Testing the Proportional Versus Additive Hazards Assumptions

Lee S. McDaniel, Menggang Yu, and Rick Chappell

Louisiana State University Health Sciences Center, New Orleans; University of Wisconsin, Madison

Objective

Develop a powerful test for model adequacy in time-to-event data with specific alternative models. We test either a null of proportional hazards or a null of additive hazards, against the alternative of additive or proportional hazards, respectively.

Introduction

Time-to-event data are commonly analyzed using an assumption of proportional hazards. In the two group case, as may occur in a clinical trial, we may wish to test

$$H_0 : \lambda_B(t) = \Delta\lambda_A(t) \qquad \text{for some } \Delta > 0$$
$$H_1 : \lambda_B(t) \neq \Delta\lambda_A(t) \qquad \text{for any } \Delta > 0.$$

Several omnibus tests of these hypotheses exist:

- Cox [1] suggests testing significance of a time-varying covariate
- Schoenfeld [2] partitions the time axis and conducts a chi-square test
- Wei [3] conducts a chi-square test without partitioning the time axis
- Nagelkerke et. al. [4] check for large autocovariances between successive contributions to the score.

However, greater power may be obtained when a specific alternative is in mind. Gill and Schumacher [5] compare two different generalized rank estimators, which should be similar if the proportional hazards model holds, and provide the form of the rank estimators that are most efficient under a specific alternative.

Chappell [6] shows that the test proposed by Gill and Schumacher [5] is equivalent to the test proposed by Cox [1], but with specific recommendations for the time varying covariate. This allows the test to be easily incorporated into existing software and generalized to situations more complex than the two sample case. In this form of the test, the time-varying covariate introduced into the model (as recommended by Gill and Schumacher [5]) is $\widehat{S}(t)$, the estimated survival curve without regard for group.

Additive Hazards in Clinical Trials

Additive hazards models have benefits that are especially salient in clinical trials. The estimator for the treatment effect in the additive hazards model:

- is asymptotically unbiased when the hazard's dependence on time or covariates is misspecified [7]
- has an asymptotic variance unaffected by adjustment for baseline covariates [7]
- has assumptions that are not affected by modification in number of covariates modeled [8].

In addition to this robustness, interpretation may be easier with the additive hazards model. This depends on the interpreter and the situation.

Proportional Null

In the two sample case, we wish to test

$$H_0 : \lambda_B(t) = \Delta\lambda_A(t) \qquad \Delta > 0$$
$$H_1 : \lambda_B(t) = \Delta\lambda_A(t) + \gamma \qquad \Delta > 0, \gamma \neq 0.$$

In order to test this null, we use a time-varying covariate, and fit the following proportional hazards model:

$$\lambda_B(t) = \lambda_A(t)e^{\beta_1+\beta_2 W(t)}.$$

If the estimate for β_2 is significantly different from 0, then we have evidence to reject the proportional hazards model.

Then, based on the results from Gill and Schumacher [5], the most efficient choice for $W(t)$ is

$$W(t) = \lambda_A(t)^{-1}.$$

Since $\lambda_A(t)$ is unknown, we estimate using the muhaz package in R to find $\widehat{\lambda}_A(t)$.

A threshold, M, is chosen which provides computational stability when $\lambda_A(t)$ is estimated to be very low. So

$$\widetilde{W}(t) = \left[\max\{\widehat{\lambda}_A(t), M\}\right]^{-1} \qquad (1)$$

We find that setting M between 1/5 and 1/10 of the maximum value of $\lambda_A(t)$ provides adequate stabilization. Type 1 error is unaffected by choice of M, but power may be reduced if M is chosen to be too high.

Additive Null

Similarly, we can test

$$H_0 : \lambda_B(t) = \lambda_A(t) + \gamma \qquad \text{for some } \gamma \neq 0$$
$$H_1 : \lambda_B(t) = \Delta\lambda_A(t) + \gamma \qquad \text{for some } \Delta > 0.$$

We use the same model

$$\lambda_B(t) = \lambda_A(t)e^{\beta_1+\beta_2 W(t)},$$

noting that, if $\beta_1 = 0$ and $W(t) = \lambda_A(t)^{-1}$, then

$$\lambda_B(t) = \lambda_A(t)e^{\beta_2\lambda_A(t)^{-1}} \approx \lambda_A(t) + \beta_2.$$

Then a value for β_1 that is significantly different from 0 provides evidence to reject the additive model.

Again, we use $\widetilde{W}(t)$ as described in equation (1) for computational stability.

Simulation Results

Proportional Null

Distribution	$W(t)$	$n =$ 100	200	300
W(3,1/3)	$(\widehat{\lambda}_A(t) \wedge M)^{-1}$	31	57	76
	$(\lambda_A(t) \wedge M)^{-1}$	31	56	73
	$\widehat{S}(t)$	33	56	74
	cox.zph	30	53	73
LL(1,3)	$(\widehat{\lambda}_A(t) \wedge M)^{-1}$	14	46	70
	$(\lambda_A(t) \wedge M)^{-1}$	11	41	64
	$\widehat{S}(t)$	11	19	27
	cox.zph	11	19	27

Table 1: Empirical percent power based on 2000 replicates. We use Weibull and log-logistic distributions for group A with subjects equally split between the two groups and $\lambda_B(t) - \lambda_A(t) = 0.25$. Censoring is at time 3 and M is 0.15.

Additive Null

Distribution	$W(t)$	$n =$ 100	200	300
W(3,1/3)	$(\widehat{\lambda}_A(t) \wedge M)^{-1}$	50	88	98
	$(\lambda_A(t) \wedge M)^{-1}$	60	88	97
LL(1,3)	$(\widehat{\lambda}_A(t) \wedge M)^{-1}$	39	71	89
	$(\lambda_A(t) \wedge M)^{-1}$	40	84	96

Table 2: Setup is similar to Table 1, except $\lambda_B(t)/\lambda_A(t) = 2$ (i.e. the proportional hazards model is true).

References

[1] D. R. Cox.
Regression models and life-tables.
Journal of the Royal Statistical Society. Series B (Methodological), 34(2):pp. 187–220, 1972.

[2] David Schoenfeld.
Chi-squared goodness-of-fit tests for the proportional hazards regression model.
Biometrika, 67(1):pp. 145–153, 1980.

[3] L. J. Wei.
Testing goodness of fit for proportional hazards model with censored observations.
Journal of the American Statistical Association, 79(387):pp. 649–652, 1984.

[4] N. J. D. Nagelkerke, J. Oosting, and A. A. M. Hart.
A simple test for goodness of fit of cox's proportional hazards model.
Biometrics, 40(2):pp. 483–486, 1984.

[5] Richard Gill and Martin Schumacher.
A simple test of the porportional hazards assumption.
Biometrika, 74(2):289–300, 1987.

[6] Rick Chappell.
A note on linear rank tests and gill and schumacher's tests of proportionality.
Biometrika, 79(1):199–201, 1992.

[7] S. Vansteelandt, T. Martinussen, and E. J. Tchetgen Tchetgen.
On adjustment for auxiliary covariates in additive hazard models for the analysis of randomized experiments.
Biometrika, 101(1):237–244, 2014.

[8] Odd O. Aalen.
Linear regression model for the analysis of life times.
Statistics in Medicine, 8:907–925, 1989.

Contact Information

- Web: http://leemcdaniel.github.io
- Email: lmcda4@lsuhsc.edu
- Phone: +1 (504) 568-5771

LSUHealthNewOrleans HEALTH SCIENCES CENTER

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON