Prileganje regresijskih modelov za relativno preživetje
Goodness of fit of relative survival models

Maja Pohar Perme

doktorsko delo

mentor prof. dr. Janez Stare

Ljubljana, marec 2007

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Additive model

Multiplicative model

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Additive model

$$\lambda_O = \lambda_P + \lambda_E$$

Multiplicative model

$$\lambda_O = \lambda_P * \nu$$

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Additive model

$$\lambda_O = \lambda_P + \lambda_E$$

Multiplicative model

$$\lambda_O = \lambda_P * \nu$$

assumptions

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Additive model

$$\lambda_O = \lambda_P + \lambda_E$$

Multiplicative model

$$\lambda_O = \lambda_P * \nu$$

assumptions

1. Avoid assumptions about

relationship between λ_O and λ_P

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Additive model

$$\lambda_O = \lambda_P + \lambda_E$$

Multiplicative model

$$\lambda_O = \lambda_P * \nu$$

assumptions

1. Avoid assumptions about

relationship between λ_O and λ_P

2. Check

proportional hazards assumption

Objectives

The goal of relative survival regression

To estimate the effect of covariates on mortality caused by a disease even though the cause of death is unknown.

Additive model

$$\lambda_O = \lambda_P + \lambda_E$$

Multiplicative model

$$\lambda_O = \lambda_P * \nu$$

assumptions

1. Avoid assumptions about

relationship between λ_O and λ_P

2. Check

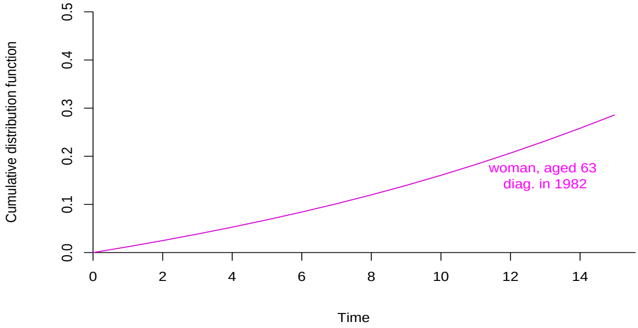
proportional hazards assumption

3. Develop software

for all existing and new methods

Transformation approach

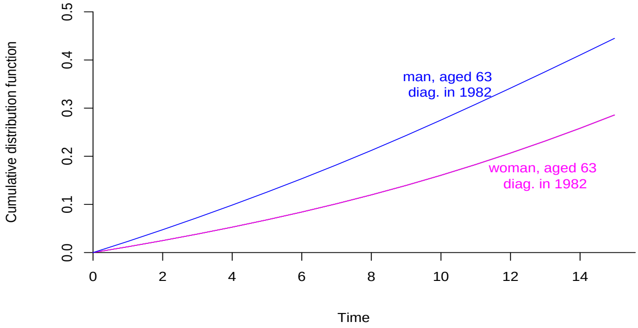
published: Stare, Henderson, Pohar, JRSS C, 2005



$$F_P$$

Transformation approach

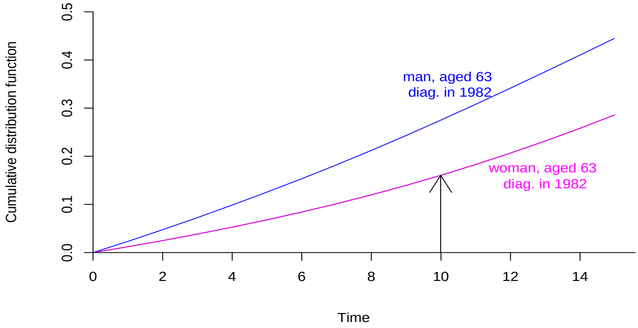
published: Stare, Henderson, Pohar, JRSS C, 2005



F_P

Transformation approach

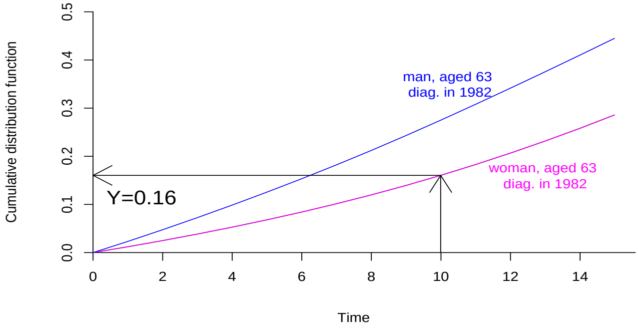
published: Stare, Henderson, Pohar, JRSS C, 2005



$$F_P(T)$$

Transformation approach

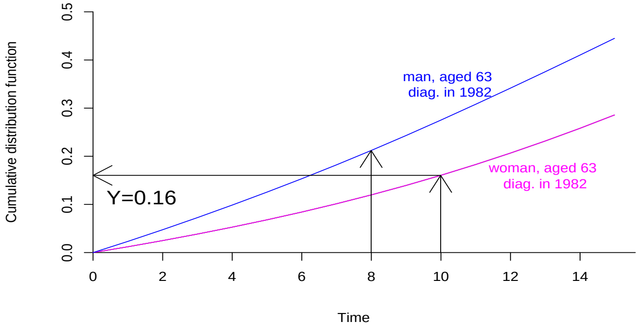
published: Stare, Henderson, Pohar, JRSS C, 2005



$$Y = F_P(T)$$

Transformation approach

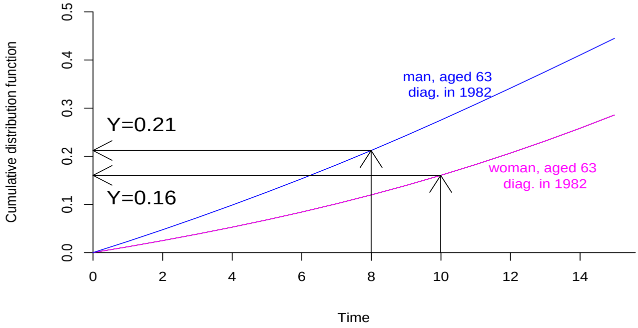
published: Stare, Henderson, Pohar, JRSS C, 2005



$$Y = F_P(T)$$

Transformation approach

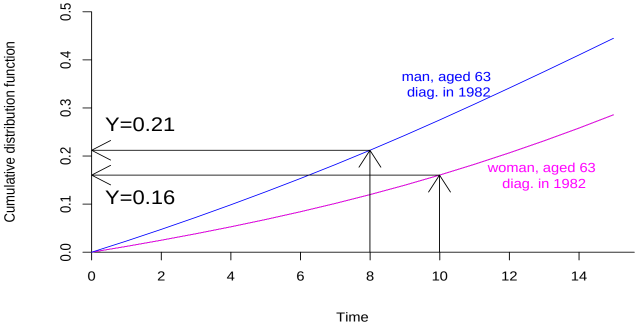
published: Stare, Henderson, Pohar, JRSS C, 2005



$$Y = F_P(T)$$

Transformation approach

published: Stare, Henderson, Pohar, JRSS C, 2005



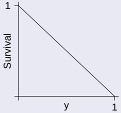
New outcome variable

$$Y = F_P(T)$$

Properties

Properties

- **New outcome variable. Censoring status and covariate values remain unchanged**
- **Any survival analysis method can be used**
- $T \sim F_P \Rightarrow Y \sim U[0, 1]$

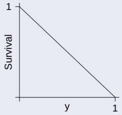


- **Patients can live better than the population**

Properties

Properties

- **New outcome variable. Censoring status and covariate values remain unchanged**
- **Any survival analysis method can be used**
- $T \sim F_P \Rightarrow Y \sim U[0,1]$



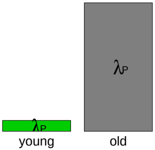
- **Patients can live better than the population**

Cox model $\lambda(y) = \lambda_0(y)e^{bx}$

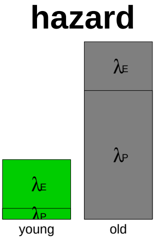
- **natural choice - new time ordering**
- **the covariates not included in the population tables keep the same coefficient**

Comparison of the models - interpretation

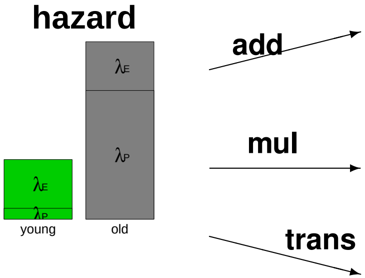
hazard



Comparison of the models - interpretation

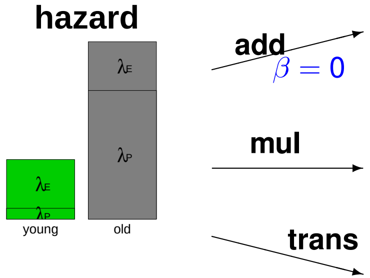


Comparison of the models - interpretation



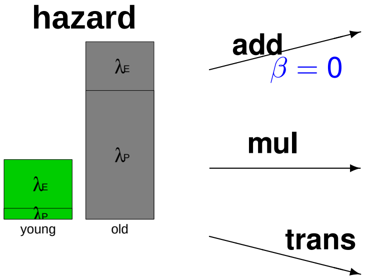
does excess hazard differ?

Comparison of the models - interpretation



does excess hazard differ?

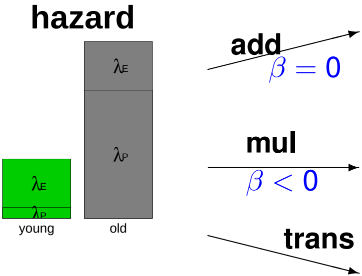
Comparison of the models - interpretation



does excess hazard differ?

does the observed hazard ratio differ from the population hazard ratio?

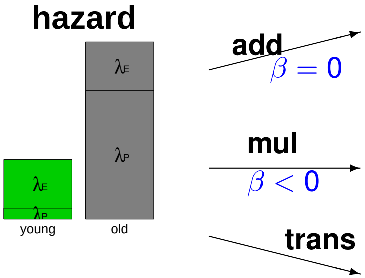
Comparison of the models - interpretation



does excess hazard differ?

does the observed hazard ratio differ from the population hazard ratio?

Comparison of the models - interpretation

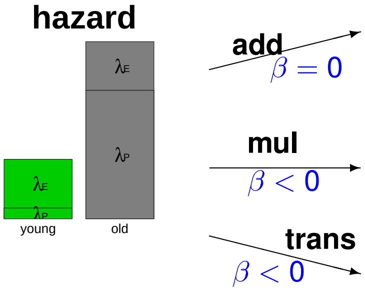


does excess hazard differ?

does the observed hazard ratio differ from the population hazard ratio?

are residual lifetimes equally affected?

Comparison of the models - interpretation



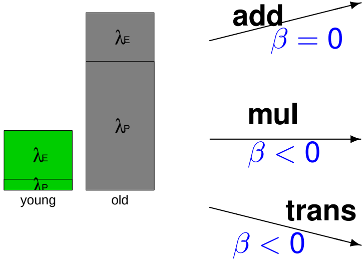
does excess hazard differ?

does the observed hazard ratio differ from the population hazard ratio?

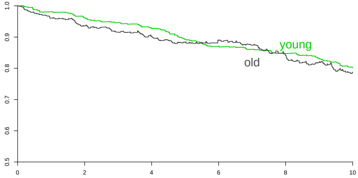
are residual lifetimes equally affected?

Comparison of the models - interpretation

hazard



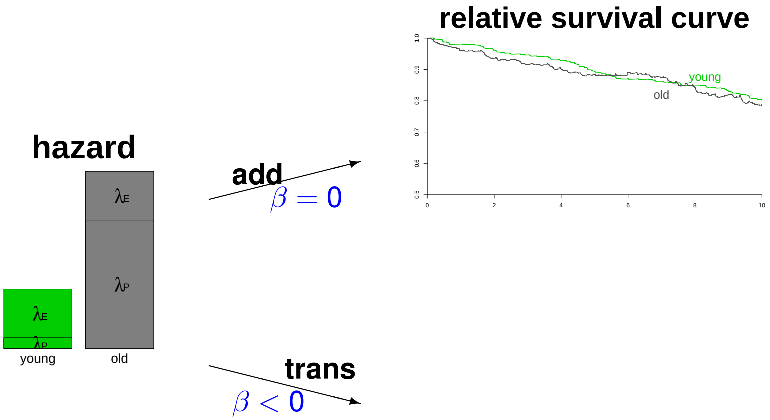
relative survival curve



does the observed hazard ratio differ from the population hazard ratio?

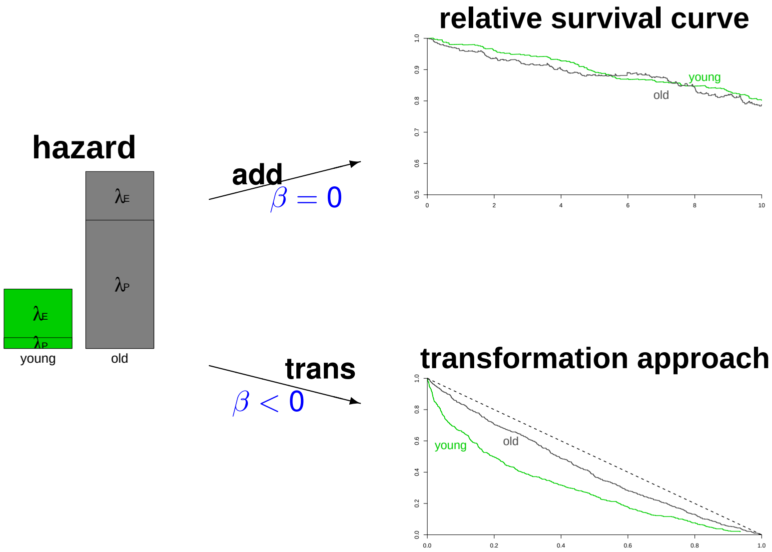
are residual lifetimes equally affected?

Comparison of the models - interpretation

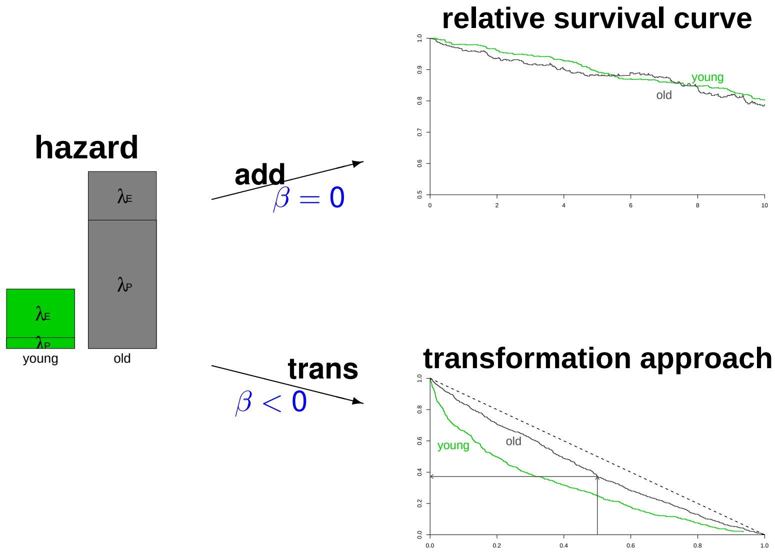


are residual lifetimes equally affected?

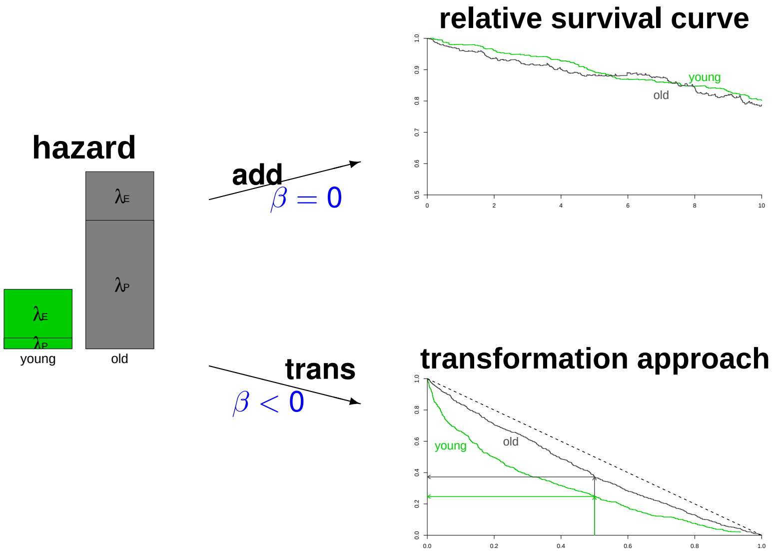
Comparison of the models - interpretation



Comparison of the models - interpretation



Comparison of the models - interpretation



Comparison of the models

The coefficient values (for the covariates used in the population tables) estimated in different models are different and require a different interpretation

- $T \sim F_P \Rightarrow$ **results of all models are equal**
- $\nu_0 = const. \Rightarrow$ **Cox in Y=multiplicative**
- **general case $\Rightarrow$ all models different, but multiplicative and Cox in Y usually closer. The proportional hazards assumption is not met simultaneously.**

Transformation approach

- **Avoids the assumption about the relationship between the observed and population hazard**
- **Provides new information about relative survival**
- **Only the outcome variable changes - all methods from the classical survival analysis can be used**
- **Can be used when certain groups of patients live better than the population (additive model can not!)**

Checking the proportional hazards assumption

published: Stare, Pohar, Henderson; SIM 2005

The Cox model

- **The multiplicative model and the Cox model in transformed time can both be seen as special cases of the Cox model**
- **Schoenfeld residuals can be used for both graphical and formal evaluation of the PH assumption**

The additive model

- **No methods exist**
- **A new kind of residuals is introduced - the partial residuals**

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X\|t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X\|t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$
Properties	$\int HdM$	✓	✓

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X\|t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$
Properties	$\int HdM$	✓	✓
	score function	✓	**X**

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X\|t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$
Properties	$\int HdM$	✓	✓
	score function	✓	**X**
Graphical inspection		✓	✓

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X \mid t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$
Properties	$\int HdM$	✓	✓
	score function	✓	**X**
Graphical inspection		✓	✓
Convergence to Brownian bridge	**in** β^0	✓	✓

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X\|t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$
Properties	$\int HdM$	✓	✓
	score function	✓	**X**
Graphical inspection		✓	✓
Convergence to Brownian bridge	**in** β^0	✓	✓
	in $\hat{\beta}$	✓	**X**

Ideas to be modified for use in relative survival

		Cox	additive
Definition	$U_i = X_i - \hat{E}_i(X\|t_i)$	$X_i - \sum X_j \frac{\lambda_j}{\sum \lambda_k}$	$X_i - \sum X_j \frac{\lambda_{Pj}+\lambda_{Ej}}{\sum(\lambda_{Pk}+\lambda_{Ek})}$
Properties	$\int HdM$	✓	✓
	score function	✓	**X**
Graphical inspection		✓	✓
Convergence to Brownian bridge	**in** β^0	✓	✓
	in $\hat{\beta}$	✓	**X**
	$\#$ **covariates** > 1	**X**	**X**

The quality of Brownian bridge approximation

In theory

- **Asymptotic distribution exists**
- **Taylor series expansion:** Brownian bridge + residual
- **Residual size depends on: variance of the covariate in time**

The quality of Brownian bridge approximation

In theory

- **Asymptotic distribution exists**
- **Taylor series expansion:** Brownian bridge + residual
- **Residual size depends on: variance of the covariate in time**

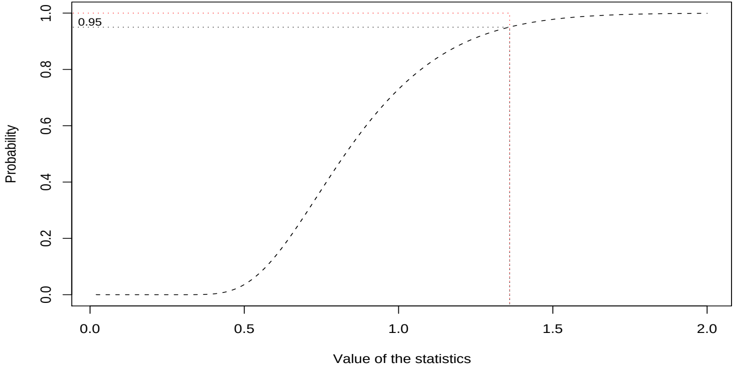
In practice

- **The proportion of tests rejecting under the null hypothesis ($\alpha = 0.05$): 0.03-0.05**
- **The test statistic can only be conservative**
- **Resampling approach (resample fom the distribution defined by the covariates and estimated coefficients)**

Speed of convergence

- Comparable to the ideal case (i.i.d. variables)
- The maximum (weighted) BB statistic: 100 (conservative otherwise)
- The Cramér - Von Mises statistic: 50
- Effective sample size depends on: censoring, baseline excess hazard

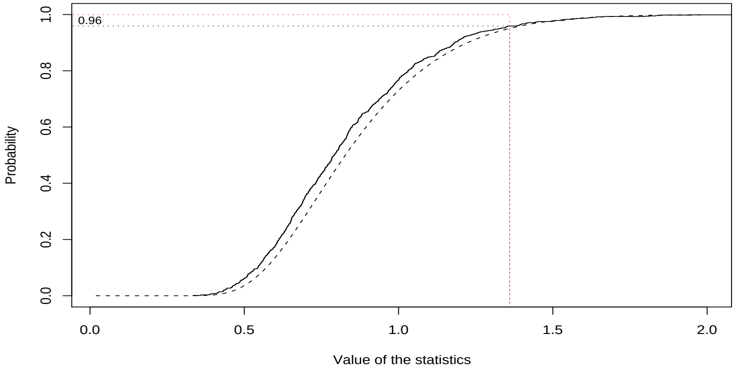
Power


0.95
Probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Value of the statistics

Power

sample size ↑ ↓ $n = 250$ **rejected**$= 4.1\%$

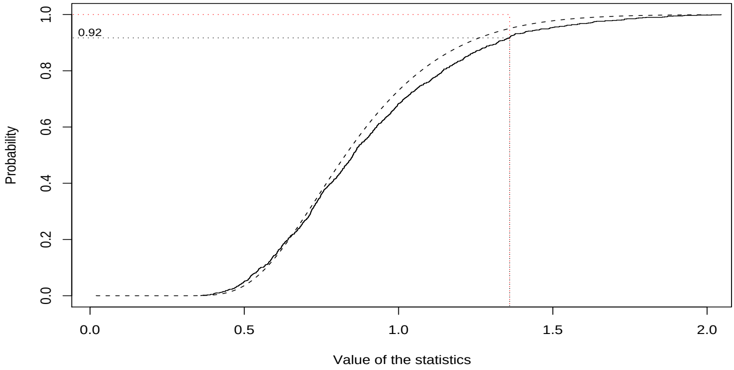
change in β ↑ ↓ $\Delta = 0$



Power

sample size ↑ ↓ $n = 250$ **rejected**$= 8.4\%$

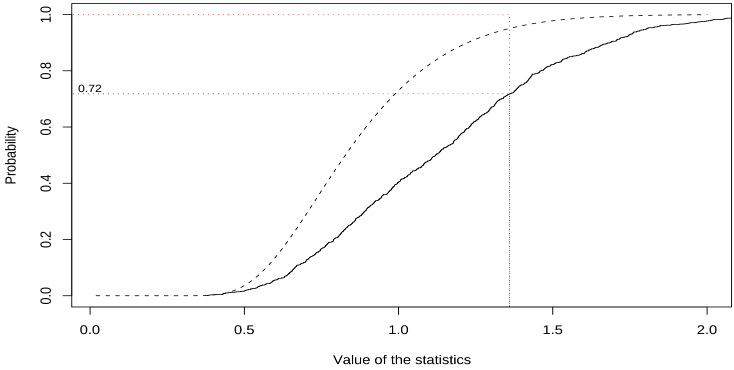
change in β ↑ ↓ $\Delta = 0.25$



Power

sample size ↑ ↓ $n = 250$ **rejected**$= 28.4\%$

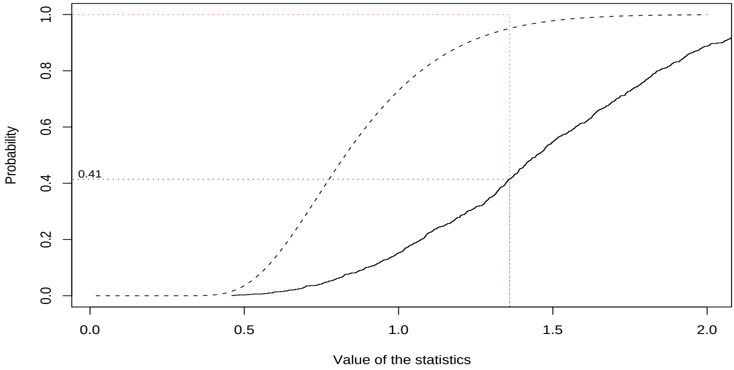
change in β ↑ ↓ $\Delta = 0.5$



Power

sample size ↑ ↓ $n = 250$ **rejected**$= 58.7\%$

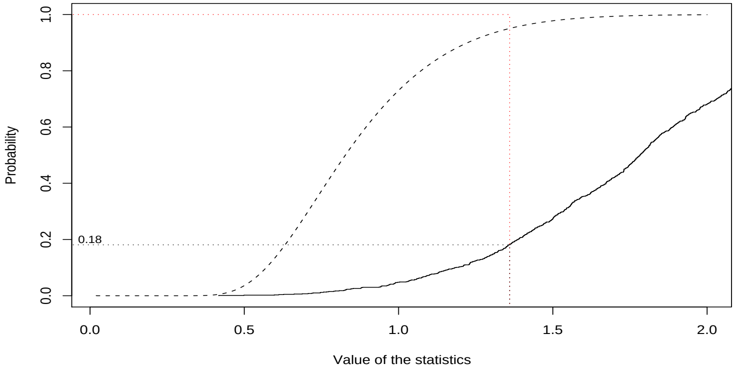
change in β ↑ ↓ $\Delta = 0.75$



Power

sample size ↑ ↓ $n = 250$ **rejected**$= 81.9\%$

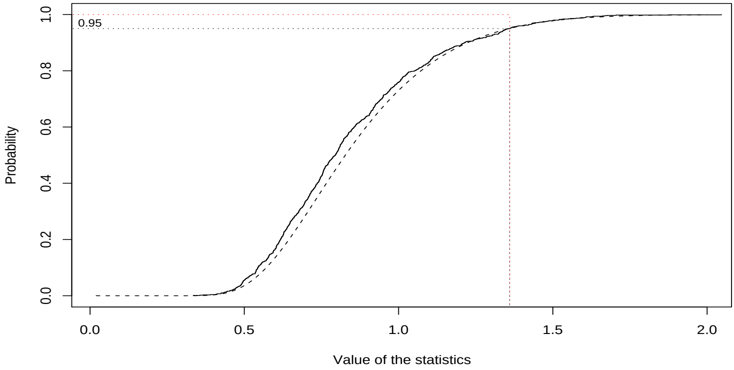
change in β ↑ ↓ $\Delta = 1$



Power

sample size ↑ ↓ $n = 500$ **rejected**$= 5\%$

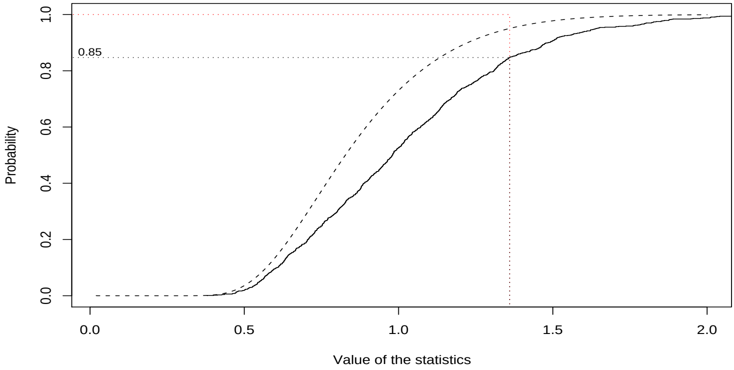
change in β ↑ ↓ $\Delta = 0$



Power

sample size ↑ ↓ $n = 500$ **rejected**$= 15.4\%$

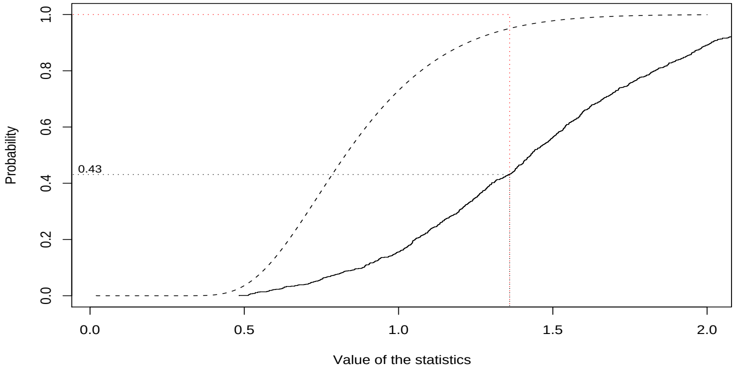
change in β ↑ ↓ $\Delta = 0.25$



Power

sample size ↑ ↓ $n = 500$ **rejected**$= 56.9\%$

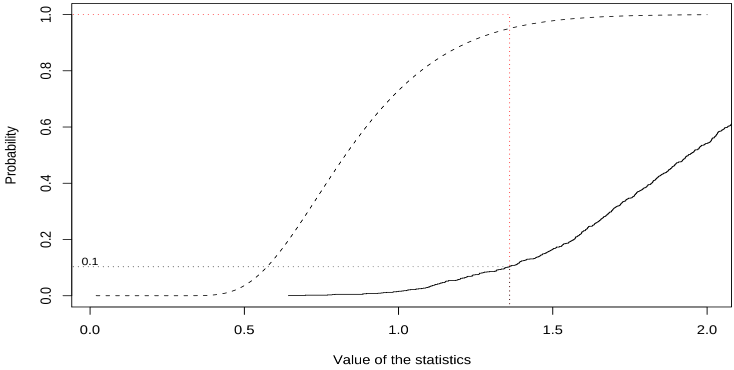
change in β ↑ ↓ $\Delta = 0.5$



Power

sample size ↑ ↓ $n = 500$ **rejected**$= 89.8\%$

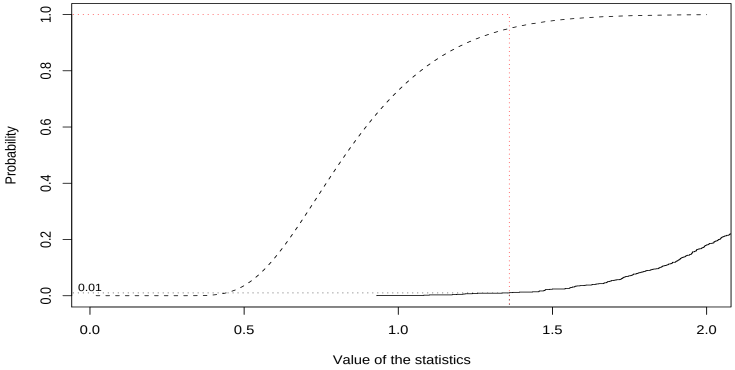
change in β ↑ ↓ $\Delta = 0.75$



Power

sample size ↑ ↓ $n = 500$ **rejected**$= 99\%$

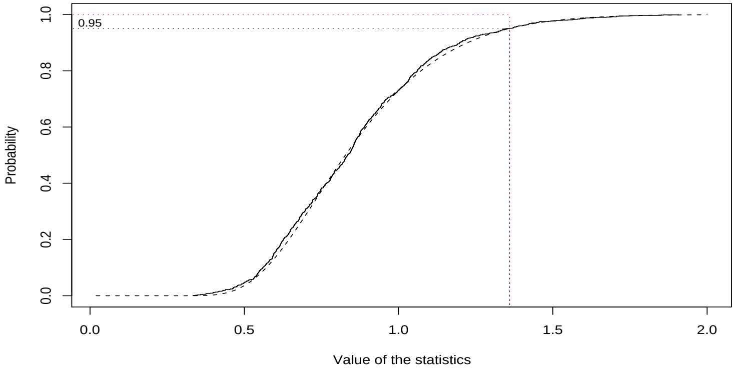
change in β ↑ ↓ $\Delta = 1$



Power

sample size ↑ ↓ $n = 1000$ **rejected**$= 5\%$

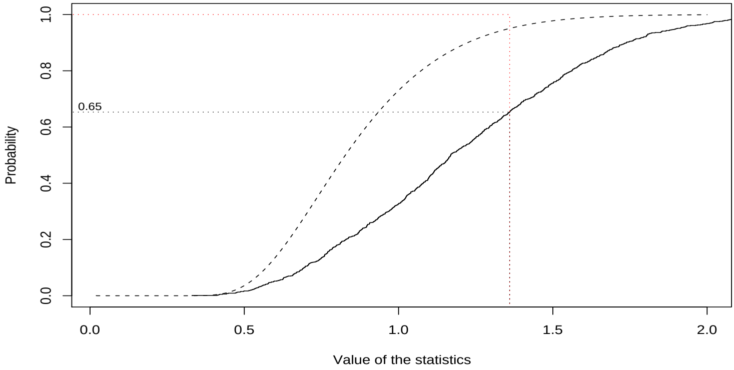
change in β ↑ ↓ $\Delta = 0$



Power

sample size ↑ ↓ $n = 1000$ **rejected**$= 34.9\%$

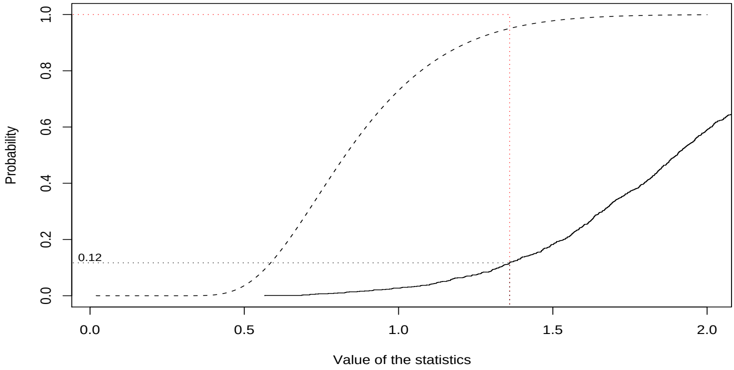
change in β ↑ ↓ $\Delta = 0.25$



Power

sample size ↑ ↓ $n = 1000$ **rejected**$= 88.7\%$

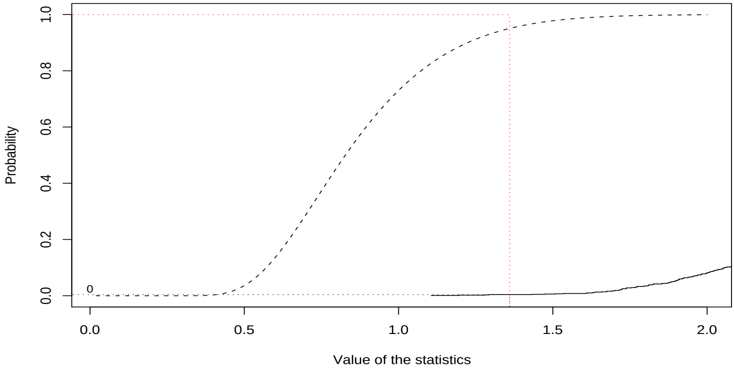
change in β ↑ ↓ $\Delta = 0.5$



Power

sample size ↑ ↓ $n = 1000$ **rejected**$= 99.6\%$

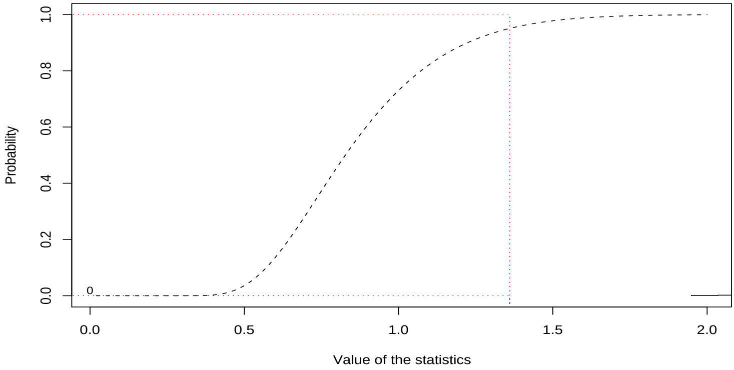
change in β ↑ ↓ $\Delta = 0.75$



Power

sample size ↑ ↓ $n = 1000$ **rejected**$= 100\%$

change in β ↑ ↓ $\Delta = 1$



To sum up

- The partial residuals are useful in checking the PH assumption in the additive model
- Theoretical deficiencies are not important for practical use - the Brownian bridge theory can be used
- The choice of the test statistic should be based on the alternative hypothesis
- The Cramér - Von Mises statistic more appropriate with small sample sizes

relsurv package

published: Pohar, Stare, CMPB 2006; CRAN

The package ensures easy use of relative survival methods

- **provides a uniform syntax for all the models using any format of population tables**

  ```
  > rs.fun(Surv(time,cens) ~ variables,data,ratetable)
  ```

  `rsadd`, `rstrans`, `rsmul`
- **simplifies transformation of population tables into R** **`transrate.hld`, `transrate.hmd`, `transrate`, `joinrate`**
- **functions for checking goodness of fit** **`rs.zph`, `rs.br`**
- **provides methods for plotting results** **`plot.rsurvfit`, `plot.rs.zph`, `plot.rs.br`**

relsurv package

- **available at CRAN**
- **the most complete and flexible package for relative survival**
- **usable with any format of population tables**
- **thoroughly checked and compared to results in other software**
- **enriched by the options required by the users**

Bibliography

- **Stare J., Henderson R., Pohar M.**
  An individual measure of relative survival
  Journal of Royal Statistical Society – C, 2005
- **Stare J., Pohar M., Henderson R.**
  Goodness of fit of relative survival models
  Statistics in Medicine, 2005
- **Pohar M., Stare J.**
  Relative survival analysis in R
  Computer methods and programs in biomedicine, 2006

Definition of residuals

Cox model
Schoenfeld residuals

X_i

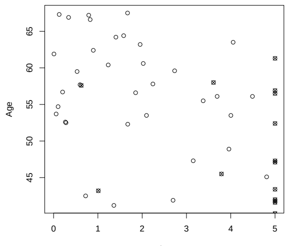
additive model
partial residuals

X_i

Notation

X **covariate**

t_i ***i*th event time**



Definition of residuals

Cox model
Schoenfeld residuals

$$X_i \quad \hat{E}(X, t_i)$$

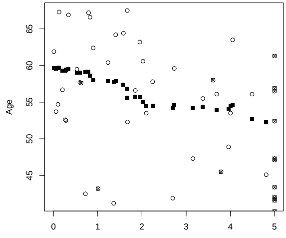
additive model
partial residuals

$$X_i \quad \hat{E}(X, t_i)$$

Notation

X covariate

t_i ith event time



Definition of residuals

Cox model
Schoenfeld residuals

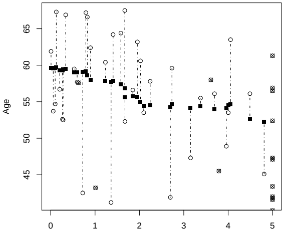
$$X_i - \hat{E}(X, t_i)$$

additive model
partial residuals

$$X_i - \hat{E}(X, t_i)$$

Notation

- X **covariate**
- t_i i**th event time**



Definition of residuals

Cox model
Schoenfeld residuals

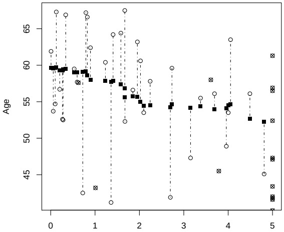
$$U_i \quad = \quad X_i - \hat{E}(X, t_i)$$

additive model
partial residuals

$$U_i : \quad = \quad X_i - \hat{E}(X, t_i)$$

Notation

X	**covariate**
t_i	***i*th event time**
U_i	**residual**



Definition of residuals

Cox model
Schoenfeld residuals

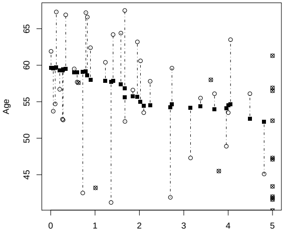
$$U_i = X_i - \hat{E}(X, t_i) = X_i - \sum_{j \in R_i} X_j \frac{\lambda_j}{\sum_{k \in R_i} \lambda_k}$$

additive model
partial residuals

$$U_i := X_i - \hat{E}(X, t_i) = X_i - \sum_{j \in R_i} X_j \frac{\lambda_{Pj} + \lambda_{Ej}}{\sum_{k \in R_i} (\lambda_{Pk} + \lambda_{Ek})}$$

Notation

- X covariate
- t_i ith event time
- U_i residual
- R_i risk set at time i
- λ_j hazard for person j
- λ_P population hazard
- λ_E excess hazard



Properties

Residuals can be expressed as martingales

$$
\begin{aligned}
U(\beta^0, t) &= \sum_{i=1}^{n} \int_0^t \left\{ \mathbf{X_i}(u) - \hat{E}(X|u, \beta^0) \right\} dN_i(u) \\
&= \sum_{i=1}^{n} \int_0^t \underbrace{\left\{ \mathbf{X_i}(u) - \hat{E}(X|u, \beta^0) \right\}}_{\textbf{predictable process}} \underbrace{dM_i(u)}_{\textbf{martingale}}
\end{aligned}
$$

Properties

Residuals can be expressed as martingales

$$
\begin{aligned}
U(\beta^0, t) &= \sum_{i=1}^{n} \int_0^t \left\{ \mathbf{X_i}(u) - \hat{E}(X|u, \beta^0) \right\} dN_i(u) \\
&= \sum_{i=1}^{n} \int_0^t \underbrace{\left\{ \mathbf{X_i}(u) - \hat{E}(X|u, \beta^0) \right\}}_{\text{predictable process}} \underbrace{dM_i(u)}_{\text{martingale}}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E\mathbf{U}(\beta^0, t) &= \mathbf{0} \\
E\mathbf{U_i}(\beta^0, t) &= \mathbf{0} \\
\operatorname{cov}\left(\mathbf{U_i}(\beta^0, \mathrm{t}), \mathbf{U_j}(\beta^0, \mathrm{t})\right) &= \mathbf{0}
\end{aligned}
$$

variance of the residual process can be computed

Properties

Residuals can be expressed as martingales

$$
\begin{aligned}
U(\beta^0, t) &= \sum_{i=1}^{n} \int_0^t \left\{ \mathbf{X_i}(u) - \hat{E}(X|u, \beta^0) \right\} dN_i(u) \\
&= \sum_{i=1}^{n} \int_0^t \left\{ \underbrace{\mathbf{X_i}(u) - \hat{E}(X|u, \beta^0)}_{\text{predictable process}} \right\} \underbrace{dM_i(u)}_{\text{martingale}}
\end{aligned}
$$

Therefore,

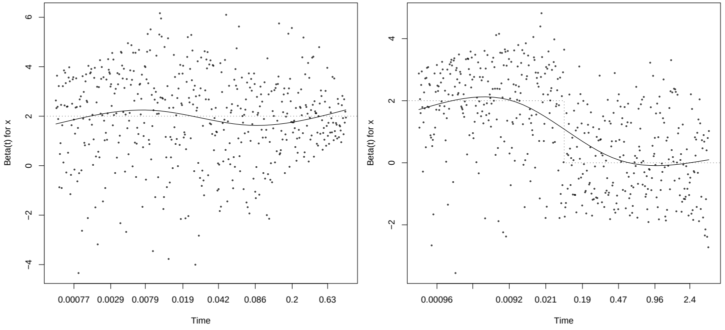
$$
\begin{aligned}
E\mathbf{U}(\beta^0, t) &= \mathbf{0} \\
E\mathbf{U_i}(\beta^0, t) &= \mathbf{0} \\
\mathrm{cov}\left(\mathbf{U_i}(\beta^0, \mathrm{t}), \mathbf{U_j}(\beta^0, \mathrm{t})\right) &= \mathbf{0}
\end{aligned}
$$

variance of the residual process can be computed

Schoenfeld residuals follow from the score function

$U(\hat{\beta}, \infty) = \sum_{i=1}^{n} U_i(\hat{\beta}) = 0$ **This is not true in the additive model case!**

Graphical inspection



$$\beta^0(t_i) \simeq \hat{\beta} + \left(\frac{\partial}{\partial\beta}\{\hat{E}(X|t_i,\hat{\beta})\}\right)^{-1} E[U(\hat{\beta},t)]$$

Brownian motion constructed as the sum of residuals

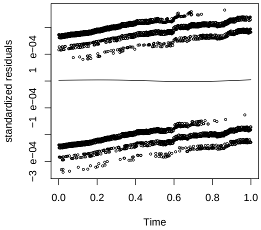
Cox model & additive model

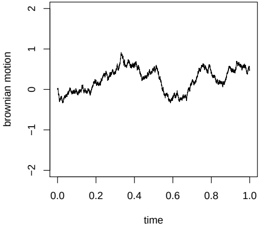
$$B(\beta_0, \frac{k}{n}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{k} \frac{U_i(\beta_0)}{\sqrt{V_i(\beta_0)}} \quad \overset{n\to\infty}{\longrightarrow} \quad \text{Brownian motion}$$

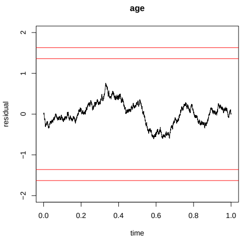
$$BB(\beta_0, \frac{k}{n}) = B(\beta_0, \frac{k}{n}) - \frac{k}{n} B(\beta_0, 1) \overset{n\to\infty}{\longrightarrow} \text{Brownian bridge}$$

Notation

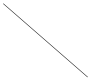
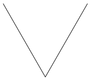
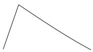
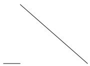
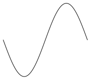
- U_i **Schoenfeld-like residuals**
- V **variance**
- n **number of deaths**
- β_0 **true regression coefficient**







Tests based on Brownian bridge properties

β^0 *in time*	*brownian bridge process*	*test statistic*
		T_1 $max(abs(BB(t))$
		T_2 *max using weighted residuals*
		T_3 Cramer–Von Mises $\int_0^1 BB^2(t)dt - (\int_0^1 BB(t)dt)^2$