Group Project 1

Huanyu Chen, Shaolei Ma, Ruiqi Xue

2024-02-20

Abstract

This study compares the performance of three hypothesis tests for time-to-event data: the conventional log-rank test and two variants of the weighted log-rank test. We assess their effectiveness under scenarios involving both proportional and non-proportional hazard functions, using Monte Carlo simulation techniques to evaluate power across a range of coefficient values. Our analysis highlights nuanced differences in their ability to detect treatment effects, providing insights into selecting appropriate statistical methodologies for analyzing time-to-event data in clinical trials.

1 Introduction

In clinical trials, treatment efficacy is typically evaluated using time-to-event outcomes and the hazard ratio under the proportional hazards assumption. The log-rank test is used to compare observed and expected event counts. However, in real-world scenarios, the proportional hazards assumption may not always hold, requiring statistical adjustments. To address this issue, researchers have proposed weighted log-rank tests that incorporate parameters for different emphases on early or late events. To test the performance of various log-rank tests, we focus on both proportional hazards and non-proportional hazards based on exponential and Weibull distribution assumptions, and then evaluate the power of different parameters.

2 Methods

2.1 Proportional-Hazard Assumption

Under proportional-hazards assumption, the hazard function (Cox model) can be written as:

$$h(t|x) = h_0(t)exp(\beta' x)$$

where t is the time, x the vector of covariates, β the vector of regression coefficients, $h_0(t)$ is the baseline hazard function. Then, the survival function is

$$S(t|x) = exp[-H_0(t)exp(\beta' x)]$$

where

$$H_0(t) = \int_0^t h_0(u)du$$

Thus, the distribution function is

$$F(t|x) = 1 - exp[-H_0(t)exp(\beta' x)]$$

Let Y be a random variable with distribution function F, then $U = F(Y) \sim U(0,1)$, $(1-U) \sim U(0,1)$, i.e.

$$U = exp[-H_0(t)exp(\beta' x)] \sim U(0,1)$$

if $h_0(t) > 0$ for all t, then H_0 can be inverted and the survival time T of the model can be written as

$$T = H_0^{-1}[-log(U)exp(-\beta' x)]$$

where $U \sim U(0,1)$.

To simply the problem, here we only consider one covariate x, which indicates whether the sample belongs to the control arm ($x = 0$) or the treatment arm ($x = 1$), and set a negative β under the assumption that the treatment has a consistent positive effect.

Now, we only need to know H_0^{-1} to simulate the survival time. To do so, we consider two commonly used survival time distributions: **Exponential distribution** and **Weibull distribution**.

2.1.1 Exponential Distribution

For exponential distribution with scale parameter $\lambda > 0$, the possibility density function is $f_0 = \lambda exp(-\lambda t)$. Thus, $T = -\lambda^{-1} log(U)exp(-\beta' x)$ where $U \sim U(0,1)$.

2.1.2 Weibull Distribution

For Weibull distribution with the scale parameter λ, and is the shape parameter γ, the possibility density function is $f_0 = \lambda\gamma t^{\gamma-1} exp(-\lambda t^{\gamma})$. Thus, $T = (-\lambda^{-1} log(U)exp(-\beta' x))^{1/\gamma}$ where $U \sim U(0,1)$.

2.2 Non-Proportional-Hazard Assumption

Under Non-Proportional-Hazard Assumption, we still consider the exponential model and Weibull model.

2.2.1 Piecewise Exponential Model

Late Effect: We suppose the hazard function for the treatment arm is: $h(t|x=1) = \begin{cases} \lambda_0 & t < 1 \\ \lambda_1 & t \geq 1 \end{cases}$

Thus, $T = \begin{cases} -\lambda_0^{-1} log(U) & U > exp(-\lambda_0) \\ \dfrac{\lambda_1 - log(U)}{\lambda_0 + \lambda_1} & U \leq exp(-\lambda_0) \end{cases}$

Early Effect: We can use the similar simulation method to generate piecewise exponential models in which the treatment arm shows early effect. The hazard function becomes: $h(t|x=1) = \begin{cases} \lambda_0 & t \geq 1 \\ \lambda_1 & t < 1 \end{cases}$

Thus, $T = \begin{cases} -\lambda_0^{-1} log(U) & U \leq exp(-\lambda_0) \\ \dfrac{\lambda_1 - log(U)}{\lambda_0 + \lambda_1} & U > exp(-\lambda_0) \end{cases}$ where $U \sim U(0,1)$.

2.2.2 Weibull Model

To simplify the problem, we assume the control and treatment arm share the same scale parameter λ. For the control arm, suppose the hazard function is: $h(t|x=0) = \lambda\gamma_0 t^{(\gamma_0 - 1)}$. Thus, $T = (-\lambda^{-1} log(U))^{1/\gamma_0}$.

Similarly, we can write the hazard function for the treatment arm as: $h(t|x=1) = \lambda\gamma_1 t^{(\gamma_1 - 1)}$. We can derive that $T = (-\lambda^{-1} log(U))^{1/\gamma_1}$.

3 Simulation Results

Table 1: Specificity of 3 Log-Rank Tests based on PH Assumption

n	baseline	lambda	beta	test1_specificity	test2_specificity	test3_specificity
200	Exponential	1.0	-5.0	1.00	1.00	1.00
100	Exponential	1.0	-5.0	1.00	1.00	1.00
200	Exponential	0.8	-5.0	1.00	1.00	1.00
100	Exponential	0.8	-5.0	1.00	1.00	1.00
200	Exponential	0.5	-5.0	1.00	1.00	1.00
100	Exponential	0.5	-5.0	1.00	1.00	1.00
200	Exponential	1.0	-1.0	1.00	1.00	1.00
100	Exponential	1.0	-1.0	0.94	0.96	0.94
200	Exponential	0.8	-1.0	1.00	1.00	1.00
100	Exponential	0.8	-1.0	0.98	0.92	0.98
200	Exponential	0.5	-1.0	1.00	1.00	1.00
100	Exponential	0.5	-1.0	0.98	0.94	0.96
200	Exponential	1.0	-0.5	0.90	0.68	0.84
100	Exponential	1.0	-0.5	0.60	0.52	0.54
200	Exponential	0.8	-0.5	0.80	0.74	0.74
100	Exponential	0.8	-0.5	0.64	0.48	0.58
200	Exponential	0.5	-0.5	0.84	0.74	0.74
100	Exponential	0.5	-0.5	0.62	0.46	0.60

Table 2: Specificity of 3 Log-Rank Tests based on PH Assumption

n	baseline	lambda	gamma	beta	test1_specificity	test2_specificity	test3_specificity
200	Weibull	1.0	1.5	-5.0	1.00	1.00	1.00
100	Weibull	1.0	1.5	-5.0	1.00	1.00	1.00
200	Weibull	0.8	1.5	-5.0	1.00	1.00	1.00
100	Weibull	0.8	1.5	-5.0	1.00	1.00	1.00
200	Weibull	0.5	1.5	-5.0	1.00	1.00	1.00
100	Weibull	0.5	1.5	-5.0	1.00	1.00	1.00
200	Weibull	1.0	1.2	-5.0	1.00	1.00	1.00
100	Weibull	1.0	1.2	-5.0	1.00	1.00	1.00
200	Weibull	0.8	1.2	-5.0	1.00	1.00	1.00
100	Weibull	0.8	1.2	-5.0	1.00	1.00	1.00
200	Weibull	0.5	1.2	-5.0	1.00	1.00	1.00
100	Weibull	0.5	1.2	-5.0	1.00	1.00	1.00
200	Weibull	1.0	1.5	-1.0	1.00	1.00	1.00
100	Weibull	1.0	1.5	-1.0	0.96	0.94	0.94
200	Weibull	0.8	1.5	-1.0	1.00	1.00	1.00
100	Weibull	0.8	1.5	-1.0	0.94	0.88	0.94
200	Weibull	0.5	1.5	-1.0	0.84	0.72	0.82
100	Weibull	0.5	1.5	-1.0	0.50	0.48	0.48
200	Weibull	1.0	1.2	-1.0	1.00	1.00	1.00
100	Weibull	1.0	1.2	-1.0	0.98	0.96	0.98
200	Weibull	0.8	1.2	-1.0	1.00	1.00	1.00
100	Weibull	0.8	1.2	-1.0	0.98	0.80	0.94
200	Weibull	0.5	1.2	-1.0	0.90	0.72	0.90
100	Weibull	0.5	1.2	-1.0	0.70	0.52	0.68
200	Weibull	1.0	1.5	-0.5	0.86	0.76	0.78
100	Weibull	1.0	1.5	-0.5	0.54	0.40	0.52
200	Weibull	0.8	1.5	-0.5	0.54	0.42	0.56
100	Weibull	0.8	1.5	-0.5	0.40	0.36	0.34
200	Weibull	0.5	1.5	-0.5	0.44	0.26	0.40
100	Weibull	0.5	1.5	-0.5	0.24	0.18	0.18
200	Weibull	1.0	1.2	-0.5	0.94	0.82	0.86
100	Weibull	1.0	1.2	-0.5	0.64	0.50	0.56
200	Weibull	0.8	1.2	-0.5	0.70	0.56	0.70
100	Weibull	0.8	1.2	-0.5	0.42	0.32	0.40
200	Weibull	0.5	1.2	-0.5	0.46	0.36	0.46
100	Weibull	0.5	1.2	-0.5	0.28	0.20	0.28

Table 3: Specificity of 3 Log-Rank Tests based on NPH Assumption (Late)

n	lambda0	lambda1	test1_specificity	test2_specificity	test3_specificity
200	0.8	0.4	0.10	0.28	0.04
100	0.8	0.4	0.14	0.28	0.06
200	0.5	0.4	0.86	1.00	0.34
100	0.5	0.4	0.58	0.80	0.18
200	0.8	0.3	0.14	0.18	0.06
100	0.8	0.3	0.10	0.08	0.08
200	0.5	0.3	0.58	0.82	0.20
100	0.5	0.3	0.38	0.60	0.18

Table 4: Specificity of 3 Log-Rank Tests based on NPH Assumption (Early)

n	lambda0	lambda1	test1_specificity	test2_specificity	test3_specificity
200	0.9	0.7	0.32	0.24	0.90
100	0.9	0.7	0.16	0.14	0.70
200	0.8	0.7	0.22	0.26	0.88
100	0.8	0.7	0.12	0.10	0.66
200	0.9	0.6	0.30	0.34	0.80
100	0.9	0.6	0.10	0.16	0.58
200	0.8	0.6	0.10	0.40	0.62
100	0.8	0.6	0.08	0.16	0.32

Table 5: Specificity of 3 Log-Rank Tests based on NPH Assumption

n	lambda	gamma0	gamma1	test1_specificity	test2_specificity	test3_specificity
200	0.8	5	1.5	0.92	0.04	0.98
100	0.8	5	1.5	0.82	0.06	0.90
200	0.5	5	1.5	1.00	0.98	1.00
100	0.5	5	1.5	0.98	0.90	0.98
200	0.8	3	1.5	0.42	0.14	0.64
100	0.8	3	1.5	0.20	0.10	0.48
200	0.5	3	1.5	0.76	0.40	0.84
100	0.5	3	1.5	0.54	0.16	0.56
200	0.8	5	1.2	0.98	0.08	1.00
100	0.8	5	1.2	0.86	0.02	0.94
200	0.5	5	1.2	1.00	1.00	1.00
100	0.5	5	1.2	1.00	0.94	1.00
200	0.8	3	1.2	0.58	0.12	0.90
100	0.8	3	1.2	0.24	0.08	0.50
200	0.5	3	1.2	0.98	0.44	0.98
100	0.5	3	1.2	0.80	0.30	0.82

4 Conclusion

5 Referrence

Austin, P. C. (2012). Generating survival times to simulate Cox proportional hazards models with time-varying covariates. Statistics in Medicine, 31(29), 3946–3958. https://doi.org/10.1002/sim.5452

Bender, R., Augustin, T., & Blettner, M. (2005). Generating survival times to simulate Cox proportional hazards models. Statistics in Medicine, 24(11), 1713–1723. https://doi.org/10.1002/sim.2059

Bardo, M. F., Huber, C., Benda, N., Brugger, J., Fellinger, T., Vaidotas Galaune, Heinz, J., Heinzl, H., Hooker, A. C., Florian Klinglmüller, Franz König, Mathes, T., Mittlböck, M., Posch, M., Ristl, R., & Friede, T. (2023). Methods for non-proportional hazards in clinical trials: A systematic review. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.2306.16858