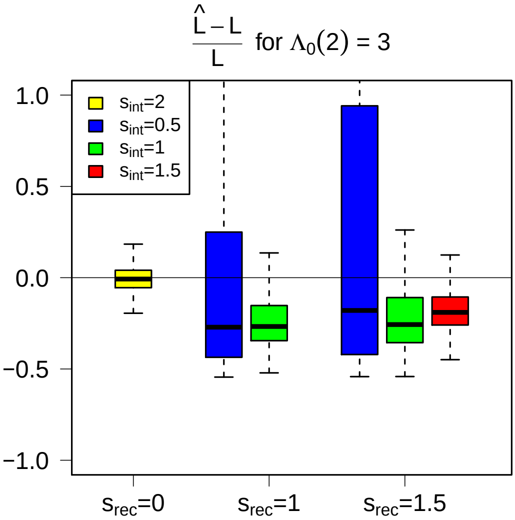
Results - Gompertz



Results - Gompertz



Results - Gompertz



Conclusions

- extrapolation depends on timing of the interim and the underlying parametric distribution
- Very early interim analyses not recommended
- simulation algorithm easy to apply by calling `simrec`:

```
devtools::install_github("katharinaingel/simrec")
library(simrec)
simdata <- simrec(...)
```

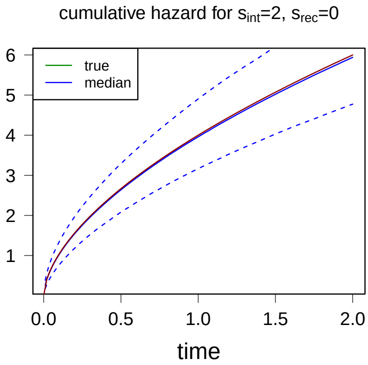
References

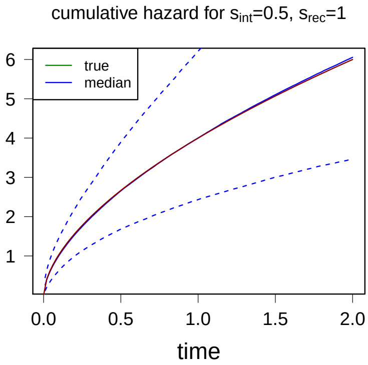
- INGEL, K. and JAHN-EIMERMACHER, A. (2014): Sample-size calculation and reestimation for a semiparametric analysis of recurrent event data taking robust standard errors into account. *Biometrical Journal, 56, 631–648*.
- JAHN-EIMERMACHER, A., INGEL, K., OZGA, A., PREUSSLER, S. and BINDER, H. (2015): Simulating recurrent event data with hazard functions defined on a total time scale. *BMC, 15:16*.
- `simrec`: An R-Package for Simulation of Recurrent Event Data. https://github.com/katharinaingel/simrec

- contact: ingel@uni-mainz.de

Results - Weibull

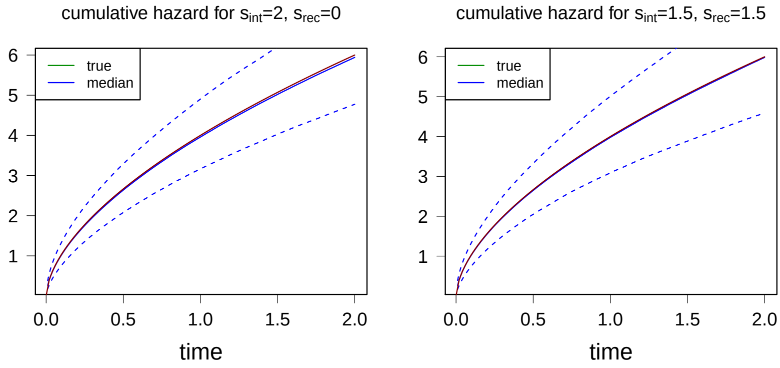
$$\Lambda_0(2) = 6$$





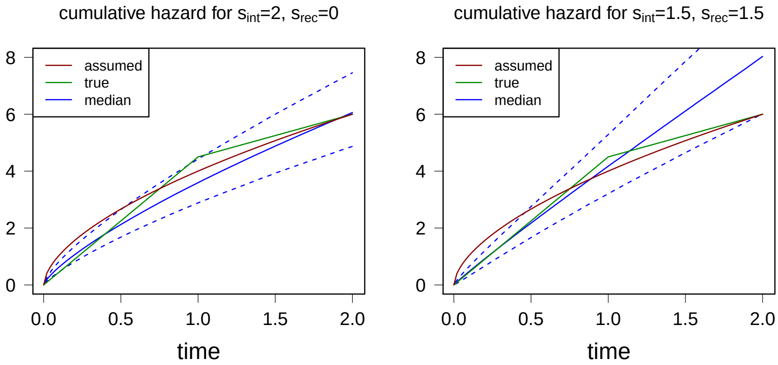
Results - Weibull

$$\Lambda_0(2) = 6$$

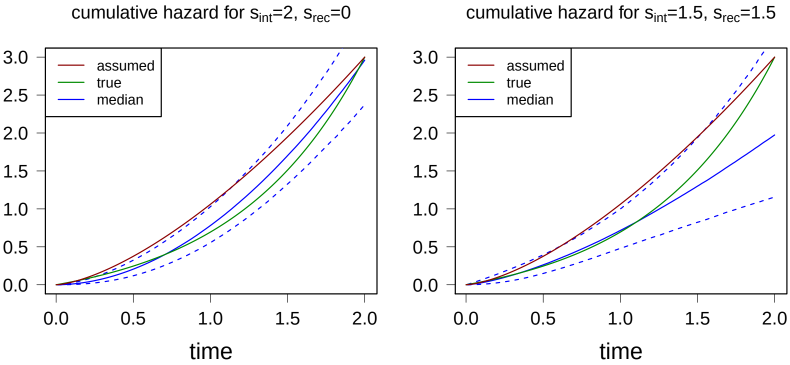


Results - Step Function

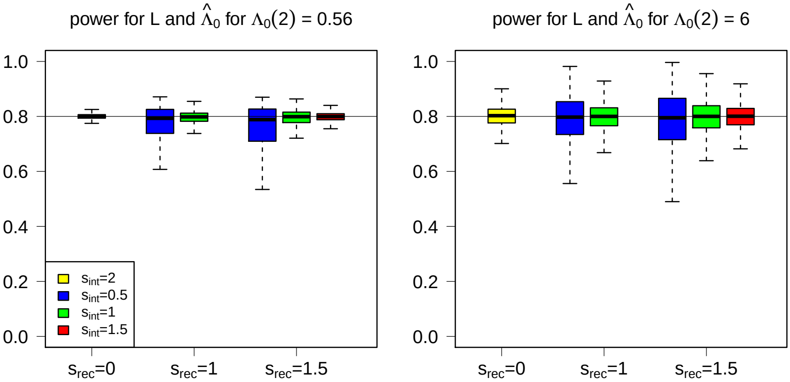
$$\Lambda_0(2) = 6$$



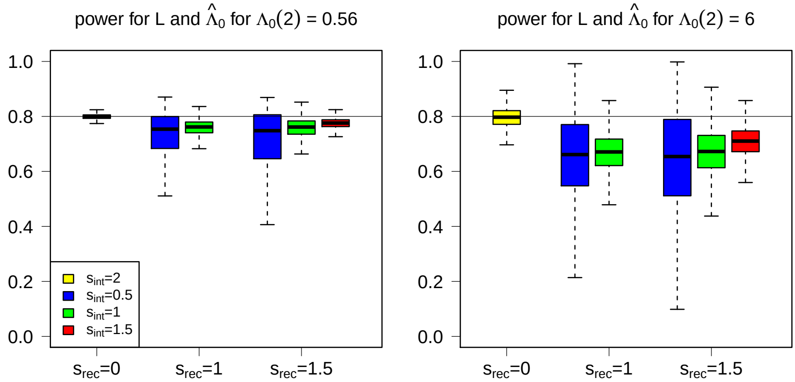
Results - Gompertz



Results - Weibull



Results - Step Function



Results - Gompertz