Randomized Survival Probability Residual for Assessing Parametric Survival Models

Tingxuan Wu

Supervisors: Longhai Li and Cindy Feng

Collaborative Biostatistics Program
School of Public Health
University of Saskatchewan

May 29, 2019

Outline

- Section 1: Introduction
- Section 2: Randomized Survival Probability Residual
- Section 3: Simulation Study
- Section 4: Real Data Analysis
- Section 5: Conclusion and Future Work

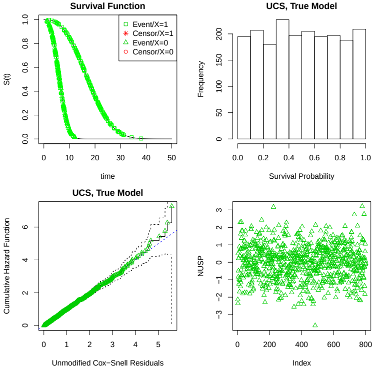
Section 1

Introduction

Residual Analysis

- Residual analysis is a standard practice in normal regression for answering these four questions:
  1. Graphical overall GOF checking, eg. qqplot
  2. Numerical overall GOF checking, eg. χ^2 test, normality test
  3. Model diagnosis for discovering directions (eg., non-linearity) for improving Models, eg. residual plots vs covariate
  4. Identifying outliers from an assumed model
- For survival models without censored observations, we can use Cox-Snell (CS) residuals. This is based on the theory that survival probabilities are uniformly distributed under the true model.
- When there are censored observations, the survival probabilities are no longer uniformly distributed, leading to difficulty in model diagnosis with CS residual and its variants.
- A novel residual based on Randomized Survival Probability (RSP) is proposed in our study for diagnosing survival models.

Cox-Snell Residuals for Uncensored Data



The green triangles: event times.

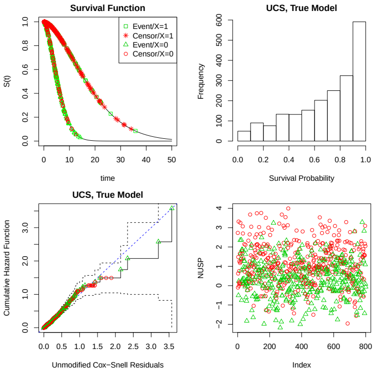
- $S_i(T_i) \sim$ Unif(0,1)
- Unmodified Cox-Snell (UCS) residual is a transformation of $S_i(T_i)$ with the quantile of exp(1), written as follows:

$$r_i^c = -\log S_i(T_i) \sim \exp(1) \qquad (1)$$

- The cumulative hazard function of r_i^C is a straight line with unit slope.
- Normally-transformed Unmodified Survival Probability (NUSP) residual:

$$r_i^{c\star} = \Phi^{-1}(S_i(T_i)) \sim N(0,1) \qquad (2)$$

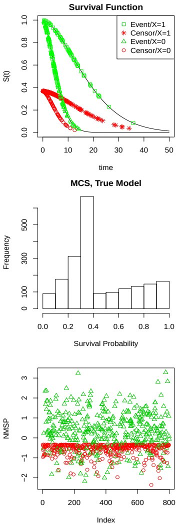
Cox-Snell Residual for Censored data



- With censored observations, the survival probabilities are no longer uniformly distributed.
- NUSP residual is no longer normally distributed.
- Cumulative hazard curve based on UCS residual can still be used to graphically check the GOF.
- However, we cannot answer other questions: numerical GOF testing, model diagnosis, outlier detection.

The green triangles: event times; the red circles: censored times.

Modified Cox-Snell Residuals



- Modified survival probabilities:

$$S_i'(T_i) = \begin{cases} S_i(T_i)/e & \text{censored } T_i \\ S_i(T_i) & \text{uncensored } T_i. \end{cases} \quad (3)$$

- $S_i'(T_i)$ is no longer uniformly distributed.
- Modified Cox-Snell (MCS) residual:

$$r^{c\prime}_i = -\log(S_i'(T_i)) = \begin{cases} r_i^c + 1 & \text{censored } T_i \\ r_i^c & \text{uncensored } T_i \end{cases} \quad (4)$$

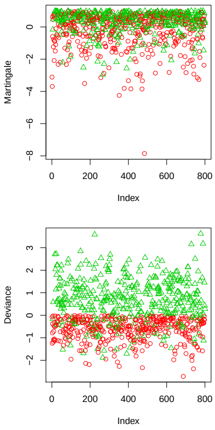
- Normal-transformed Modified Survival Probabilitiy (NMSP) residual:

$$r^{c*}_i = \Phi^{-1}(S_i'(T_i)). \quad (5)$$

The green triangles: event times, the red circles: censored times.

Other modifications: Martingale and Deviance Residuals

Residual plot



- The martingale residuals are defined as

$$r_i^M = \delta_i - r_i^c \tag{6}$$

where δ_i is the event indicator, and r_i^c is the UCS residual.

- The deviance residuals are defined as

$$r_i^D = sgn(r_i^M)[-2(r_i^M + \delta_i \log(\delta_i - r_i^M))]^{\frac{1}{2}} \tag{7}$$

where r_i^M is the martingale residual, the function $sgn(.)$ is the sign function.

The green triangles: event times, the red circles: censored times.

Necessity of Improving Residual Analysis for Censored Data

- When censored observations are present, the reference (null) distributions for modified CS residuals, martingale, deviance residuals under the true model are not clearly characterized.
- The Kaplan-Meier (KM) cumulative hazard function curve of unmodified Cox-Snell (UCS) is still valid since the KM method can account for the censoring effect. We can check the overall GOF by visually comparing the curve against the straight line with 45° angle.
- However, the cumulative hazard curve cannot be used to answer other questions demanded in practice: numerical overall GOF checking, model diagnosis with residual plot, and outlier detection.

Section 2

Randomized Survival Probability Residual

Definition of Randomized Survival Probability Residual

Let $S(T_i)$ be the survival function, if the random survival time T_i has survival function $S_i(T_i)$, then $S_i^*(T_i; u_i)$ is defined as follows:

$$S^*(T_i; u_i) = \begin{cases} S_i(T_i^*) & T_i^* < C_i(T_i \text{ is event time}) \\ u_i S_i(C_i) & T_i^* \geq C_i(T_i \text{ is censored time}) \end{cases} \tag{8}$$

where u_i is a uniform random variable on (0, 1]. The Randomized Survival Probability (RSP) follows:

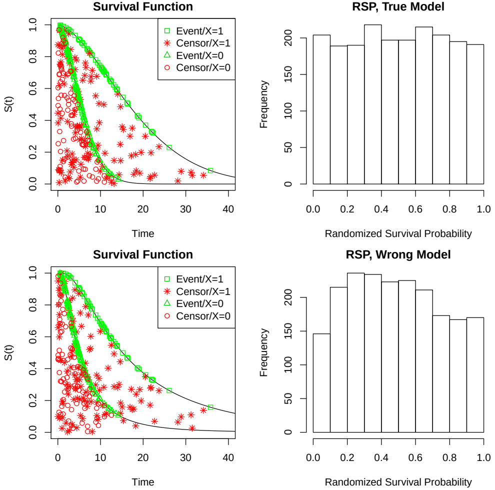
$$S^*(T_i; u_i) \sim Uniform(0, 1) \tag{9}$$

Then, the RSP residuals are defined as

$$q_i = q(T_i; u_i) = \Phi^{-1}(S^*(T_i; u_i)) \sim N(0, 1) \tag{10}$$

where $\Phi()$ is the cumulative distribution function (CDF) of a standard normal distribution.

Illustrative Examples: Assessing Distributional Assumption for Survival Time



- RSP is uniformly distributed under the true model (first row), and
- RSP is no longer uniformly distributed under the wrong model (second row).

Section 3

Simulation Study

Accelerated Failure Time Regression Model

- Accelerated failure time (AFT) model is one type of survival model. Weibull, log-logistic and log-normal distributions for the survival time are most commonly used in AFT models.
- The log-linear form of the AFT model is given by

$$\log T_i = \mu + a_1 X_{1i} + a_2 X_{2i} + \ldots + a_p X_{pi} + \sigma \epsilon_i, \tag{11}$$

  where μ is intercept, σ is scale parameter, ϵ_i is a random variable and $a_1, \cdots, a_p$ represent the effects of covariates on the survival time.
- The survival function based on an AFT model is written as

$$S(t|X) = S_0\left(\frac{t}{\exp \eta(X)}\right) \tag{12}$$

  where $S_0(t)$ is the baseline survival function and the acceleration factor is given according to the formula $\exp(\eta(X)) = \exp(a_1 X_1 + a_2 X_2 + \ldots + a_p X_p)$.

Simulation Setup

- Covariate x_i is simulated from Bern(0.5).
- Censored time C_i is simulated from Exp(13).
- **True model:**

$$\log(T_i^*) = \beta_0 + \beta_1 x_i + \epsilon_i \tag{13}$$

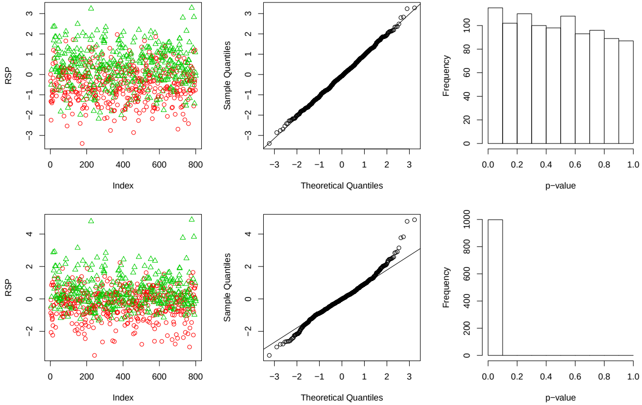
  where $\epsilon_i \sim$ standard extreme value distribution (γ), and real life time $T_i^* \sim$ Weibull (λ, γ). We set $\beta_0 = 2$, $\beta_1 = 1$. The sample size n is 800 and the censorship c is 50%.
- $T_i = \min(T_i^*, C_i)$ is the observed survival time.
- **Wrong model:**

$$\log(T_i^*) = \beta_0 + \beta_1 x_i + \epsilon_i \tag{14}$$

  where $\epsilon_i \sim N(\mu, \sigma^2)$, and real life time $T_i^* \sim$ lognormal $(\mu + \beta_1 x_i, \sigma)$.

RSP Residual for a Single Dataset

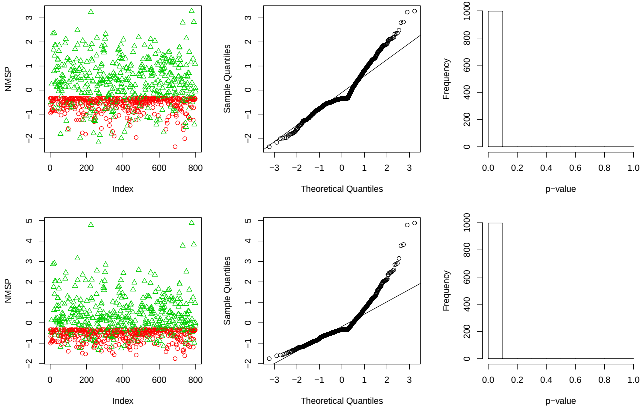
Residual plot | QQ plot | SW p-values



The green triangles: event times; the red circles: censored times.

NMSP Residual for a Single Dataset

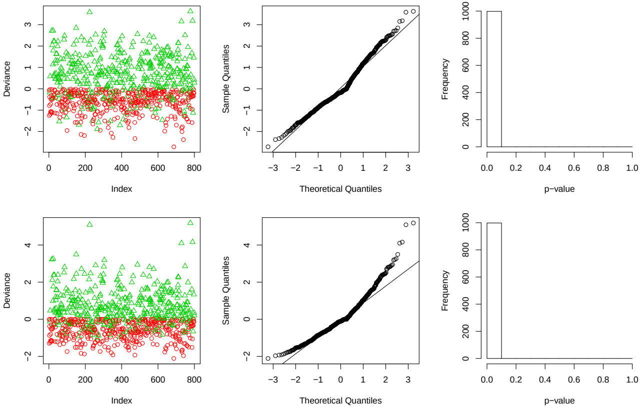
Residual plot　　QQ plot　　SW p-values



The green triangles: event times; the red circles: censored times.

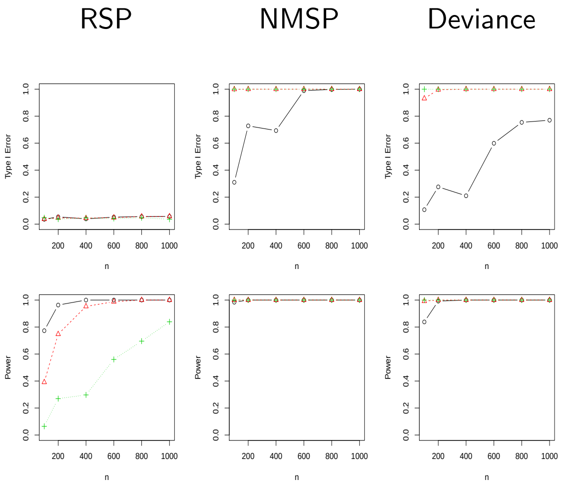
Deviance Residual for a Single Dataset

Residual plot QQ plot SW p-values



The green triangles: event times; the red circles: censored times.

Power Analysis



Comparison of the type I errors and powers of the SW tests for the RSP, NMSP, and deviance residuals. Response variable is simulated from the true model at varying sample sizes of $n = 100$, 200, 400, 600, 800 and 1000, and the percentage of censorship $c = 20\%$ (black circles), 50% (red triangles) and 80% (green crosses).

Section 4

Real Data Analysis

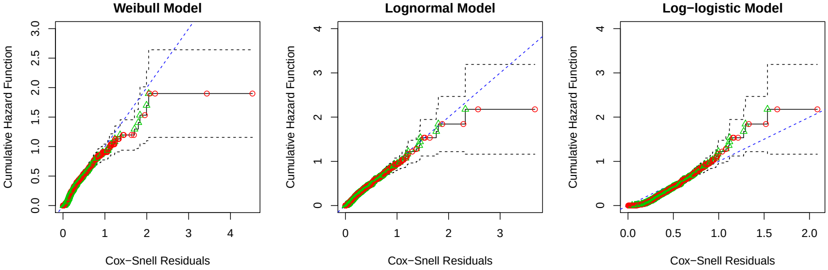
Data Description

- A real data analysis was conducted using a survey dataset on 686 individuals with 56.5% censorship, which was for studying the recurrence-free survival in breast cancer patients.
- The response variable of interest is recurrence-free survival, which is the time from entry to the study until a recurrence of the cancer or death.
- Weibull, Log-logistic and Lognormal AFT models with all of variables included as covariates are fitted to this data.
- RSP residual is applied to examine the goodness-of-fits of the fitted AFT models in comparison with other types of residuals.

List of Variables

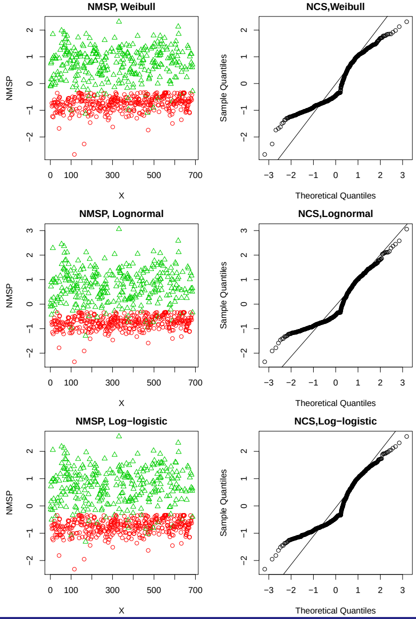
Variable	Definition
Time	Recurrence-free survival time (days)
Status	Event indicator (0 = censored, 1 = relapse or death)
Treat	Hormonal treatment (0 = no tamoxifen, 1 = tamoxifen)
Age	Patient age (years)
Men	Menopausal status (1 = premenopausal, 2 = postmenopausal)
Size	Tumour size (mm)
Grade	Tumour grade (1, 2, 3)
Nodes	Number of positive lymph nodes
Prog	Progesterone receptor status (femtomoles)
Oest	Oestrogen receptor status, (femtomoles)

UCS Residual for the Fitted Models



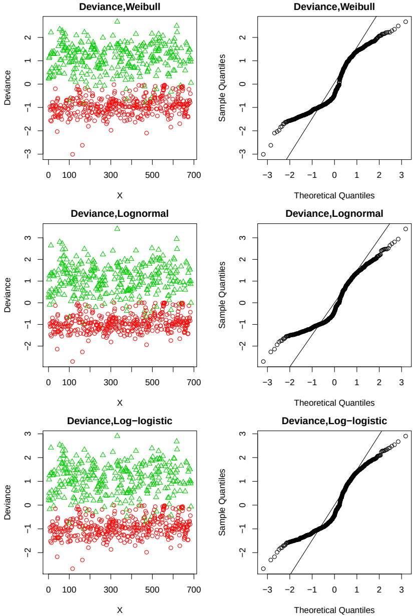
Cox-Snell residuals for Weibull, Log-logistic, and Lognormal AFT models fitted to the breast cancer patients dataset. The green triangles correspond to the event times and the red circles correspond to the censored times.

NMSP Residual for the Fitted Models



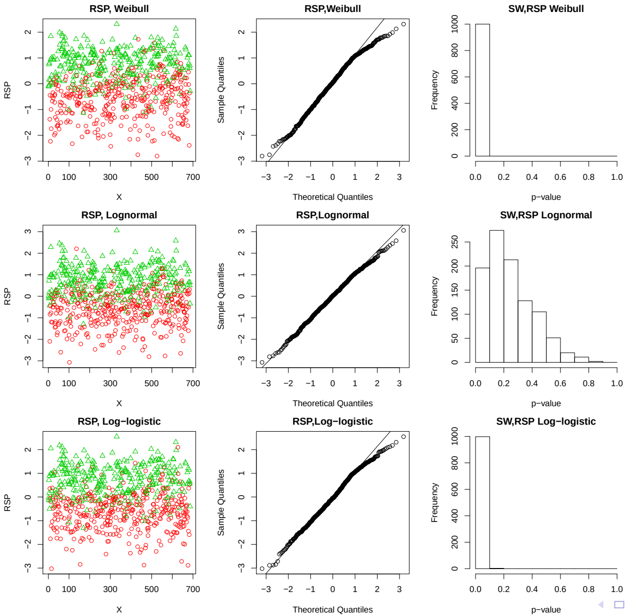
- The p-values of the SW test $< 10^{-3}$ for all of models.
- NMSP residuals for the Weibull, Log-logistic, and Lognormal AFT models fitted to the breast cancer patients dataset. The green triangles correspond to the event times and the red circles correspond to the censored times.

Deviance Residual for the Fitted Models



- The p-values of the SW test $< 10^{-3}$ for all of models.
- Deviance residuals for the Weibull, Log-logistic, and Lognormal AFT models fitted to the breast cancer patients dataset. The green triangles correspond to the event times and the red circles correspond to the censored times.

RSP Residual for the Fitted Models



RSP residuals for the Weibull, Log-logistic, and Lognormal AFT models fitted to the breast cancer patients dataset. The green triangles correspond to the event times and the red circles correspond to the censored times.

Sensitivity Analysis of Randomize in RSP

Model	Weibull	Lognormal	Log-logistic
RSP	100%	7.8%	99.3%
AIC	5182	5140	5154

Percentages of P-values smaller than 0.05 for the SW test of the RSP residuals and AIC comparisons for Weibull, Log-normal and Log-logistic models in the breast cancer data analysis.

Section 5

Conclusion and Future Work

Conclusion

Model diagnosis in survival analysis is very important but difficult. This study empirically demonstrates that RSP residual is an excellent tool that can be used to diagnose AFT models. We have empirically demonstrated that

- under any true model, RSP residuals are normally distributed.
- the overall GOF tests by applying normality test to RSP residuals are well-calibrated.
- RSP residual has great power in detecting many kinds of model inadequacy.

Future Work

For further study, RSP residuals will be extended to diagnose AFT models with random effects. We will extend RSP residuals to diagnose proportional hazard models with and without random effects.