

ON A ROBUST SIGNIFICANCE TEST FOR THE COX REGRESSION MODEL

TADEUSZ BEDNARSKI AND FILIP BOROWICZ

Institute of Economic Sciences
Wrocław University
Pl. Uniwersytecki 1, 50–137 Wrocław, Poland
e-mail: bednarski@prawo.uni.wroc.pl
e-mail: borowicz@prawo.uni.wroc.pl

Abstract

A robust significance testing method for the Cox regression model, based on a modified Wald test statistic, is discussed. Using Monte Carlo experiments the asymptotic behavior of the modified robust versions of the Wald statistic is compared with the standard significance test for the Cox model based on the log likelihood ratio test statistic.

Keywords: robust estimation Wald test.

2000 Mathematics Subject Classification: 62F35, 62N03.

1. INTRODUCTION

The statistical relationship between survival time T and the vector of covariates Z , in the Cox proportional hazards model, is described by the *hazard function* which is defined for fixed $Z = z$ as

$$\lambda(t, z) = \lambda_0(t) \exp(\beta' z),$$

where $\lambda_0(t)$ is the baseline hazard function and $\beta \in R^k$ is a vector of regression parameters.

Cox's (1972, 1975) proposed the method of estimation of the regression parameters β which does not require specification of the baseline hazard function $\lambda_0(t)$. The method consists in maximization of the partial likelihood function and it is equivalent to solving the following score function equation

$$\sum_{i=1}^n \left[Z_i - \frac{\sum_{T_j \geq T_i} Z_j \exp(\beta' Z_j)}{\sum_{T_j \geq T_i} \exp(\beta' Z_j)} \right] I_{T_i \leq C_i} = 0,$$

where T_i , C_i , Z_i denote respectively the time, censoring and covariate variables in the sample.

It was shown in Samuels (1978) and in Bednarski (1989) that inference for the Cox model based on the partial likelihood can be very sensitive to even small departures from the model (see also Reid and Crépeau (1985)). A preliminary examination of model's fit to the data is typically done via the likelihood ratio significance test. Test statistics based on the log likelihood ratio are necessarily bound to maximum likelihood estimation – a method which is usually not robust. This is also the case for the Cox model, where the partial likelihood and the partial likelihood estimation is used. This simple method of testing significance fails under data contamination. Mimicking the likelihood ratio methodology to objective functions – they have the potential of producing robust estimates – leads to non tractable limiting distributions under the null hypothesis. A natural way to bypass the difficulties is in application of proper variants of the Wald test statistic, based on the asymptotic standardization of $\sqrt{n}(\hat{\beta} - \beta_0)$, where $\hat{\beta}$ is the estimator of β_0 and β_0 stands either for the true parameter value or the null hypotheses parameter value.

Since significance testing may be viewed as a preliminary fit testing, it is worth mentioning that the goodness of fit testing for the Cox model, based on estimates of β_0 , was first proposed in Lin (1991) and its robustified version was given in Krug (1998). The idea was in using a standardized difference of two different estimates of β_0 , weighted and non-weighted ones, as test statistic. Minder and Bednarski (1996) proposed a heuristic way of studying goodness of fit for the Cox model under robust estimation. The method was based on computation of normalized differences between the Kaplan – Meier and Cox model estimators of survival functions for conveniently stratified data.

In the following section a detailed derivation of the robustified Wald test statistic along with its asymptotic behavior is given. The final section gives extensive Monte Carlo comparison of distributions of the partial likelihood ratio test statistic with two variants of robustified Wald statistics. The study is carried out for samples from the Cox model and for contaminated samples from the model.

2. MODIFIED WALD STATISTIC

Bednarski (1993), using the Fréchet differentiability of statistical functionals, proposed a robust method of estimation of regression coefficients based on a modification of the partial likelihood estimator. The method consists in solving the following equation

$$L(F_n, \beta, A) = \int A(w, y) \left[y - \frac{\int A(w, z) z I_{a \wedge t \geq w} \exp(\beta' z) dF_n(t, a, z)}{\int A(w, z) I_{a \wedge t \geq w} \exp(\beta' z) dF_n(t, a, z)} \right] I_{w \leq c} dF_n(w, c, y) = 0,$$

where F_n denotes the empirical distribution function of a sample $(T_1, C_1, Z_1), \dots, (T_n, C_n, Z_n)$, t is the time failure variable, c is the censoring variable and z is the k -dimensional covariate vector. Conditions defining weight functions A that lead to the differentiable estimator-functional $\beta(F)$ are given in Bednarski (1993). If weights are equal 1 then the method reduces to the partial likelihood estimation.

The differentiability property ensures the following expansion

$$\sqrt{n} (\beta(F_n) - \beta(F)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_{\beta_0, A}(X_i) + o_P(1),$$

where $\psi_{\beta_0, A}(x) = Lin'_{\beta_0, A}(\delta_x - F)H^{-1}(F, \beta_0, A)$, where $H(F, \beta, A)$ is a matrix with i -th row equal to the vector

$$\frac{L(F, \beta, A)}{\beta_i} =$$

$$\int A \left[\frac{\int AI_{t \geq w} z z_i \exp(\beta' z) dF \int AI_{t \geq w} \exp(\beta' z) dF}{\left[\int AI_{t \geq w} \exp(\beta' z) dF \right]^2} \right.$$

$$\left. - \frac{\int AI_{t \geq w} z \exp(\beta' z) dF \int AI_{t \geq w} z_i \exp(\beta' z) dF}{\left[\int AI_{t \geq w} \exp(\beta' z) dF \right]^2} \right] dF,$$

$$Lin_{A, \beta}(G - F) =$$

$$\int A \left[y - \frac{\int AI_{t \geq w} z \exp(\beta' z) dF}{\int AI_{t \geq w} \exp(\beta' z) dF} \right] d(G - F)$$

$$- \int A \left[\frac{\int AI_{t \geq w} z \exp(\beta' z) dF \int AI_{t \geq w} \exp(\beta' z) d(G - F)}{\left[\int AI_{t \geq w} \exp(\beta' z) dF \right]^2} \right.$$

$$\left. - \frac{\int AI_{t \geq w} \exp(\beta' z) dF \int AI_{t \geq w} z \exp(\beta' z) d(G - F)}{\left[\int AI_{t \geq w} \exp(\beta' z) dF \right]^2} \right] dF.$$

The expansion holds for the Cox model distribution with true parameter value β_0 and for its infinitesimal neighborhoods.

By the central limit theorem it follows that statistic

$$\sqrt{n}(\beta(F_n) - \beta_0)' \times \left(\frac{1}{n} \sum_{i=1}^n \psi_{\beta_0,A}(X_i) \psi'_{\beta_0,A}(X_i) \right)^{-1} \times \sqrt{n}(\beta(F_n) - \beta_0)$$

has, under the Cox model, the asymptotic chi-square distribution with k degrees of freedom, where k is the dimension of the vector of regression parameters. The statistic has a noncentral chi-square distribution under small departures from the model. If $\psi_{\beta,A}$ is a sufficiently smooth function of β , then the following statistic

$$\sqrt{n}(\hat{\beta} - \beta_0)' \times \left(\frac{1}{n} \sum_{i=1}^n \psi_{\hat{\beta},A}(X_i) \psi'_{\hat{\beta},A}(X_i) \right)^{-1} \times \sqrt{n}(\hat{\beta} - \beta_0)$$

has the same limiting behavior and it can then be used to significance testing for the Cox model: $H_0 : \beta = \beta_0$ versus $H_1 : \beta \neq \beta_0$.

3. SIMULATIONS

A Monte Carlo experiment was carried out to compare the behavior of the likelihood ratio test statistic (PLE) with two modified robust versions of the Wald statistic. One of them - based on the asymptotic influence function expansion - was described in the previous section (IF). The second one uses the hessian matrix instead of analytically derived covariance matrix (HESS). The "hessian" method is included here since it is a frequently applied variance assessment method in statistical packages.

The asymptotic behavior of the statistics is studied via the performance of their distribution functions under the null and alternative hypotheses, with and without contamination.

In the first step samples from the Cox model were generated. They were of size 100 and 300 for $\lambda_0(t) \equiv 1$, $\lambda_0(t) = t$ and for β equal respectively to $(0, 0)'$, $(1, -1)'$, $(0.1, 0.1)'$, $(0.1, -0.1)'$, $(0.5, 0.5)'$, $(0.5, -0.5)'$. Covariates were generated from the standard normal distribution. In each model case 1000 repetitions were used to approximate the distribution functions of the statistic. In the contaminated samples, independent $2N(0, 1) + 1$ random variables replaced covariates from the original Cox model sample in a fixed

fraction of items. Then, for each sample, the following three test statistics were computed:

- the PLE statistic equal to $-2 \left[\log(L(\beta_0)) - \log(L(\hat{\beta}_{PLE})) \right]$, where $\hat{\beta}_{PLE}$ is the partial likelihood estimator and $L(\beta)$ is the partial likelihood function,
- the HESS statistic given by

$$\sqrt{n} (\hat{\beta} - \beta_0)' \times H (F_n, \hat{\beta}, A) \times \sqrt{n} (\hat{\beta} - \beta_0),$$

where $\hat{\beta}$ is the robust estimator of regression parameters and H denotes the hessian of the logarithm of the weighted partial likelihood function.

- the “influence function” statistic IF given by

$$\sqrt{n} (\hat{\beta} - \beta_0)' \times \left(\frac{1}{n} \sum_{i=1}^n \psi_{\hat{\beta}}(X_i) \psi'_{\hat{\beta}}(X_i) \right)^{-1} \times \sqrt{n} (\hat{\beta} - \beta_0).$$

The following two graphs (Figures 1 and 2) show results of simulations for $\beta_0 = (0, 0)'$ – the true parameter value. We can see that the empirical distribution functions of the three test statistics coincide very well with the theoretical chi-square distribution (gray dashed curve) for the uncontaminated and for the two percent contaminated data.

Figures 3 and 4 give simulation results for $\beta_0 = (1, -1)'$. For uncontaminated data the empirical distribution of PLE shows visibly better fit to the chi-square distribution than HESS and IF. However, the 2% contamination changes dramatically the situation in favor of the robustified statistics. The contamination effect magnifies with the sample size as seen on Figure 5 and Figure 6.

The following figures show the behavior of the test statistics under alternatives when $H_0 : \beta = (0, 0)'$. Figures 7 and 8 correspond to simulation results when the true vector parameter $\beta = (0.3, 0.3)'$. To expose better the distributional effects 5% contamination was applied with contaminating covariates constant and equal 3. As expected the power of the PLE significance test is noticeably superior under the model. However it becomes inferior under the slight contamination.

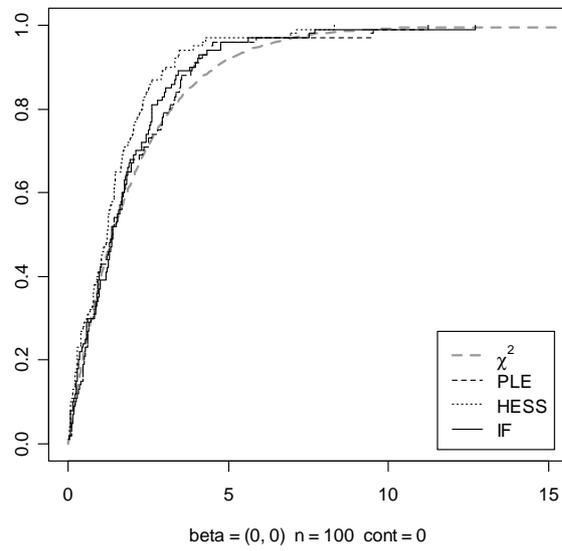


Figure 1.

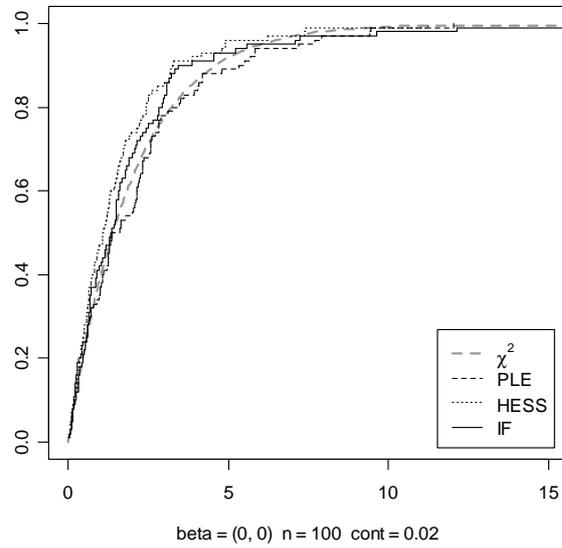


Figure 2.

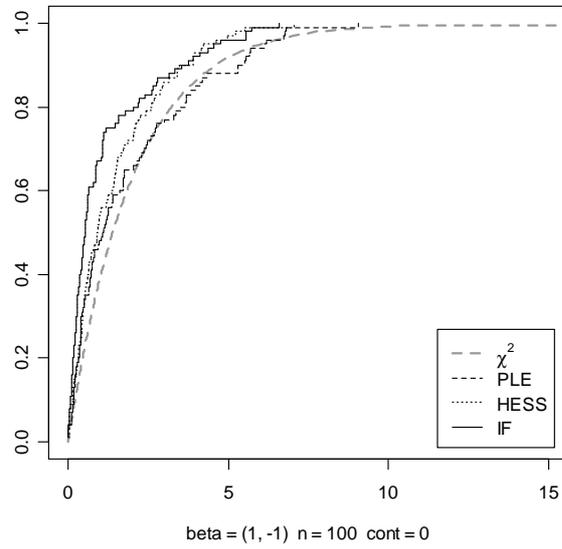


Figure 3.

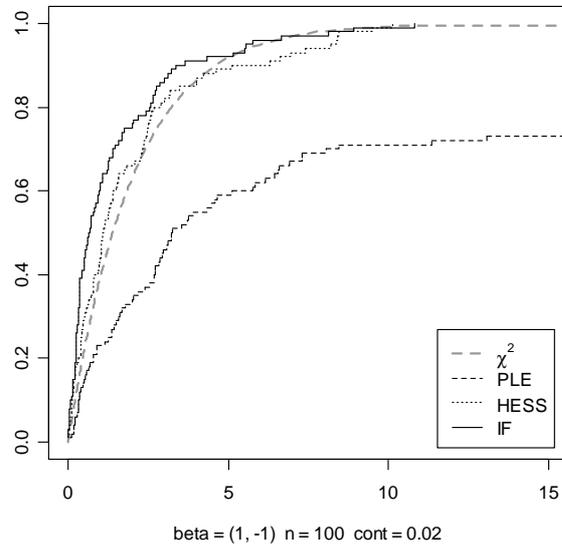


Figure 4.

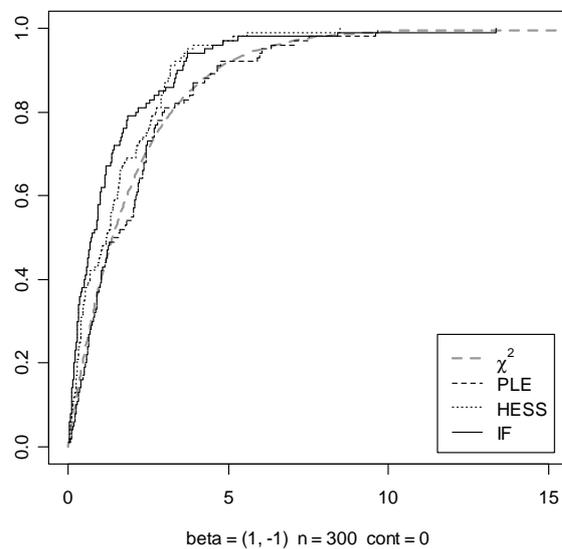


Figure 5.

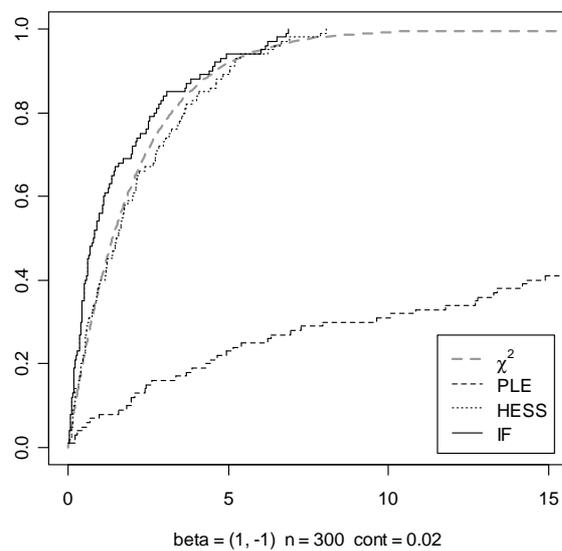


Figure 6.

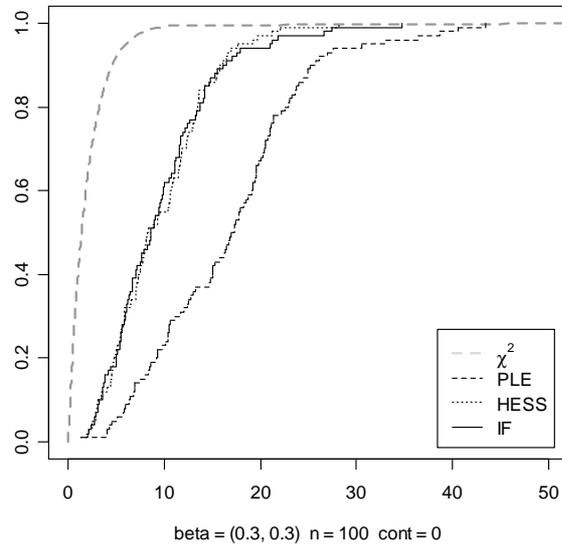


Figure 7.

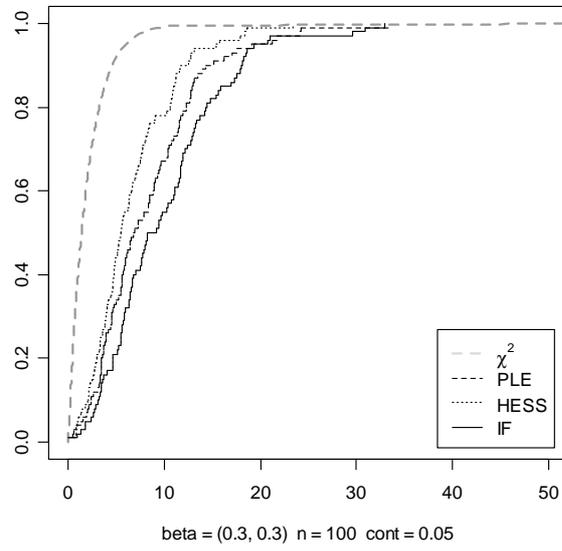


Figure 8.

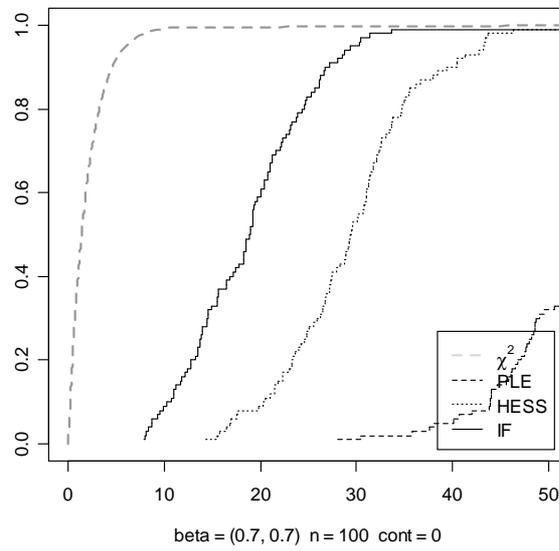


Figure 9.

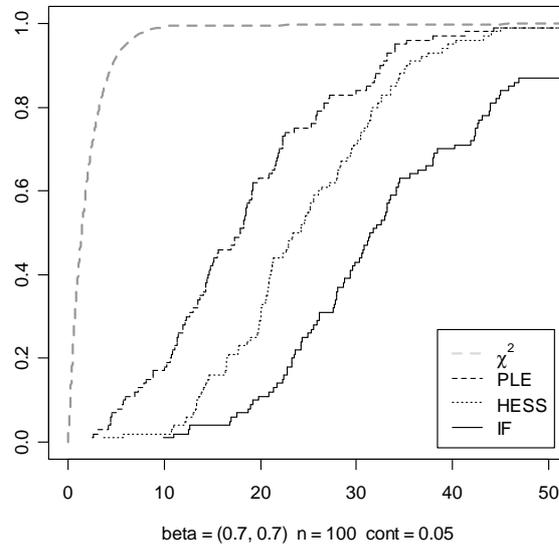


Figure 10.

The effects just described sharpen when the true parameter becomes more distant from zero. It is largely demonstrated in Figures 9 and 10 where β was taken $(0.7, 0.7)'$.

The overall conclusion is that the robust method of testing significance in the Cox regression model, using the robustified Wald test, may be very useful in practice as a supplementary method. Even though it gives tests less efficient than the PLE under the model, it leads to more stable distributions of the test statistic under small departures from the Cox model. In practical terms it means that we have a better control of the significance level and the power of tests in the inference process.

The above discussed test based on the influence function (IF) was implemented into a statistical package for robust inference in the Cox regression model. The package called "coxrobust" is part of the R project.

REFERENCES

- [1] T. Bednarski, *On sensitivity of Cox's estimator*, Statistics and Decisions **7** (1989), 215–228.
- [2] T. Bednarski, *Robust estimation in Cox's regression model*, Scandinavian Journal of Statistics **20** (1993), 213–225.
- [3] D.R. Cox, *Regression models and life tables*, Journal of the Royal Statistical Society. Series B, **34** (1972), 187–220.
- [4] D.R. Cox, *Partial likelihood*, Biometrika **62** (2) (1975), 269–276.
- [5] B. Krug, *Robust Estimation in Selected Additive Models*, PhD thesis (1998), Institute of Mathematics of the Polish Academy of Sciences.
- [6] D.Y. Lin, *Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators*, J. Amer. Statist. Assoc. **86** (1991), 725–729.
- [7] Ch. Minder and T. Bednarski, *A robust method for proportional hazards regression*, Statistics in Medicine **15** (1996), 1033–1047.
- [8] N. Reid and H. Crépeau, *Influence functions for proportional hazards regression*, Biometrika **72** (1985), 1–9.

- [9] S. Samuels, *Robustness for survival estimators*, Unpublished PhD thesis (1978), Dept. of Biostatistics, Univ. of Washington.

Received 26 November 2006

Revised 8 February 2007