Generating Survival Times to Simulate Cox Proportional Hazards Models

Ralf Bender[1], Thomas Augustin[2], Maria Blettner[1]

[1]Dept. of Epidemiology and Medical Statistics, School of Public Health
University of Bielefeld, Germany

[2]Department of Statistics, University of Munich, Germany

Corresponding Author:

Ralf Bender, Ph.D., statistician
Department of Epidemiology and Medical Statistics
School of Public Health, University of Bielefeld
P.O. Box 10 01 31, D–33501 Bielefeld, Germany

Phone: +49 521 106-3803
Fax: +49 521 106-6465
Email: Ralf.Bender@uni-bielefeld.de

Contract/grant sponsor: German Research Foundation (DFG)

Contract/grant numbers: BE 2485/1-1, SFB 386

Short Title: Generating Survival Times

Number of words: 2795 (excluding summary, references and tables)

SUMMARY

This paper discusses techniques to generate survival times for simulation studies regarding Cox proportional hazards models. In linear regression models, the response variable is directly connected with the considered covariates, the regression coefficients and the simulated random errors. Thus, the response variable can be generated from the regression function, once the regression coefficients and the error distribution are specified. However, in the Cox model, which is formulated via the hazard function, the effect of the covariates have to be translated from the hazards to the survival times, because the usual software packages for estimation of Cox models require the individual survival time data. A general formula describing the relation between the hazard and the corresponding survival time of the Cox model is derived. It is shown how the exponential, the Weibull and the Gompertz distribution can be used to generate appropriate survival times for simulation studies. Additionally, the general relation between hazard and survival time can be used to develop own distributions for special situations and to handle flexibly parameterized proportional hazards models. The use of other distributions than the exponential distribution only is indispensable to investigate the characteristics of the Cox proportional hazards model, especially in non-standard situations, where the partial likelihood depends on the baseline hazard.

KEY WORDS: Cox proportional hazards model; exponential distribution; Gompertz distribution; simulation; survival times; Weibull distribution

1. INTRODUCTION

Simulation studies represent an important statistical tool to investigate the performance, properties and adequacy of statistical models, test statistics and estimation techniques considering pre-specified conditions. One of the most important statistical models in medical research is the Cox proportional hazards model [1]. The Cox model and the corresponding partial likelihood [2] are intensively investigated by means of simulation studies to get information about bias and efficiency of the estimated regression coefficients for a variety of situations, in particular when fundamental model assumptions are violated. For example, Hu, Tsiatis and Davidian [3] compared several approaches to estimate the parameters of a Cox model when covariates are measured with error. They performed a number of simulations with exponentially distributed survival times.

Because of censoring, it is convenient to model survival times through the hazard function. The Cox proportional hazards model is given by

$$\mathrm{h}(t\,|\,x) \;=\; \mathrm{h}_0(t) \times \exp(\beta' x) \tag{1}$$

where t is the time, x the vector of covariates, β the vector of regression coefficients and $\mathrm{h}_0(t)$ is the so-called baseline hazard function, i.e. the hazard function under x=0. Because the model is formulated through the hazard function, the simulation of appropriate survival times for the Cox model is not straightforward. One important issue in simulation studies regarding regression models is the knowledge of the true regression coefficients. This is no problem in a linear regression model, where the simulated variables are directly connected with the pre-specified regression coefficients. However, in the Cox model, the effect of the covariates have to be translated from the hazards to the survival times, because the usual software packages for Cox models require the individual survival time data, not the hazard function. The translation of the regression effects from hazard to survival time is easy if the baseline hazard function is constant, i.e. the survival times are exponentially distributed. This may be the reason why most simulation studies regarding the Cox model consider only the exponential distribution. Another frequently used distribution for survival times is the Weibull distribution [4]. In simulation studies, a common practice is to consider only binary covariates such as group 1 and group 2. For example, Schemper [5] simulated Weibull distributed survival times for the situation of two binary covariates in order to compare strategies for analysis with the Cox model in the presence of non-proportional hazards. In the case of discrete covariates, the Weibull distributions can be specified with different sets of parameters for each group. The Weibull parameters can be chosen such that the hazards are proportional and the true hazard ratio (HR) for the comparison of the two groups can be calculated from the Weibull parameters. Then, the true regression coefficient for the Cox model can be obtained from log(HR).

Considering only the exponential and/or the Weibull distribution may be sufficient for some applications. However, for a realistic description of various survival time data, other distributions are required. One important field in medicine is the modeling of human mortality for which frequently the Gompertz distribution is used. Other commonly used distributions in survival time analysis are the gamma, the lognormal and the log-logistic distribution [4]. The latter distributions, however, do not have the proportional hazards property. Among the known parametric distributions, only the exponential, the Weibull and the Gompertz model share the assumption of proportional hazards with the Cox regression model [4]. In this paper, it is shown how survival times can be generated to simulate Cox models with known regression coefficients considering especially the exponential, the Weibull and the Gompertz distribution. The general relation between the hazard and the corresponding survival time of the Cox model is developed, from which the required relation for these three distributions can be derived as special cases. Additionally, the general relation between hazard and survival time can be used to develop own distributions for special situations, and to study proportional hazards models with flexibly parameterized baseline hazard functions.

2. SIMULATING SURVIVAL TIMES

2.1. General considerations

The survival function of the Cox proportional hazards model (1) is given by

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

where

$$H_o(t) = \int_0^t h_o(u)\, du$$

is the cumulative baseline hazard function [6]. Thus, the distribution function of the Cox model is

$$F(t|x) = 1 - \exp[-H_o(t) \times \exp(\beta' x)] \qquad (2)$$

Let Y be a random variable with distribution function F, then $U = F(Y)$ follows a uniform distribution on the interval from 0 to 1 [7], abbreviated as $U \sim \mathrm{Uni}[0,1]$. Moreover, if $U \sim \mathrm{Uni}[0,1]$, then $(1-U) \sim \mathrm{Uni}[0,1]$, too [7]. Thus, let T be the survival time of the Cox model (1), then it follows from (2) that

$$U = \exp[-H_o(T) \times \exp(\beta' x)] \sim \mathrm{Uni}[0,1]$$

If $h_o(t) > 0$ for all t, then H_o can be inverted and the survival time T of the Cox model (1) can be expressed as

$$T = H_o^{-1}[-\log(U) \times \exp(-\beta' x)] \qquad (3)$$

where U is a random variable with $U \sim \mathrm{Uni}[0,1]$. Equation (3) is suitable for the generation of survival times, because random numbers following a $\mathrm{Uni}[0,1]$ distribution are frequently available in statistical program packages. For example in SAS, uniformly distributed random numbers can be generated by means of the function RANUNI [8]. As $-\log(U)$ is exponentially distributed with parameter 1 if $U \sim \mathrm{Uni}[0,1]$, we can also use exponentially distributed random numbers. In SAS, the function RANEXP can be used to generate random numbers following an exponential distribution [8]. After the generation of appropriate random numbers, these can be transformed into survival times for Cox models by applying formula (3). It is just required to insert the inverse of an appropriate cumulative baseline hazard function into equation (3). Thus, it is possible to generate survival times for simulating any Cox model with positive baseline hazard function by transformations of uniformly or exponentially distributed random numbers.

2.2. Exponential distribution

The exponential distribution with scale parameter $\lambda > 0$ has a constant hazard function for $t \geq 0$ [6]. The inverse of the cumulative hazard function (see Table I) is given by

$$\mathrm{H}_{\mathrm{o}}^{-1}(t) = \lambda^{-1} t \tag{4}$$

By inserting (4) into equation (3), we get the following expression for the survival time of a Cox model with constant baseline hazard

$$T = \lambda^{-1} \left[-\log(U) \times \exp(-\beta' x)\right] = -\frac{\log(U)}{\lambda \times \exp(\beta' x)} \tag{5}$$

The corresponding hazard function of the Cox model is given by

$$\mathrm{h}(t \,|\, x) = \lambda \times \exp(\beta' x) \tag{6}$$

Thus, the Cox model (1) with constant baseline hazard results in exponentially distributed survival times with scale parameters $\lambda(x) = \lambda \times \exp(\beta' x)$, which are dependent on the regression coefficients and the covariates considered.

2.3. *Weibull distribution*

In practice, the assumption of a constant hazard function is only rarely tenable. A more general form of the hazard function is given by the Weibull distribution, which is characterized by two positive parameters [9]. In the formulation shown in Table I, the parameter λ is known as the *scale* parameter, while ν is the *shape* parameter. In the particular case where $\nu = 1$ the hazard function reduces to that of the exponential distribution. For $\nu > 1$, the hazard function increases and for $0 < \nu < 1$, it decreases monotonically. The inverse of the cumulative hazard function is given by

$$\mathrm{H}_o^{-1}(t) = (\lambda^{-1} t)^{1/\nu} \tag{7}$$

By inserting (7) into equation (3), it follows that the survival time of a Cox model with the baseline hazard of a Weibull distribution can be expressed as

$$T = \lambda^{-1} \left[-\log(U) \times \exp(-\beta' x)\right]^{1/\nu} = \left(-\frac{\log(U)}{\lambda \times \exp(\beta' x)}\right)^{\frac{1}{\nu}} \tag{8}$$

The corresponding hazard function is given by

$$\mathrm{h}(t \mid x) = \lambda\, \nu\, t^{\nu - 1} \exp(\beta' x) = \lambda \exp(\beta' x)\, \nu\, t^{\nu - 1} \tag{9}$$

This means that the corresponding survival times are Weibull distributed with varying scale parameter $\lambda(x) = \lambda \times \exp(\beta' x)$ and fixed shape parameter ν.

2.4. *Gompertz distribution*

The Gompertz distribution represents another extension of the exponential distribution. Like the Weibull, the Gompertz distribution is characterized by two parameters [10]. In the formulation shown in Table I, when $\alpha < 0$ (> 0), the hazard function decreases (increases) from $\exp(\alpha)$, and when $\alpha = 0$, it reduces to the constant hazard function of an exponential distribution. The inverse of the cumulative hazard function is given by

$$H_0^{-1}(t) \; = \; \frac{1}{\alpha} \; \log\left(\frac{\alpha}{\lambda} t + 1\right) \tag{10}$$

By inserting (10) into equation (3), it follows that the survival time of a Cox model with the baseline hazard of a Gompertz distribution can be expressed as

$$T \; = \; \frac{1}{\alpha} \log\left[-\frac{\alpha}{\lambda}\left(\log(U)\exp(-\beta' x)\right)+1\right] \; = \; \frac{1}{\alpha} \log\left[1 - \frac{\alpha \times \log(U)}{\lambda \times \exp(\beta' x)}\right] \tag{11}$$

The corresponding hazard function is given by

$$h(t) = \lambda \times \exp(\alpha t) \times \exp(\beta' x) = \lambda \times \exp(\beta' x) \times \exp(\alpha t) \tag{12}$$

This means that the corresponding survival times are also Gompertz distributed with varying parameter $\lambda(x) = \lambda \times \exp(\beta' x)$ and fixed parameter α.

2.5. Proportional hazards models with other distributions

While up to now techniques to run simulations based on standard parametric distributions have been reported, the result in (3) is also of great value in the whole generality described there. Firstly, it allows to design comprehensive simulation studies for all the variants of the Cox model where the baseline hazard rate is modeled in a flexible parametric way. Then, instead of (1), one considers

$$h(t \mid x) \; = \; g(\eta, t) \; \times \; \exp(\beta' x) \tag{13}$$

where $g(\cdot)$ is a function known up to a multidimensional parameter η. Model (13) contains the approaches following Kalbfleisch and Prentice [11], where the hazard is assumed to be constant within intervals fixed in advance, the polynomial proportional hazard models proposed by Ciampi and Etezadi-Amoli [12] and Taulbee [13], as well as the spline method introduced by Whittemore and Keller [14], all developed in a biometrical context. In econometrics, special types of model (13) are particularly attractive for modelling dynamics in unemployment spells [15,16]. On the handling of these models in the presence of measurement error, see Augustin [17].

Secondly, result (3) enables the investigation of the behavior and stability of estimators in the Cox model under certain additional assumptions on the baseline hazard rate. For instance, the function

$$\mathrm{h_o}(t) \;=\; \left| \, a \times \sin(t) \, \right|$$

can be used to model regularly recurring periods of high baseline hazard of magnitude a > 0. Integration leads to

$$\mathrm{H_o}(t) \;=\; a \times \left(2 \left[\!\left[\frac{t}{\pi} \right]\!\right] + 1 + (-1)^{\left[\!\left[\frac{t}{\pi} \right]\!\right] + 1} \times \cos(t) \right)$$

where $[\![y]\!]$ is the truncation function, which returns the largest integer less or equal to *y*. Inverting $\mathrm{H_o}(t)$ finally yields

$$\mathrm{H_o^{-1}}(t) \;=\; \left[\!\left[\frac{t}{4a} \right]\!\right] \times \pi + \left[\!\left[0.5 + \frac{t}{4a} \right]\!\right] \times \pi + \arccos\left[-\left(t - 2 \times \left[\!\left[\frac{t}{2a} \right]\!\right] - 1 \right) \right]$$

which provides the basis for designing and performing simulations in this model.

For practical applications notice also that model (13) can be used for (almost) arbitrary functions. Neither $\mathrm{H_o}(t)$ nor its inverse function $\mathrm{H_o^{-1}}(t)$ are indispensably needed in analytical form. So, in principle, it is sufficient to determine both numerically, i.e., by numerical integration and inversion, respectively.

3. EXAMPLE

3.1. Simulations for the German Uranium Miners Cohort Study

The German Uranium Miners Cohort Study is one of the largest cohort studies on uranium miners with the purpose of evaluating the risks of cancer and mortality associated with low and high levels of radon exposure [18]. To investigate the effect of measurement error in the exposure values on hazard ratios estimated by means of Cox proportional hazards models, a simulation study has been performed [19]. As the generated survival times in the simulation study should have a similar distribution like the observed survival times in the cohort study, the Gompertz distribution was applied. By using the exponential distribution it was impossible to generate realistic survival times. Either the number of deaths or the attained age were too high in the simulated data. By using the Gompertz distribution, realistic survival times reflecting the mortality of German men can be generated.

One task in generating survival times with specific features is to find appropriate parameters for the model considered. Here, an obvious approach is to relate the expected value of the Cox-Gompertz model (11) to the tables for life expectancy of German men. Unfortunately, the expected value of a Gompertz distributed random variable T is given by a formula containing an integral which has to be evaluated numerically (Table I) [20]. Thus, appropriate parameter values of the Cox-Gompertz model can not be calculated directly. However, as the Gompertz distribution represents a left truncated extreme value distribution at time point t=0 [20], the extreme value distribution can be used as approximation to the Gompertz distribution. The extreme value distribution is defined for $-\infty < t < \infty$, but the hazard function for $t \geq 0$ is identical to that of the Gompertz distribution. The density and survival function of the extreme value distribution with parameters λ and α are given by [21]

$$f_o(t) = \lambda \times \exp(\alpha t) \times \exp\left(-\tfrac{\lambda}{\alpha} \exp(\alpha t)\right)$$

$$S_o(t) = \exp\left(-\tfrac{\lambda}{\alpha} \exp(\alpha t)\right)$$

The mean and variance of an extreme value distributed random variable *T* are given by [21]

$$E(T) = \mu_o = -\frac{1}{\alpha}\left(\log\left(\tfrac{\lambda}{\alpha}\right) + \gamma\right) \tag{14}$$

$$\mathrm{Var}(T) = \sigma_o^2 = \frac{\pi^2}{6\alpha^2} \tag{15}$$

where $\gamma \approx 0.5772$ is Euler's constant and $\pi \approx 3.14159$. Solving (14) and (15) for λ and α leads to

$$\alpha = \frac{\pi}{\sqrt{6}\,\sigma_o} \quad , \qquad \lambda = \alpha \exp(-\gamma - \alpha\mu_o) \tag{16}$$

which can be used to calculate approximately the parameters of the Gompertz distribution in dependence on the mean and variance of the considered survival time. For the mean life expectancy of μ_o=66.86 years and a standard deviation of σ_o=6 years, we get the values $\lambda = 7\times10^{-8}$ and $\alpha = 0.2138$. By using the regression coefficient $\beta_{age} = 0.15$ for age in the Cox-Gompertz model (11), we can generate survival times leading to realistic attained age values and numbers of deaths similar to those observed in the German Uranium Miners Cohort Study [19].

3.2. Comparison of the exponential and the Gompertz distribution

To assess the importance of using a realistic survival distribution in simulation studies, we compare the results of simulations of the Cox-Gompertz model (11) with those of the Cox model with exponentially distributed survival times (5). Frequently, it is argued that the choice of the distribution for the generated survival times is rather unimportant in simulations studies regarding the Cox model. Here, we present one example, where the results are dependent on the chosen survival distribution. Again, we consider simulations concerning the German Uranium Miners Cohort Study [18].

The goal of the simulation study was to investigate the effect of measurement error in the radon exposure values on the estimated hazard ratios [19]. In radon epidemiology, both classical measurement error and Berkson type errors may play a role [22]. The data situation of the German Uranium Miners Cohort Study is used with the following characteristics: sample size n=58721, total study time 1946-1998, mean (SD) age at study entry 24.3 (8.38) years, mean (SD) cumulative radon exposure 266.84 (507.82) working level months (WLM) [18]. Cox proportional hazard models considering the covariates age at baseline and radon exposure were simulated by using the Gompertz and the exponential distribution for the baseline hazard, respectively. Only a small part of the simulations are considered here for demonstrating purposes. We present the results of additive Berkson type measurement error and additive classical measurement error for radon exposure for one parameter situation.

Normally distributed measurement errors with μ_e=0 and σ_e=359.1 were generated, so that the measurement error variance amounts to 50% of the variance of the observed radon exposure. For each situation 1000 simulations were performed. In Table II the relative bias of the estimated Cox regression coefficients is shown. In all cases, measurement error leads to an attenuation of the true effect for both covariates, shown by the negative relative bias values. For the Berkson type error, the bias values are quite similar. Thus, in this situation, the conclusions for both distributions would be the same, although the exponentially distributed survival times did not fit the observed survival times of the German Uranium Miners Cohort Study. However, in the case of classical measurement error, the relative bias for the estimated exposure effect is much higher in the simulated Cox model with constant baseline hazard (–41.49%) in comparison to the

Cox-Gompertz model (–25.73%). Hence, the use of the exponential distribution would lead to incorrect conclusions about the amount of attenuation due to measurement error in the German Uranium Miners Cohort Study.

4. CONCLUSION

The high capacity of performing calculations by means of modern computers allows the evaluation of statistical methods via simulation studies. One of the most important statistical models in medical research is the Cox proportional hazards model, which is intensively investigated by means of simulation studies. While the exponential distribution is widely used for the generation of survival times in simulation studies, other distributions seem to be underutilized. One reason may be that the generation of survival times in dependence on pre-specified Cox regression coefficients is not obvious. We have developed the general relation between the hazard and the survival time of the Cox model, which can be used to generate survival times following the exponential, Weibull and Gompertz distribution as well as own empirical distributions for special situations.

As the partial likelihood in the classical Cox model does not depend on the baseline hazard, not much attention is paid to the choice of the distribution of the generated survival times in simulations studies regarding the Cox model. However, there are a lot of practical situations, where the use of more flexible distributions than the exponential distribution is required in simulation studies investigating the characteristics of the Cox proportional hazards model. When fundamental assumptions of the Cox model are violated so that the partial likelihood depends on the baseline hazard, e.g. in the presence of measurement error, the results of the simulation study may substantially depend on the distribution of the generated survival times. In these cases, the distribution of the generated survival times should reflect the considered data situation to get appropriate simulation results. The general relation between the hazard and the survival time of the Cox model presented in this paper can be used for this purpose.

ACKNOWLEDGEMENTS

Supported in part by German Research Foundation (DFG) grants BE 2485/1-1 and SFB 386. We thank Drs Bernd Grosche, Michaela Kreuzer and Annemarie Brachner for providing the data of the German Uranium Miners Cohort Study. We thank also Dr Jürgen Wellmann and Dirk Taeger for useful discussions on simulation techniques.

References

1. Cox DR. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society - B* 1972; **34**: 187-220.

2. Cox DR. Partial likelihood. *Biometrika* 1975; **62**: 269-276.

3. Hu P, Tsiatis AA, Davidian M. Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics* 1998; **54**: 1407-1419.

4. Lee ET, Go OT. Survival analysis in public health research. *Annual Review of Public Health* 1997; **18**: 105-134.

5. Schemper M. Cox analysis of survival data with non-proportional hazard functions. *Statistician* 1992; **41**: 455-465.

6. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Wiley: New York, 1980.

7. Mood AM, Graybill FA, Boes DC. *Introduction to the Theory of Statistics*. McGraw-Hill: New York, 1974.

8. SAS. *SAS Language Guide for Personal Computers, Version 6 Edition*. SAS Institute Inc.: Cary, NC, 1985.

9. Collett D. *Modelling Survival Data in Medical Research*. Chapman and Hall: London, 1994.

10. Lee ET. *Statistical Methods for Survival Data Analysis*. Wiley: New York, 1992.

11. Kalbfleisch JD, Prentice RL. Marginal likelihood based on Cox's regression and life model. *Biometrika* 1973; **60**: 267-278.

12. Ciampi A, Etezadi-Amoli J. A general model for testing the proportional hazards and the accelerated failure time hypothesis in the analysis of censored survival data with covariates. *Communications in Statistics - A: Theory and Methods* 1985; **14**: 651-557.

13. Taulbee JD. A general model for the hazard rate with covariables. *Biometrics* 1979; **35**: 439-450.

14. Whittemore AS, Keller JB. Survival estimation using splines. *Biometrics* 1986; **42**: 495-506.

15. Flinn CJ, Heckmann JJ. Models for the analysis of labor force dynamics. In *Advances in Econometrics*, Basman RL, Rhodes GF (eds). JAI Press: Greenwich, CT, 1982; 35-95.

16. Gritz M. The impact of training on the frequency and duration of employment. *Journal of Econometrics* 1993; **57**: 21-51.

17. Augustin T. An exact corrected log-likelihood function for Cox's proportional hazards model under measurement error and some extensions. *Scandinavian Journal of Statistics* 2003; **30** (in press).

18. Kreuzer M, Brachner A, Lehmann F, Martignoni K, Wichmann HE, Grosche B. Characteristics of the German Uranium Miners Cohort Study. *Health Physics* 2002; **83**: 26-34.

19. Bender R, Blettner M. Diskussion der Messfehlerproblematik durch Verwendung einer Job-Exposure-Matrix (JEM). In *Stand der Forschung zu den "Deutschen Uranbergarbeiterstudien" 1. Fachgespräch am 7./8. Mai 2001 in St. Augustin*, Geschäftsstelle der Strahlenschutzkommission beim Bundesamt für Strahlenschutz (ed). Urban und Fischer: München, 2002; 97-105.

20. Elandt-Johnson RC, Johnson NL. *Survival Models and Data Analysis*. Wiley: New York, 1980.

21. Lawless JE. *Statistical Models and Methods for Lifetime Data*. Wiley: New York, 1982.

22. Heid IM, Küchenhoff H, Wellmann J, Gerken M, Kreienbrock L, Wichmann HE. On the potential of measurement error to induce differential bias on odds ratio estimates: An example from radon epidemiology. *Statistics in Medicine* 2002; **21**: 3261-3278.

TABLES

Table I. Characteristics of the exponential, the Weibull and the Gompertz distribution

	Distribution		
Characteristic	Exponential	Weibull	Gompertz
Parameter	scale parameter $\lambda > 0$	scale parameter $\lambda > 0$ shape parameter $\nu > 0$	scale parameter $\lambda > 0$ shape parameter $\alpha \in (-\infty, \infty)$
Range	$[0,\infty)$	$[0,\infty)$	$[0,\infty)$
Hazard function	$h_o(t) = \lambda$	$h_o(t) = \lambda \nu t^{\nu-1}$	$h_o(t) = \lambda \exp(\alpha t)$
Cumulative hazard function	$H_o(t) = \lambda t$	$H_o(t) = \lambda t^{\nu}$	$H_o(t) = \frac{\lambda}{\alpha}\left(\exp(\alpha t) - 1\right)$
Density function	$f_o(t) = \lambda \exp(-\lambda t)$	$f_o(t) = \lambda \nu t^{\nu-1} \exp(-\lambda t^{\nu})$	$f_o(t) =$ $\lambda \exp(\alpha t) \exp\left(\frac{\lambda}{\alpha}\left(1 - \exp(\alpha t)\right)\right)$
Survival function	$S_o(t) = \exp(-\lambda t)$	$S_o(t) = \exp(-\lambda t^{\nu})$	$S_o(t) = \exp\left(\frac{\lambda}{\alpha}\left(1 - \exp(\alpha t)\right)\right)$
Mean	$E(T) = \frac{1}{\lambda}$	$E(T) = \frac{1}{\sqrt[\nu]{\lambda}}\,\Gamma\left(\frac{1}{\nu} + 1\right)$ Γ denotes the gamma function	$E(T) = \frac{1}{\lambda} G\left(\frac{\lambda}{\alpha}\right)$, where $G(x) = \int_x^{\infty} \frac{1}{y} \exp(-y)\, dy$
Variance	$\mathrm{Var}(T) = \frac{1}{\lambda^2}$	$\mathrm{Var}(T) =$ $\frac{1}{\sqrt[\nu]{\lambda^2}}\left[\Gamma\left(\frac{2}{\nu} + 1\right) - \Gamma^2\left(\frac{1}{\nu} + 1\right)\right]$	

Table II. Relative bias (in %) of estimated Cox regression coefficients due to measurement error in exposure (Berkson type error and classical measurement error) by using the Gompertz and the exponential distribution

Covariate	Berkson measurement error		Classical measurement error	
	Gompertz	exponential	Gompertz	exponential
radon exposure	−6.74	−6.80	−25.73	−41.49
age	−5.78	−6.54	−5.31	−2.98