6

Extension of the Cox Proportional Hazards Model for Time-Dependent Variables

D.G. Kleinbaum and M.Klein, *Survival Analysis: A Self-Learning Text, Third Edition,*
Statistics for Biology and Health, DOI 10.1007/978-1-4419-6646-9_6,

Introduction

We begin by defining a time-dependent variable and providing some examples of such a variable. We also state the general formula for a Cox model that is extended to allow time dependent variables, followed by a discussion of the characteristics of this model, including a description of the hazard ratio.

In the remainder of the presentation, we give examples of models with time-dependent variables, including models that allow for checking the PH assumption for time-independent variables. In particular, we describe a method that uses "heaviside functions" to evaluate the PH assumption for time-independent variables. We also describe two computer applications of the extended Cox model, one concerning a study on the treatment of heroin addiction and the other concerning the Stanford heart transplant study.

Abbreviated Outline

The outline below gives the user a preview of the material to be covered by the presentation. A detailed outline for review purposes follows the presentation.

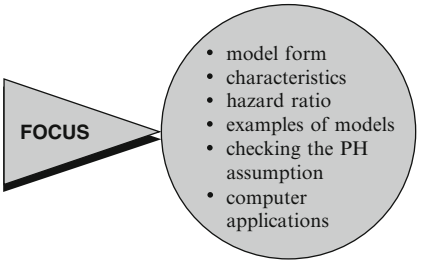
Objectives

Upon completing the chapter, the learner should be able to:

1. State or recognize the general form of the Cox model extended for time-dependent variables.
2. State the specific form of an extended Cox model appropriate for the analysis, given a survival analysis scenario involving one or more time-dependent variables.
3. State the formula for a designated hazard ratio of interest, given a scenario describing a survival analysis using an extended Cox model.
4. State the formula for an extended Cox model that provides a method for checking the PH assumption for one more of the time-independent variables in the model, given a scenario describing a survival analysis involving time-independent variables.
5. State the formula for an extended Cox model that uses one or more heaviside functions to check the PH assumption for one more of the time-independent variables in the model, given a scenario describing a survival analysis involving time-independent variables.
6. State the formula for the hazard ratio during different time interval categories specified by heaviside function(s) that are contained in an extended cox model.
7. Carry out an appropriate analysis of the data to evaluate the effect of one or more of the explanatory variables in the model(s) being used, given computer results for a survival analysis involving time-dependent variables. Such an analysis will involve:
   - computing and interpreting any hazard ratio(s) of interest;
   - carrying out and interpreting appropriate test(s) of hypotheses for effects of interest;
   - obtaining confidence intervals for hazard ratios of interest;
   - evaluating interaction and confounding involving one or more covariates.

Presentation

I. Preview



This presentation describes how the Cox proportional hazards (PH) model can be extended to allow time-dependent variables as predictors. Here, we focus on the model form, characteristics of this model, the formula for and interpretation of the hazard ratio, and examples of the extended Cox model. We also show how the extended Cox model can be used to check the PH assumption for time-independent variables, and we provide computer applications to illustrate different types of time-dependent variables. Finally, we describe the extended cox likelihood and how it contrasts with the Cox PH likelihood function.

II. Review of the Cox PH Model

$$h(t, \mathbf{X}) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$$

$\mathbf{X} = (X_1, X_2, \ldots, X_p)$
Explanatory/predictor variables

The general form of the Cox PH model is shown here. This model gives an expression for the hazard at time t for an individual with a given specification of a set of explanatory variables denoted by the bold $\mathbf{X}$. That is, the bold $\mathbf{X}$ represents a collection (sometimes called a "vector") of predictor variables that is being modeled to predict an individual's hazard.

$$h_0(t) \times \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$$

Baseline hazard	Exponential
Involves t but not X's	Involves X's but not t (X's are time-independent)

The Cox model formula says that the hazard at time t is the product of two quantities. The first of these, $h_0(t)$, is called the **baseline hazard** function. The second quantity is the exponential expression e to the linear sum of $\beta_i X_i$, where the sum is over the p explanatory X variables.

An important feature of this formula, which concerns the proportional hazards (PH) assumption, is that the baseline hazard is a function of t but does not involve the X's, whereas the exponential expression involves the X's but does not involve t. The X's here are called **time-independent** X's.

X's involving t: time-dependent

Requires extended Cox model (no PH)

It is possible, nevertheless, to consider X's that do involve t. Such X's are called **time-dependent** variables. If time-dependent variables are considered, the Cox model form may still be used, but such a model no longer satisfies the PH assumption and is called the **extended Cox model**. We will discuss time-dependent variables and the corresponding extended Cox model beginning in the next section.

Hazard ratio formula:

$$\widehat{HR} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

where $\mathbf{X}^* = \left(X_1^*, X_2^*, \ldots, X_p^*\right)$ and $\mathbf{X} = (X_1, X_2, \ldots, X_p)$ denote the two sets of X's.

From the Cox PH model, we can obtain a general formula, shown here, for estimating a hazard ratio that compares two specifications of the X's, defined as $\mathbf{X}^*$ and $\mathbf{X}$.

PH assumption:

$$\frac{\hat{h}(t, \mathbf{X}^*)}{\hat{h}(t, \mathbf{X})} = \hat{\theta} \text{ (a constant over } t)$$

i.e., $\hat{h}(t, \mathbf{X}^*) = \hat{\theta}\hat{h}(t, \mathbf{X})$

The (PH) assumption underlying the Cox PH model is that the hazard ratio comparing any two specifications of $\mathbf{X}$ predictors is constant over time. Equivalently, this means that the hazard for one individual is proportional to the hazard for any other individual, where the proportionality constant is independent of time.

Hazards cross $\Rightarrow$ PH not met

Hazards don't cross $\not\Rightarrow$ PH met

An example of when the PH assumption is not met is given by any study situation in which the hazards for two or more groups cross when graphed against time. However, even if the hazard functions do not cross, it is possible that the PH assumption is not met.

Three approaches:

- graphical
- time-dependent variables
- goodness-of-fit test

As described in more detail in Chapter 4, there are three general approaches for assessing the PH assumption. These are

- a graphical approach;
- the use of time-dependent variables in an extended Cox model; and
- the use of a goodness-of-fit test.

Time-dependent covariates used to assess PH for time-indep. X's

$\Downarrow$

Extend Cox model: add product term(s) involving function of t

When time-dependent variables are used to assess the PH assumption for a time-independent variable, the Cox model is extended to contain **product** (i.e., interaction) **terms** involving the time-independent variable being assessed and some function of time.

EXAMPLE

$h(t,\mathbf{X}) = h_0(t)\exp[\beta_1 \text{sex} + \beta_2(\text{sex} \times t)]$
H_0: $\beta_2 = 0 \Rightarrow$ PH assumption satisfied

For example, if the PH assumption is being assessed for sex, a Cox model might be extended to include the variable sex $\times$ t in addition to sex. If the coefficient of the product term turns out to be non-significant, we can conclude that the PH assumption is satisfied for sex provided that the variable sex $\times$ t is an appropriate choice of time-dependent variable.

Options when PH assumption not satisfied:

- Use a stratified Cox (SC) model.
- Use time-dependent variables.

There are two options to consider if the PH assumption is not satisfied for one or more of the predictors in the model. In Chapter 5, we described the option of using a stratified Cox (SC) model, which stratifies on the predictor(s) not satisfying the PH assumption, while keeping in the model those predictors that satisfy the PH assumption. In this chapter, we describe the other option, which involves using time-dependent variables.

Time-dependent variables may be:

- inherently time-dependent
- defined to analyze a time-independent predictor not satisfying the PH assumption.

Note that a given study may consider predictors that are inherently defined as time-dependent, as we will illustrate in the next section. Thus, in addition to considering time-dependent variables as an option for analyzing a time-independent variable not satisfying the PH assumption, we also discuss predictors which are inherently defined as time-dependent.

III. Definition and Examples of Time-Dependent Variables

Definition:

Time-dependent	Time-independent
Value of variable differs over time	Value of variable is constant over time
Example:	
Race $\times$ t	Race

A time-dependent variable is defined as any variable whose value for a given subject may differ over time (t). In contrast, a time-independent variable is a variable whose value for a given subject remains constant over time.

As a simple example, the variable RACE is a time-independent variable, whereas the variable RACE $\times$ time is a time-dependent variable.

EXAMPLE OF DEFINED VARIABLES

Defined variable: RACE $\times$ t

Time-independent

Race $= 1 \Rightarrow$ Race $\times\, t = t$

Race $= 0 \Rightarrow$ Race $\times\, t = 0$ (at any t)

$E \times \underbrace{(\log t - 3)}$

Function of t

[E denotes a (0,1) exposure variable].

$E \times g(t)$ where $g(t) = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t < t_0 \end{cases}$

Heaviside function

$t \geq t_0 : E \times \overbrace{g(t)}^{\boxed{1}} = E$

$t < t_0 : E \times \overbrace{g(t)}^{\boxed{0}} = 0$

Heaviside functions used when PH assumptions not met.

The variable RACE $\times$ time is an example of what is called a "defined" time-dependent variable. Most defined variables are of the form of the product of a time-independent variable (e.g., RACE) multiplied by time or some function of time. Note that after RACE is determined for a given subject, all the values of the RACE $\times$ time variable are completely defined over a specified time interval of study.

A second example of a defined variable is given by $E \times (\log t - 3)$, where E denotes, say, a (0,1) exposure status variable determined at one's entry into the study. Notice that here we have used a function of time, $\log t - 3$, rather than time t alone.

Yet another example of a defined variable, which also involves a function of time, is given by $E \times g(t)$, where $g(t)$ is defined to take on the value 1 if t is greater than or equal to some specified value of t, called t_0, and takes on the value 0 if t is less than t_0.

The function $g(t)$ is called a "heaviside" function. Note that whenever t is greater than or equal to t_0, $g(t)$ equals 1, so $E \times g(t) = E$; however, whenever t is less than t_0, $g(t) = 0$, so the value of E $\times$ $g(t)$ is always 0. We will later return to illustrate how heaviside functions may be used as one method for the analysis when a time-independent variable like E does not satisfy the proportional hazards assumption.

Internal variable:

EXAMPLES OF INTERNAL VARIABLES

$E(t)$, $EMP(t)$, $SMK(t)$, $OBS(t)$,

Values change because of "internal" characteristics or behavior of the individual.

Another type of time-dependent variable is called an "internal" variable. Examples of such a variable include exposure level E at time t, employment status (EMP) at time t, smoking status (SMK) at time t, and obesity level (OBS) at time t.

All these examples consider variables whose values may change over time for any subject under study; moreover, for internal variables, the reason for a change in value depends on "internal" characteristics or behavior specific to the individual.

"Ancillary" variable:
Value changes because of "external" characteristics.

In contrast, a variable is called an "ancillary" variable if its value changes primarily because of "external" characteristics of the environment that may affect several individuals simultaneously. An example of an ancillary variable is air pollution index at time t for a particular geographical area. Another example is employment status (EMP) at time t, if the primary reason for whether someone is employed or not depends more on general economic circumstances than on individual characteristics.

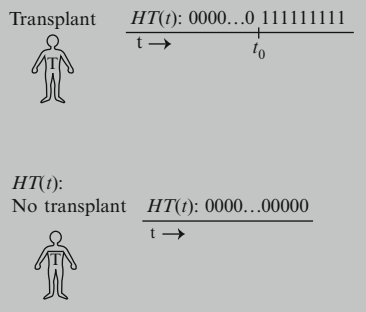
EXAMPLES OF ANCILLARY VARIABLES

Air pollution index at time t; $EMP(t)$

ANOTHER EXAMPLE

Heart transplant status at time t:

$$HT(t) = \begin{cases} 1 & \text{if received transplant at some time } t_0 \le t \\ 0 & \text{if did not receive transplant by time } t \end{cases}$$



As another example, which may be part internal and part ancillary, we consider heart transplant status (HT) at time t for a person identified to have a serious heart condition, making him or her eligible for a transplant. The value of this variable HT at time t is 1 if the person has already received a transplant at some time, say t_0, prior to time t. The value of HT is 0 at time t if the person has not yet received a transplant by time t.

Note that once a person receives a transplant, at time t_0, the value of HT remains at 1 for all subsequent times. Thus, for a person receiving a transplant, the value of HT is 0 up to the time of transplant, and then remains at 1 thereafter. In contrast, a person who never receives a transplant has HT equal to 0 for all times during the period he or she is in the study.

The variable "heart transplant status," $HT(t)$, can be considered essentially an internal variable, because individual traits of an eligible transplant recipient are important determinants of the decision to carry out transplant surgery. Nevertheless, the availability of a donor heart prior to tissue and other matching with an eligible recipient can be considered an "ancillary" characteristic external to the recipient.

Computer commands differ for defined vs. internal vs. ancillary.

But, the form of extended Cox model and procedures for analysis are the same regardless of variable type.

The primary reason for distinguishing among defined, internal, or ancillary variables is that the computer commands required to define the variables for use in an extended Cox model are somewhat different for the different variable types, depending on the computer program used. Nevertheless, the form of the extended Cox model is the same regardless of variable type, and the procedures for obtaining estimates of regression coefficients and other parameters, as well as for carrying out statistical inferences, are also the same.

IV. The Extended Cox Model for Time-Dependent Variables

$$h(t,\mathbf{X}(t)) = h_0(t)\exp\left[\sum_{i=1}^{p_1}\beta_i X_i + \sum_{j=1}^{p_2}\delta_j X_j(t)\right]$$

$$\mathbf{X}(t) = \underbrace{(X_1, X_2, \ldots X_{p_1}}_{\text{Time-independent}}, \underbrace{X_1(t), X_2(t), \ldots X_{p_2}(t))}_{\text{Time-dependent}}$$

Given a survival analysis situation involving both time-independent and time-dependent predictor variables, we can write the extended Cox model that incorporates both types as shown here at the left. As with the Cox PH model, the extended model contains a baseline hazard function $h_0(t)$ which is multiplied by an exponential function. However, in the extended model, the exponential part contains both time-independent predictors, as denoted by the X_i variables, and time-dependent predictors, as denoted by the $X_j(t)$ variables. The entire collection of predictors at time t is denoted by the bold $\mathbf{X}(t)$.

EXAMPLE

$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta(E \times t)]$,
$p_1 = 1, p_2 = 1$,
$\mathbf{X}(t) = (X_1 = E, X_1(t) = E \times t)$

As a simple example of an extended Cox model, we show here a model with one time-independent variable and one time-dependent variable. The time-independent variable is exposure status E, say a (0,1) variable, and the time-dependent variable is the product term $E \times t$.

Estimating regression coefficients:
ML procedure:
Maximize (partial) L.
Risk sets more complicated than for PH model.

As with the simpler Cox PH model, the regression coefficients in the extended Cox model are estimated using a maximum likelihood (ML) procedure. ML estimates are obtained by maximizing a (partial) likelihood function L. However, the computations for the extended Cox model are more complicated than for the Cox PH model, because the risk sets used to form the likelihood function are more complicated with time-dependent variables. The extended Cox likelihood is described later in this chapter.

Computer programs for the extended Cox model:

Stata (Stcox)
SAS (PHREG)
SPSS (COXREG)
R
} Computer Appendix

Computer packages that include programs for fitting the extended Cox model include Stata, SAS, SPSS, and R. See the Computer Appendix at the end of this text for a comparison of the Stata, SAS, SPSS, and R procedures applied to the same dataset.

Statistical inferences:
Wald and/or *LR* tests
Large sample confidence intervals

Methods for making statistical inferences are essentially the same as for the PH model. That is, one can use Wald and/or likelihood ratio (*LR*) tests and large sample confidence interval methods.

Assumption of the model:
The hazard at time t depends on the value of $X_j(t)$ at that same time.

An important assumption of the extended Cox model is that the effect of a time-dependent variable $X_j(t)$ on the survival probability at time t depends on the value of this variable at that *same* time t, and not on the value at an earlier or later time.

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p_2} \delta_j X_j(t)\right]$$

One coefficient for $X_j(t)$

Note that even though the values of the variable $X_j(t)$ may change over time, the hazard model provides only one coefficient for each time-dependent variable in the model. Thus, at time t, there is only one value of the variable $X_j(t)$ that has an effect on the hazard, that value being measured at time t.

Can modify for lag-time effect

It is possible, nevertheless, to modify the definition of the time-dependent variable to allow for a "lag-time" effect.

Lag-time effect:

EXAMPLE

$EMP(t)$ = employment status at week t

Model without lag-time:
$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta EMP(t)]$
↑ Same week

Model with 1-week lag-time:
$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta^* EMP(t-1)]$
↑ One-week earlier

To illustrate the idea of a lag-time effect, suppose, for example, that employment status, measured weekly and denoted as $EMP(t)$, is the time-dependent variable being considered. Then, an extended Cox model that does *not* consider lag-time assumes that the effect of employment status on the probability of survival at week t depends on the observed value of this variable at the same week t, and not, for example, at an earlier week.

However, to allow for, say, a time-lag of one week, the employment status variable may be modified so that the hazard model at time t is predicted by the employment status at week $t-1$. Thus, the variable $EMP(t)$ is replaced in the model by the variable $EMP(t-1)$.

General lag-time extended model:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp\left[\sum_{i=1}^{p_1}\beta_i X_i + \sum_{j=1}^{p_2}\delta_j X_j\big(t-L_j\big)\right]$$

$X_j(t-L_j)$ replaces $X_j(t)$

More generally, the extended Cox model may be alternatively written to allow for a lag-time modification of any time-dependent variable of interest. If we let L_j denote the lag-time specified for time-dependent variable j, then the general "lag-time extended model" can be written as shown here. Note that the variable $X_j(t)$ in the earlier version of the extended model is now replaced by the variable $X_j(t - L_j)$.

V. The Hazard Ratio Formula for the Extended Cox Model

PH assumption is not satisfied for the extended Cox model.

We now describe the formula for the hazard ratio that derives from the extended Cox model. The most important feature of this formula is that the proportional hazards assumption is no longer satisfied when using the extended Cox model.

$$\widehat{HR}(t) = \frac{\hat{h}(t,\mathbf{X}^*(t))}{\hat{h}(t,\mathbf{X}(t))} = \exp\left[\sum_{i=1}^{p_1}\hat{\beta}_i\left[X_i^* - X_i\right] + \sum_{j=1}^{p_2}\delta_j\left[X_j^*(t) - X_j(t)\right]\right]$$

The general hazard ratio formula for the extended Cox model is shown here. This formula describes the ratio of hazards at a particular time t, and requires the specification of two sets of predictors at time t. These two sets are denoted as bold $\mathbf{X}^*(t)$ and bold $\mathbf{X}(t)$.

Two sets of predictors:

$$\mathbf{X}^*(t) = \left(X_1^*, X_2^*, \ldots, X_{p_1}^*, X_1^*(t), X_2^*(t), \ldots, X_{p_2}^*(t)\right)$$

$$\mathbf{X}(t) = \left(X_1, X_2, \ldots, X_{p_1}, X_1(t), X_2(t), \ldots, X_{p_2}(t)\right)$$

The two sets of predictors, $\mathbf{X}^*(t)$ and $\mathbf{X}(t)$, identify two specifications at time t for the combined set of predictors containing both time-independent and time-dependent variables. The individual components for each set of predictors are shown here.

EXAMPLE

$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta(E \times t)]$

$$E = \begin{cases} 1 & \text{if exposed} \\ 0 & \text{if unexposed} \end{cases}$$

$\mathbf{X}^*(t) = (E = 1, E \times t = t)$
$\mathbf{X}(t) = (E = 0, E \times t = 0)$

$$\widehat{HR}(t) = \frac{\hat{h}(t,E=1)}{\hat{h}(t,E=0)} = \exp\left[\hat{\beta}(1-0) + \hat{\delta}((1\times t) - (0\times t))\right] = \exp\left[\hat{\beta} + \hat{\delta}t\right]$$

$\hat{\delta}_0 \Rightarrow \widehat{HR}(t) \uparrow as\ t \uparrow$
PH assumption not satisfied

As a simple example, suppose the model contains only one time-independent predictor, namely, exposure status E, a (0,1) variable, and one time-dependent predictor, namely, $E \times t$. Then, to compare exposed persons, for whom $E = 1$, with unexposed persons, for whom $E = 0$, at time t, the bold $\mathbf{X}^*(t)$ set of predictors has as its two components $E = 1$ and $E \times t = t$; the bold $\mathbf{X}(t)$ set has as its two components $E = 0$ and $E \times t = 0$.

If we now calculate the estimated hazard ratio that compares exposed to unexposed persons at time t, we obtain the formula shown here; that is, *HR* "hat" equals the exponential of β "hat" plus δ "hat" times t. This formula says that the hazard ratio is a function of time; in particular, if δ "hat" is positive, then the hazard ratio increases with increasing time. Thus, the hazard ratio in this example is certainly not constant, so that the PH assumption is not satisfied for this model.

$$\widehat{HR}(t) = \exp\left[\sum_{i=1}^{p_1} \hat{\beta}_i\left[X_i^* - X_i\right] + \sum_{j=1}^{p_2} \boxed{\hat{\delta}_j}\boxed{\left[X_j^*(t) - X_j(t)\right]}\right]$$

A function of time

More generally, because the general hazard ratio formula involves differences in the values of the time-dependent variables at time t, this hazard ratio is a function of time. Thus, in general, the extended Cox model does not satisfy the PH assumption if any δ_j is not equal to zero.

In general, PH assumption not satisfied for extended Cox model.

$\hat{\delta}_j$ is not time-dependent.
$\hat{\delta}_j$ represents "overall" effect of $X_j(t)$.

Note that, in the hazard ratio formula, the coefficient δ_j "hat" of the difference in values of the jth time-dependent variable is itself not time-dependent. Thus, this coefficient represents the "overall" effect of the corresponding time-dependent variable, considering all times at which this variable has been measured in the study.

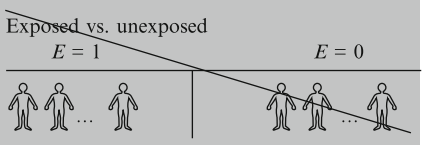
EXAMPLE

$E(t)$ = chemical exposure status at time t (weekly)

$$= \begin{cases} 0 & \text{if unexposed at time } t \\ 1 & \text{if exposed at time } t \end{cases}$$

$E(t)$	0 1 0 1 1
t	1 2 3 4 5…

$E(t)$	1 1 0 1 0
t	1 2 3 4 5…



$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta E(t)]$$

One coefficient

δ represents the overall effect of $E(t)$.

$$\widehat{HR}(t) = \frac{\hat{h}(t, E(t)=1)}{\hat{h}(t, E(t)=0)} = \exp\left[\hat{\delta}[1-0]\right] = e^{\hat{\delta}}, \text{ a fixed number}$$

But, PH is *not* satisfied:
$\widehat{HR}(t)$ is time-dependent because $E(t)$ is time-dependent.

As another example to illustrate the formula for the hazard ratio, consider an extended Cox model containing only one variable, say a weekly measure of chemical exposure status at time t. Suppose this variable, denoted as $E(t)$, can take one of two values, 0 or 1, depending on whether a person is unexposed or exposed, respectively, at a given weekly measurement.

As defined, the variable $E(t)$ can take on different patterns of values for different subjects. For example, for a 5-week period, subject A's values may be 01011, whereas subject B's values may be 11010.

Note that in this example, we do not consider two separate groups of subjects, with one group always exposed and the other group always unexposed throughout the study. This latter situation would require a (0,1) time-independent variable for exposure, whereas our example involves a time-dependent exposure variable.

The extended Cox model that includes only the variable $E(t)$ is shown here. In this model, the values of the exposure variable may change over time for different subjects, but there is only one coefficient, δ, corresponding to the one variable in the model. Thus, δ represents the overall effect on survival time of the time-dependent variable $E(t)$.

Notice, also, that the hazard ratio formula, which compares an exposed person to an unexposed person at time t, yields the expression e to the δ "hat."

Although this result is a fixed number, the PH assumption is not satisfied. The fixed number gives the hazard ratio at a given time, assuming that the exposure status at that time is 1 in the numerator and is 0 denominator. Thus, the hazard ratio is time-dependent, because exposure status is time-dependent, even though the formula yields a single fixed number.

VI. Assessing Time-Independent Variables That Do Not Satisfy the PH Assumption

Use an extended Cox model to

- check PH assumption;
- assess effect of variable not satisfying PH assumption.

We now discuss how to use an extended Cox model to check the PH assumption for time-independent variables and to assess the effect of a variable that does not satisfy the PH assumption.

Three methods for checking PH assumption:

1. graphical
2. extended Cox model
3. GOF test

As described previously (see Chapter 4), there are three methods commonly used to assess the PH assumption: (1) graphical, using, say, log–log survival curves; (2) using an extended Cox model; and (3) using a goodness-of-fit (GOF) test. We have previously (in Chapter 4) discussed items 1 and 3, but only briefly described item 2, which we focus on here.

Cox PH model for p time-independent X's:

$$h(t, \mathbf{X}) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$$

If the dataset for our study contains several, say p, time-independent variables, we might wish to fit a Cox PH model containing each of these variables, as shown here.

Extended Cox model:
Add product terms of the form: $X_i \times g_i(t)$

However, to assess whether such a PH model is appropriate, we can extend this model by defining several product terms involving each time-independent variable with some function of time. That is, if the ith time-independent variable is denoted as X_i, then we can define the ith product term as $X_i \times g_i(t)$ where $g_i(t)$ is some function of time for the ith variable.

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i + \sum_{i=1}^{p} \delta_i X_i g_i(t)\right]$$

The extended Cox model that simultaneously considers all time-independent variables of interest is shown here.

EXAMPLE

$g_i(t) = 0$ for all i implies no time-dependent variable involving X_i, i.e.,

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$$

In using this extended model, the crucial decision is the form that the functions $g_i(t)$ should take. The simplest form for $g_i(t)$ is that all $g_i(t)$ are identically 0 at any time; this is another way of stating the original PH model, containing no time-dependent terms.

EXAMPLE 2

$g_i(t) = t \Rightarrow X_i g(t) = X_\mathrm{i} \times \mathrm{t}$

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i + \sum_{i=1}^{p} \delta_i (X_i \times t)\right]$$

Another choice for the $g_i(t)$ is to let $g_i(t) = t$. This implies that for each X_i in the model as a main effect, there is a corresponding time-dependent variable in the model of the form $X_\mathrm{i} \times \mathrm{t}$. The extended Cox model in this case takes the form shown here.

EXAMPLE 3: one variable at a time

$$X_L \text{ only} \Rightarrow \begin{cases} g_L(t) = t, \\ g_i(t) = 0 \text{ for other } i \end{cases}$$

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{\substack{\mathrm{i}=1 \\ \mathrm{i} \neq \mathrm{L}}}^{p} \beta_i \mathbf{X}_i + \delta_L (\mathbf{X}_L \times t)\right]$$

Suppose, however, we wish to focus on a particular time-independent variable, say, variable X_L. Then $g_i(t) = t$ for $i = L$, but equals 0 for all other i. The corresponding extended Cox model would then contain only one product term $X_L \times \mathrm{t}$, as shown here.

EXAMPLE 4

$g_i(t) = \ln t \Rightarrow X_i g(t) = X_i \times \ln t$

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i + \sum_{i=1}^{p} \delta_i (X_i \times \ln t)\right]$$

Another choice for the $g_i(t)$ is the log of t, rather than simply t, so that the corresponding time-dependent variables will be of the form $X_i \times \ln t$.

EXAMPLE 5: Heaviside function

$$g_i(t) = \begin{cases} 0 & \text{if } t \geq t_0 \\ 1 & \text{if } t > \leq t_0 \end{cases}$$

And yet another choice would be to let $g_i(t)$ be a "heaviside function" of the form $g_i(t) = 1$ when t is at or above some specified time, say t_0, and $g_i(t) = 0$ when t is below t_0. We will discuss this choice in more detail shortly.

Extended Cox model:

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i + \sum_{i=1}^{p} \delta_i X_i g_i(t)\right]$$

Given a particular choice of the $g_i(t)$, the corresponding extended Cox model, shown here again in general form, may then be used to check the PH assumption for the time-independent variables in the model. Also, we can use this extended Cox model to obtain a hazard ratio formula that considers the effects of variables not satisfying the PH assumption.

- Check PH assumption.
- Obtain hazard ratio when PH assumption not satisfied.

H_0: $\delta_1 = \delta_2 = \ldots = \delta_p = 0$

To check the PH assumption using a statistical test, we consider the null hypothesis that all the δ terms, which are coefficients of the $X_i g_i(t)$ product terms in the model, are zero.

Under H_0, the model reduces to PH model:

$$h(t, \mathbf{X}) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$$

Under this null hypothesis, the model reduces to the PH model.

$$LR = -2 \ln L_{\text{PH model}} - (-2 \ln L_{\text{ext. Cox model}}) \dot{\sim} \chi^2_p \text{ under } H_0$$

This test can be carried out using a likelihood ratio (LR) test which computes the difference between the log likelihood statistic, $-2 \ln L$, for the PH model and the log likelihood statistic for the extended Cox model. The test statistic thus obtained has approximately a chi-square distribution with p degrees of freedom under the null hypothesis, where p denotes the number of parameters being set equal to zero under H_0.

EXAMPLE

$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta(E \times t)]$
H_0: $\delta = 0$ (i.e., PH assumption is satisfied)

Reduced model:
$h(t,\mathbf{X}) = h_0(t)\exp[\beta E]$

$LR = -2 \ln L_R - (-2 \ln L_F) \dot{\sim} \chi^2$ with 1 df under H_0

F = full (extended), R = reduced (PH)

As an example of this test, suppose we again consider an extended Cox model that contains the product term $E \times t$ in addition to the main effect of E, where E denotes a (0,1) time-independent exposure variable.

For this model, a test for whether or not the PH assumption is satisfied is equivalent to testing the null hypothesis that $\delta = 0$. Under this hypothesis, the reduced model is given by the PH model containing the main effect E only. The likelihood ratio statistic, shown here as the difference between log-likelihood statistics for the full (i.e., extended model) and the reduced (i.e., PH) model, will have an approximate chi-square distribution with one degree of freedom in large samples.

SAS: **PHREG** fits both PH and extended Cox models.
Stata: **Stcox** fits both PH and extended Cox models.

Note that to carry out the computations for this test, two different types of models, a PH model and an extended Cox model, need to be fit. Nevertheless, some computer packages such as SAS and Stata use the same computer program to fit both models.

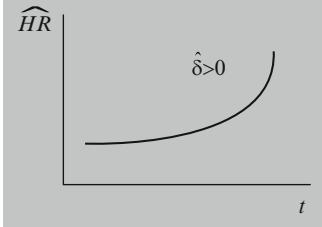
If PH test significant: Extended Cox model is preferred; HR is time-dependent.

If the result of the test for the PH assumption is significant, then the extended Cox model is preferred to the PH model. Thus, the hazard ratio expression obtained for the effect of an exposure variable of interest is time-dependent. That is, the effect of the exposure on the outcome cannot be summarized by a single HR value, but can only be expressed as a function of time.

EXAMPLE

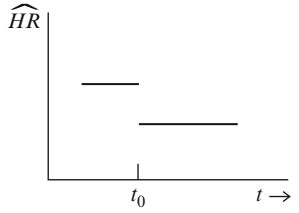
$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta(E \times t)]$

$\widehat{HR} = \exp\left[\hat{\beta} + \hat{\delta}t\right]$



We again consider the previous example, with the extended Cox model shown here. For this model, the estimated hazard ratio for the effect of exposure is given by the expression e to the quantity β "hat" plus δ "hat" times t. Thus, depending on whether δ "hat" is positive or negative, the estimated hazard ratio will increase or decrease exponentially as t increases. The graph shown here gives a sketch of how the hazard ratio varies with time if δ "hat" is positive.

Heaviside function:



We now provide a description of the use of a "heaviside" function. When such a function is used, the hazard ratio formula yields constant hazard ratios for different time intervals, as illustrated in the accompanying graph.

$$g(t) = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t \leq t_0 \end{cases}$$

$h(t, \mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta E g(t)]$

Recall that a heaviside function is of the form $g(t)$, which takes on the value 1 if t is greater than or equal to some specified value of t, called t_0, and takes on the value 0 if t is less than t_0. An extended Cox model which contains a single heaviside function is shown here.

$t \geq t_0$: $g(t) = 1 \Rightarrow E \times g(t) = E$

$h(t, \mathbf{X}) = h_0(t)\exp[(\beta + \delta)E]$

$\widehat{HR} = \exp\left[\hat{\beta} + \hat{\delta}\right]$

Note that if $t \geq t_0$, $g(t) = 1$, so the value of $E \times g(t) = E$; the corresponding hazard function is of the form $h_0(t) \times e$ to the quantity $(\beta + \delta)$ times E, and the estimated hazard ratio for the effect of E has the form e to the sum of β "hat" plus δ "hat."

$t < t_0$: $g(t) = 0 \Rightarrow E \times g(t) = 0$

$h(t,\mathbf{X}) = h_0(t)\exp[\beta E]$

$\widehat{HR} = \exp\left[\hat{\beta}\right]$

If $t < t_0$, $g(t) = 0$, the corresponding hazard ratio is simplified to e to the β "hat."

A single heaviside function in the model

$h(t,\mathbf{X})$

$= h_0(t)\exp[\beta E + \delta(E \times g(t))]$

yields two hazard ratios:

$t \geq t_0$: $\widehat{HR} = \exp\left[\hat{\beta} + \hat{\delta}\right]$

$t \leq t_0$: $\widehat{HR} = \exp\left[\hat{\beta}\right]$

Thus, we have shown that the use of a single heaviside function results in an extended Cox model which gives two hazard ratio values, each value being constant over a fixed time interval.

Alternative model with two heaviside functions:

$$h(t,\mathbf{X}) = h_0(t)\exp[\delta_1(E \times g_1(t)) + \delta_2(E \times g_2(t))]$$

$$g_1(t) = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t \leq t_0 \end{cases}$$

$$g_2(t) = \begin{cases} 1 & \text{if } t \leq t_0 \\ 0 & \text{if } t \geq t_0 \end{cases}$$

There is actually an equivalent way to write this model that uses two heaviside functions in the same model. This alternative model is shown here. The two heaviside functions are called $g_1(t)$ and $g_2(t)$. Each of these functions are in the model as part of a product term with the exposure variable E. Note that this model does not contain a main effect term for exposure.

Note: Main effect for E not in model.

Two *HR*'s from the alternative model:

$$t \geq t_0 : \; g_1(t) = 1, \; g_2(t) = 0$$
$$h(t,\mathbf{X}) = h_0(t) \exp[\delta_1(E \times 1) + \delta_2(E \times 0)]$$
$$= h_0(t)\exp[\delta_1 E]$$

so that $\widehat{HR} = \exp\left(\hat{\delta}_1\right)$

For this alternative model, as for the earlier model with only one heaviside function, two different hazard ratios are obtained for different time intervals. To obtain the first hazard ratio, we consider the form that the model takes when $t \geq t_0$. In this case, the value of $g_1(t)$ is 1 and the value of $g_2(t)$ is 0, so the exponential part of the model simplifies to $\delta_1 \times E$; the corresponding formula for the estimated hazard ratio is then e to the δ_1 "hat."

$$t > t_0 : g_1(t) = \; 0, \; g_2(t) = 1$$
$$h(t,\mathbf{X}) = h_0(t)\exp[\delta_1(E \times 0) + \delta_2(E \times 1)]$$
$$= h_0(t)\exp[\delta_2 E]$$

so that $\widehat{HR} = \exp\left(\hat{\delta}_2\right)$

When $t < t_0$, the value of $g_1(t)$ is 0 and the value of $g_2(t)$ is 1. Then, the exponential part of the model becomes $\delta_2 \times E$, and the corresponding hazard ratio formula is e to the δ_2 "hat."

Alternative model:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta_1(E \times g_1(t)) + \delta_2(E \times g_2(t))]$$

Original model:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta(E \times g(t))]$$

$$t \geq t_0 : \widehat{HR} = \exp\left(\hat{\delta}_1\right) = \exp\left(\hat{\beta} + \hat{\delta}\right)$$

$$t \leq t_0 : \widehat{HR} = \exp\left(\hat{\delta}_2\right) = \exp\left(\hat{\beta}\right)$$

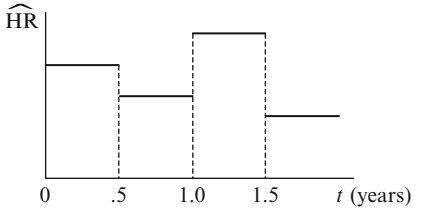
Thus, using the alternative model, again shown here, we obtain two distinct hazard ratio values. Mathematically, these are the same values as obtained from the original model containing only one heaviside function. In other words, δ_1 "hat" in the alternative model equals β "hat" plus δ "hat" in the original model (containing one heaviside function), and δ_2 "hat" in the alternative model equals β "hat" in the original model.

Heaviside functions:

- two $\widehat{HR}$'s constant within two time intervals
- *extension*: several $\widehat{HR}$'s constant within several time intervals

We have thus seen that heaviside functions can be used to provide estimated hazard ratios that remain constant within each of two separate time intervals of follow-up. We can also extend the use of heaviside functions to provide several distinct hazard ratios that remain constant within several time intervals.

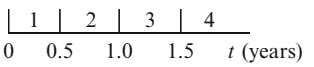
Four time intervals:



Suppose, for instance, that we wish to separate the data into *four* separate time intervals, and for each interval, we wish to obtain a different hazard ratio estimate as illustrated in the graph shown here.

Extended Cox model contains either

- E, $E \times g_1(t)$, $E \times g_2(t)$, $E \times g_3(t)$ or
- $E \times g_1(t)$, $E \times g_2(t)$, $E \times g_3(t)$, $E \times g_4(t)$



We can obtain four different hazard ratios using an extended Cox model containing *a main effect of exposure and three heaviside functions* in the model as products with exposure. Or, we can use a model containing *no main effect* exposure term, but with product terms involving exposure with *four heaviside functions*.

To illustrate the latter model, suppose, as shown on the graph, that the first time interval goes from time 0 to 0.5 of a year; the second time interval goes from 0.5 to 1 year; the third time interval goes from 1 year to a year and a half; and the fourth time interval goes from a year and a half onward.

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta_1 E g_1(t) + \delta_2 E g_2(t) + \delta_3 E g_3(t) + \delta_4 E g_4(t)]$$

where

$$g_1(t) = \begin{cases} 1 & \text{if } 0 \le t < 0.5 \text{ year} \\ 0 & \text{if otherwise} \end{cases}$$

$$g_2(t) = \begin{cases} 1 & \text{if } 0.5 \text{ year } \le t < 1.0 \text{ year} \\ 0 & \text{if otherwise} \end{cases}$$

$$g_3(t) = \begin{cases} 1 & \text{if } 1.0 \text{ year } \le t < 1.5 \text{ years} \\ 0 & \text{if otherwise} \end{cases}$$

$$g_4(t) = \begin{cases} 1 & \text{if } t \ge 1.5 \text{ years} \\ 0 & \text{if otherwise} \end{cases}$$

Then, an appropriate extended Cox model containing the four heaviside functions $g_1(t)$, $g_2(t)$, $g_3(t)$, and $g_4(t)$ is shown here. This model assumes that there are four different hazard ratios identified by three cutpoints at half a year, one year, and one and a half years. The formulae for the four hazard ratios are given by separately exponentiating each of the four estimated coefficients, as shown below:

$$4\,\widehat{\text{HR}}\text{'s} \begin{cases} 0 \le t < 0.5: \ \widehat{HR} = \exp\left(\hat{\delta}_1\right) \\ 0.5 \le t < 1.0: \ \widehat{HR} = \exp\left(\hat{\delta}_2\right) \\ 1.0 \le t < 1.5: \ \widehat{HR} = \exp\left(\hat{\delta}_3\right) \\ t \ge 1.5: \ \widehat{HR} = \exp\left(\hat{\delta}_4\right) \end{cases}$$

VII. An Application of the Extended Cox Model to An Epidemiologic Study on the Treatment of Heroin Addiction

EXAMPLE

1991 Australian study (Caplehorn et al.) of heroin addicts

- two methadone treatment clinics
- T = days remaining in treatment (= days until drop out of clinic)
- clinics differ in treatment policies

Dataset name: ADDICTS
Column 1: Subject ID
Column 2: Clinic (1 or 2) ← E
Column 3: Survival status (0 = censored, 1 = departed clinic)
Column 4: Survival time in days
Column 5: Prison Record (0 = none, 1 = any) ← covariates
Column 6: Maximum Methadone Dose (mg/day)

A 1991 Australian study by Caplehorn et al., compared retention in two methadone treatment clinics for heroin addicts. A patient's survival time (T) was determined as the time in days until the patient dropped out of the clinic or was censored at the end of the study clinic. The two clinics differed according to their overall treatment policies.

A listing of some of the variables in the dataset for this study is shown here. The dataset name is called "ADDICTS," and survival analysis programs in the Stata package are used in the analysis. Note that the survival time variable is listed in column 4 and the survival status variable, which indicates whether a patient departed from the clinic or was censored, is listed in column 3. The primary exposure variable of interest is the clinic variable, which is coded as 1 or 2. Two other variables of interest are prison record status, listed in column 5 and coded as 0 if none and 1 if any, and maximum methadone dose, in milligrams per day, which is listed in column 6. These latter two variables are considered as covariates.

$$h(t,\mathbf{X}) = h_0(t)\exp[\beta_1(\text{clinic}) + \beta_2(\text{prison}) + \beta_3(\text{dose})]$$

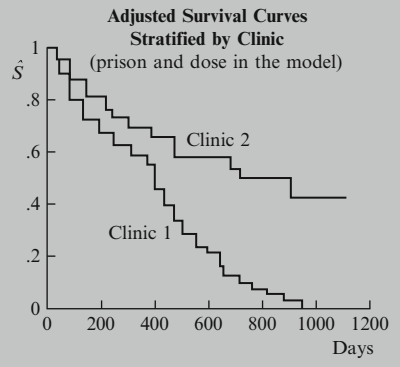
	Coef.	Std. Err.	p > |z|	Haz. Ratio	P (PH)
Clinic	−1.009	0.215	0.000	0.365	0.001
Prison	0.327	0.167	0.051	1.386	0.332
Dose	−0.035	0.006	0.000	0.965	0.347

One of the first models considered in the analysis of the addicts dataset was a Cox PH model containing the three variables, clinic, prison record, and dose. An edited printout of the results for this model is shown here. What stands out from this printout is that the $P(PH)$ value for the clinic variable is zero to three significant places, which indicates that the clinic variable does not satisfy the proportional hazard assumption.

$P(PH)$ for the variables prison and dose are nonsignificant ⇒ remain in model.

Since the $P(PH)$ values for the other two variables in the model are highly nonsignificant, this suggests that these two variables, namely, prison and dose, can remain in the model.

EXAMPLE: (continued)



Results:

- Curve for clinic 2 consistently lies above curve for clinic 1.
- Curves diverge, with clinic 2 being vastly superior after one year.

Stratifying by **clinic**: cannot obtain hazard ratio for **clinic**

Hazard ratio for **clinic** requires **clinic** in the model.

Extended Cox model:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{clinic}) + \beta_2(\text{prison}) + \beta_3(\text{dose}) + \delta(\text{clinic})\, g(t)]$$

where

$$g(t) = \begin{cases} 1 & \text{if } t \geq 365 \text{ days} \\ 0 & \text{if } t < 365 \text{ days} \end{cases}$$

and

$$\text{clinic} = \begin{cases} 1 & \text{if clinic 1} \\ 0 & \text{if clinic 2} \end{cases}$$

None Previously clinic = 2 for clinic 2

$t \geq 365$ days: $HR = \exp\left(\hat{\beta}_1 + \hat{\delta}\right)$

$t < 365$ days: $HR = \exp\left(\hat{\beta}_1\right)$

Further evidence of the PH assumption not being satisfied for the clinic variable can be seen from a graph of adjusted survival curves stratified by clinic, where the prison and dose variables have been kept in the model. Notice that the two curves are much closer together at earlier times, roughly less than 1 year (i.e., 365 days), but the two curves diverge greatly after 1 year. This indicates that the hazard ratio for the clinic variable will be much closer to one at early times but quite different from one later on.

The above graph, nevertheless, provides important results regarding the comparison of the two clinics. The curve for clinic 2 consistently lies above the curve for clinic 1, indicating that clinic 2 does better than clinic 1 in retaining its patients in methadone treatment. Further, because the two curves diverge after about a year, it appears that clinic 2 is vastly superior to clinic 1 after one year but only slightly better than clinic 1 prior to one year.

Unfortunately, because the clinic variable has been stratified in the analysis, we cannot use this analysis to obtain a hazard ratio expression for the effect of clinic, adjusted for the effects of prison and dose. We can only obtain such an expression for the hazard ratio if the clinic variable is in the model.

Nevertheless, we can obtain a hazard ratio using an alternative analysis with an extended Cox model that contains a heaviside function, $g(t)$, together with the clinic variable, as shown here. Based on the graphical results shown earlier, a logical choice for the cutpoint of the heaviside function is one year (i.e., 365 days). The corresponding model then provides two hazard ratios: one that is constant above 365 days and the other that is constant below 365 days.

Note that in the extended Cox model here, we have coded the clinic variable as 1 if clinic 1 and 0 if clinic 2, whereas previously we had coded clinic 2 as 2. The reason for this change in coding, as illustrated by computer output below, is to obtain hazard ratio estimates that are greater than unity.

EXAMPLE: (continued)

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_2(\text{prison}) + \beta_3(\text{dose}) + \beta_1(\text{clinic})g_1(t) + \delta_2(\text{clinic})g_2(t)]$$

where

$$g_1(t) = \begin{cases} 1 & \text{if } t > 365 \text{ days} \\ 0 & \text{if } t \geq 365 \text{ days} \end{cases}$$

and

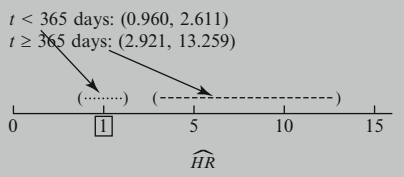
$$g_2(t) = \begin{cases} 1 & \text{if } t \geq 365 \text{ days} \\ 0 & \text{if } t > 365 \text{ days} \end{cases}$$

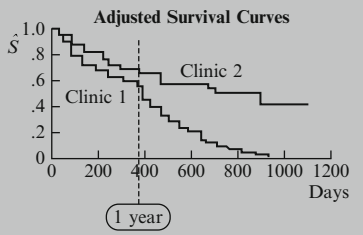
$t < 365$ days: $\widehat{HR} = \exp(\hat{\delta}_1)$
$t \geq 365$ days: $\widehat{HR} = \exp(\hat{\delta}_2)$

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Prison	0.378	0.168	0.025	1.459	1.049	2.029
Dose	−0.036	0.006	0.000	0.965	0.953	0.977
Clinic × g_1	0.460	0.255	0.072	1.583	0.960	2.611
Clinic × g_2	1.828	0.386	0.000	6.223	2.921	13.259

$t < 365$ days: $\widehat{HR} = e^{0.460} = 1.583$
$t \geq 365$ days: $[\widehat{HR} = e^{1.828} = 6.223$

95% confidence intervals for clinic effect:





An equivalent way to write the model is to use two heaviside functions, $g_1(t)$ and $g_2(t)$, as shown here. This latter model contains product terms involving clinic with each heaviside function, and there is no main effect of clinic.

Corresponding to the above model, the effect of clinic is described by two hazard ratios, one for time less than 365 days and the other for greater than 365 days. These hazard ratios are obtained by separately exponentiating the coefficients of each product term, yielding e to the δ_1 "hat" and e to the δ_2 "hat," respectively.

A printout of results using the above model with two heaviside functions is provided here. The results show a borderline nonsignificant hazard ratio ($P = 0.072$) of 1.6 for the effect of clinic when time is less than 365 days in contrast to a highly significant (P = 0.000 to three decimal places) hazard ratio of 6.2 when time exceeds 365 days.

Note that the estimated hazard ratio of 1.583 from the printout is computed by exponentiating the estimated coefficient 0.460 of the product term "clinic × g_1" and that the estimated hazard ratio of 6.223 is computed by exponentiating the estimated coefficient 1.828 of the product term "clinic × g_2".

Note also that the 95% confidence interval for the clinic effect prior to 365 days (that is, for the product term "clinic × $g_1(t)$") is given by the limits 0.960 and 2.611, whereas the corresponding confidence interval after 365 days (that is, for the product term "clinic × $g_2(t)$") is given by the limits 2.921 and 13.259. The latter interval is quite wide, showing a lack of precision when t exceeds 365 days; however, when t precedes 365 days, the interval includes the null hazard ratio of 1, suggesting a chance effect for this time period.

The results we have just shown support the observations obtained from the graph of adjusted survival curves. That is, these results suggest a large difference in clinic survival times after 1 year in contrast to a small difference in clinic survival times prior to 1 year, with clinic 2 always doing better than clinic 1 at any time.

EXAMPLE: (continued)

One other analysis:
Use an extended Cox model that provides for diverging survival curves

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{clinic}) + \beta_2(\text{prison}) + \beta_3(\text{dose}) + \delta(\text{clinic} \times \text{t})]$$

$$\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + \hat{\delta}\text{t}\right)$$

$\widehat{\text{HR}}$ changes over time.

$t = 91$ days

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{clinic}) + \beta_2(\text{prison}) + \beta_3\,(\text{dose}) + \delta(\text{clinic})(91)]$$

So

$$\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + 91\hat{\delta}\right)$$

$t = 274$:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{clinic}) + \beta_2(\text{prison}) + \beta_3(\text{dose}) + \delta(\text{clinic})(274)]$$

$$\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + 274\hat{\delta}\right)$$

$t = 458.5$:

$$\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + 458.5\hat{\delta}\right)$$

$t = 639$:

$$\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + 639\hat{\delta}\right)$$

$t = 821.5$:

$$\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + 821.5\hat{\delta}\right)$$

$\hat{\delta} > 0 \Rightarrow \widehat{HR}\uparrow$ as time $\uparrow$

There is, nevertheless, at least one other approach to the analysis using time-dependent variables that we now describe. This approach considers our earlier graphical observation that the survival curves for each clinic continue to diverge from one another even after 1 year. In other words, it is reasonable to consider an extended Cox model that allows for such a divergence, rather than a model that assumes the hazard ratios are constant before and after 1 year.

One way to define an extended Cox model that provides for diverging survival curves is shown here. This model includes, in addition to the clinic variable by itself, a time-dependent variable defined as the product of the clinic variable with time (i.e., clinic $\times$ t). By including this product term, we are able to estimate the effect of clinic on survival time, and thus the hazard ratio, for any specified time t.

To demonstrate how the hazard ratio changes over time for this model, we consider what the model and corresponding estimated hazard ratio expression are for different specified values of t.

For example, if we are interested in the effect of clinic on survival on day 91, so that t = 91, the exponential part of the model simplifies to terms for the prison and dose variables plus β_1 times the clinic variable plus δ times the clinic variable times 91: the corresponding estimated hazard ratio for the clinic effect is then e to the power β_1 "hat" plus δ "hat" times t = 91.

At 274 days, the exponential part of the model contains the prison, dose, and clinic main effect terms as before, plus δ times the clinic variable times 274: the corresponding hazard ratio for the clinic effect is then e to β_1 "hat" plus 274 δ "hat".

The formulae for the estimated hazard ratio for other specified days are shown here. Notice that the estimated hazard ratio appears to be increase over the length of the follow-up period. Thus, if δ "hat" is a positive number, then the estimated hazard ratios will increase over time.

EXAMPLE: (continued)

Computer results for extended Cox model involving $T(t)$:

	Coef.	Std. Err.	P>|z|	Haz. Ratio	[95% Conf.	Interval]
prison	0.390	0.169	0.021	1.476	1.060	2.056
dose	−0.035	0.006	0.000	0.965	0.953	0.978
clinic	(−0.0183)	0.347	0.958	0.982	0.497	1.939
clinic × t	(0.003)	0.001	0.001	(1.003)	1.001	1.005

$\widehat{\text{cov}}\left(\hat{\beta}_1, \hat{\delta}\right) = -.000259$ Log likelihood = −667.642

$\hat{\beta}_1 = -0.0183 \qquad \hat{\delta} = 0.003$

$\widehat{\text{HR}}$ depends on $\hat{\beta}_1$ and $\hat{\delta}$.

$t = 91.5$: $\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + \hat{\delta}t\right) = 1.292$

$t = 274$: $\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + \hat{\delta}t\right) = 2.233$

$t = 458.5$: $\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + \hat{\delta}t\right) = 3.862$

$t = 639$: $\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + \hat{\delta}t\right) = 6.677$

$t = 821.5$: $\widehat{\text{HR}} = \exp\left(\hat{\beta}_1 + \hat{\delta}t\right) = 11.544$

$$\exp\left[\hat{\beta}_1 + \hat{\delta}t \pm 1.96\sqrt{\widehat{\text{var}}\left(\hat{\beta}_1 + \hat{\delta}t\right)}\right]$$

$$\text{Var}\left(\hat{\beta}_1 + \hat{\delta}t\right) = s^2_{\hat{\beta}_1} + t^2 s^2_{\hat{\delta}} + 2t\,\widehat{\text{cov}}\left(\hat{\beta}_1, \hat{\delta}\right)$$

$(0.347)^2 \quad (0.001)^2 \quad (-.000259)$

Time (days)	$\widehat{\text{HR}}$	95% CI
91.5	1.292	(0.741, 2.250)
274	2.233	(1.470, 3.391)
458.5	3.862	(2.298, 6.491)
639	6.677	(3.102, 14.372)
821.5	11.544	(3.976, 33.513)

We now show edited results obtained from fitting the extended Cox model we have just been describing, which contains the product of clinic with time. The covariance estimate shown at the bottom of the table will be used below to compute confidence intervals.

From these results, the estimated coefficient of the clinic variable is β_1 "hat" equals −0.0183, and the estimated coefficient δ "hat" obtained for the product term equals 0.003. For the model being fit, the hazard ratio depends on the values of both β_1 "hat" and δ "hat."

On the left, the effect of the variable clinic is described by five increasing hazard ratio estimates corresponding to each of five different values of t. These values, which range between 1.292 at 91.5 days to 11.544 at 821.5 days, indicate how the effect of clinic diverges over time for the fitted model.

We can also obtain 95% confidence intervals for each of these hazard ratios using the large sample formula shown here. The variance expression in the formula is computed using the variances and covariances which can be obtained from the computer results given above. In particular, the variances are $(0.347)^2$ and $(0.001)^2$ for β_1 "hat" and δ "hat," respectively; the covariance value is −0.000259.

A table showing the estimated hazard ratios and their corresponding 95% confidence intervals for the clinic effect is given here. Note that all confidence intervals are quite wide.

VIII. An Application of the Extended Cox Model to the Analysis of the Stanford Heart Transplant Data

EXAMPLE

Patients identified as eligible for heart transplant:
T = time until death or censorship
65 patients receive transplants
38 patients do not receive transplants
n = 103 patients

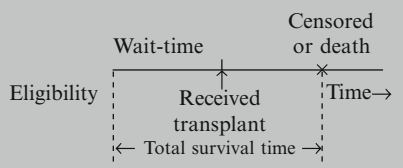
Goal: Do patients receiving transplants survive longer than patients not receiving transplants?

We now consider another application of the extended Cox model which involves the use of an internally defined time-dependent variable. In a 1977 report (Crowley and Hu, *J. Amer. Statist. Assoc.*) on the Stanford Heart Transplant Study, patients identified as being eligible for a heart transplant were followed until death or censorship. Sixty-five of these patients received transplants at some point during follow-up, whereas thirty-eight patients did not receive a transplant. There were, thus, a total of n = 103 patients. The goal of the study was to assess whether patients receiving transplants survived longer than patients not receiving transplants.

One approach:
Compare two separate groups: 65 transplants vs. 38 nontransplants

One approach to the analysis of these data was to separate the dataset into two separate groups, namely, the 65 heart transplant patients and the 38 patients not receiving transplants, and then to compare survival times for these groups.

Problem:



Note: Wait-time contributes to survival time for nontransplants.

A problem with this approach, however, is that those patients who received transplants had to wait from the time they were identified as eligible for a transplant until a suitable transplant donor was found. During this "wait-time" period, they were at risk for dying, yet they did not have the transplant. Thus, the wait-time accrued by transplant patients contributes information about the survival of nontransplant patients. Yet, this wait-time information would be ignored if the *total* survival time for each patient were used in the analysis.

Covariates:

Tissue mismatch score, Age at transplant } prognostic only for transplants

Another problem with this approach is that two covariates of interest, namely, *tissue mismatch score* and *age at transplant*, were considered as prognostic indicators of survival only for patients who received transplants. Note that *age at eligibility* was not considered an important prognostic factor for the nontransplant group.

Age at eligibility: not considered prognostic for nontransplants

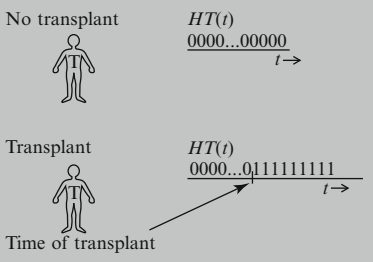
EXAMPLE: (continued)

Problems:
- wait-time of transplant recipients
- prognostic factors for transplants only

Alternative approach:
Uses an extended Cox model

Exposure variable:
Heart transplant status at time t, defined as

$$HT(t) = \begin{cases} 0 & \text{if did not receive transplant by time } t, \text{ i.e., if } t < \text{wait-time} \\ 1 & \text{if received transplant prior to time } t, \text{ i.e., if } t \geq \text{wait-time} \end{cases}$$



Wait-time for transplants contributes to survival for nontransplants.

In addition to $HT(t)$, two time-dependent covariates included in model.

Because of the problems just described, which concern the wait-time of transplants and the effects of prognostic factors attributable to transplants only, an alternative approach to the analysis is recommended. This alternative involves the use of time-dependent variables in an extended Cox model.

The exposure variable of interest in this extended Cox model is heart transplant status at time t, denoted by $HT(t)$. This variable is defined to take on the value 0 at time t if the patient has not received a transplant at this time, that is, if t is less than the wait-time for receiving a transplant. The value of this variable is 1 at time t if the patient has received a transplant prior to or at time t, that is, if t is equal to or greater than the wait-time.

Thus, for a patient who did not receive a transplant during the study, the value of $HT(t)$ is 0 at all times. For a patient receiving a transplant, the value of $HT(t)$ is 0 at the start of eligibility and continues to be 0 until the time at which the patient receives the transplant; then, the value of $HT(t)$ changes to 1 and remains 1 throughout the remainder of follow-up.

Note that the variable $HT(t)$ has the property that the wait-time for transplant patients contributes to the survival experience of nontransplant patients. In other words, this variable treats a transplant patient as a nontransplant patient prior to receiving the transplant.

In addition to the exposure variable $HT(t)$, two other time-dependent variables are included in our extended Cox model for the transplant data. These variables are covariates to be adjusted for in the assessment of the effect of the $HT(t)$ variable.

EXAMPLE: (continued)

Covariates:

$$TMS(t) = \begin{cases} 0 \text{ if } t < \text{wait-time} \\ TMS \text{ if } t \geq \text{wait-time} \end{cases}$$

$$AGE(t) = \begin{cases} 0 \text{ if } t < \text{wait-time} \\ AGE \text{ if } t \geq \text{wait-time} \end{cases}$$

These covariates are denoted as $TMS(t)$ and $AGE(t)$ and they are defined as follows: $TMS(t)$ equals 0 if t is less than the wait-time for a transplant but changes to the "tissue mismatch score" (TMS) at the time of the transplant if t is equal to or greater than the wait-time. Similarly, $AGE(t)$ equals 0 if t is less than the wait-time but changes to AGE at time of transplant if t is equal to or greater than the wait-time.

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta_1 HT(t) + \delta_2 TMS(t) + \delta_3 AGE(t)]$$

The extended Cox model for the transplant data is shown here. The model contains the three time-dependent variables $HT(t)$, $TMS(t)$ and $AGE(t)$ as described above.

Focus:
Asscesing the effect of $HT(t)$ adjusted for $TMS(t)$ and $AGE(t)$

Note: $HT(t)$ does not satisfy PH assumption.

For this model, since $HT(t)$ is the exposure variable of interest, the focus of the analysis concerns assessing the effect of this variable adjusted for the two covariates. Note, however, that because the $HT(t)$ variable is time-dependent by definition, this variable does not satisfy the PH assumption, so that any hazard ratio estimate obtained for this variable is technically time-dependent.

Variable	Coef.	Std. Err.	P>|z|	Haz. Ratio
$HT(t)$	−3.1718	1.1861	0.008	0.0417
$TMS(t)$	0.4442	0.2802	0.112	1.5593
$AGE(t)$	0.0552	0.0226	0.014	1.0567

A summary of computer results for the fit of the above extended Cox model is shown here. These results indicate that the exposure variable $HT(t)$ is significant below the one percent significance level (i.e., the two-sided p-value is 0.008). Thus, transplant status appears to be significantly associated with survival.

$$\widehat{\text{HR}} = e^{-3.1718} = 0.0417 = \frac{1}{23.98}$$

$$\widehat{\text{HR}} = \frac{\hat{h}(\text{transplants})}{\hat{h}(\text{nontransplants})} \approx \frac{1}{24}?$$

Not appropriate!

To evaluate the strength of the association, note that e to the coefficient of $HT(t)$ equals 0.0417. Since 1 over 0.0417 is 23.98, it appears that there is a 24-fold increase in the hazard of nontransplant patients to transplant patients. The preceding interpretation of the value 0.0417 as a hazard ratio estimate is not appropriate, however, as we shall now discuss further.

EXAMPLE (continued)

23.98 is inappropriate as a $\widehat{HR}$:

- does not compare two *separate* groups
- exposure variable is *not* time-independent
- wait-time on transplants contributes to survival on nontransplants

First, note that the value of 23.98 inappropriately suggests that the hazard ratio is comparing two separate groups of patients. However, the exposure variable in this analysis is *not* a time-independent variable that distinguishes between two separate groups. In contrast, the exposure variable is time-dependent, and uses the wait-time information on transplants as contributing to the survival experience of nontransplants.

Alternative interpretation:
At time t,
h("not yet received transplant")
$\approx 24\,h\hat{}$ ("already received transplant")

Since the exposure variable is time-dependent, an alternative interpretation of the hazard ratio estimate is that, at any given time t, the hazard for a person *who has not yet received a transplant* (but may receive one later) is approximately 24 times the hazard for a person *who already has received a transplant by that time*.

More appropriate:

Hazard ratio formula should account for *TMS* and *AGE*.

Transplant?	$HT(t)$	$TMS(t)$	$AGE(t)$
Yes	1	*TMS*	*AGE*
No	0	0	0

Actually, we suggest that a more appropriate hazard ratio expression is required to account for a transplant's *TMS* and *AGE* score. Such an expression would compare, at time t, the values of each of the three time-dependent variables in the model. For a person who received a transplant, these values are 1 for $HT(t)$ and *TMS* and *AGE* for the two covariates. For a person who has not received a transplant, the values of all three variables are 0.

i denotes ith transplant patient

$\mathbf{X}^*(t) = (HT(t) = 1, TMS(t) = TMS_i, \text{AGE}(t) = AGE_i)$
$\mathbf{X}(t) = (HT(t) = 0, TMS(t) = 0, AGE(t) = 0)$

Using this approach to compute the hazard ratio, the $\mathbf{X}^*(t)$ vector, which specifies the predictors for a patient i who received a transplant at time t, has the values 1, TMS_i and AGE_i for patient i; the $\mathbf{X}(t)$ vector, which specifies the predictors at time t for a patient who has not received a transplant at time t, has values of 0 for all three predictors.

$$\begin{aligned}\widehat{HR}(t) &= \exp\left[\hat{\delta}_1(1-0) + \hat{\delta}_2(TMS_i - 0) + \hat{\delta}_3(AGE_i - 0)\right] \\ &= \exp\left[\hat{\delta}_1 + \hat{\delta}_2\, TMS_i + \hat{\delta}_3\, AGE_i\right] \\ &= \exp[-3.1718 + 0.4442\, TMS_i + 0.0552\, AGE_i]\end{aligned}$$

The hazard ratio formula then reduces to e to the sum of δ_1 "hat" plus δ_2 "hat" times TMS_i plus δ_3 "hat" times AGE_i, where the δ "hat's" are the estimated coefficients of the three time-dependent variables. Substituting the numerical values for these coefficients in the formula gives the exponential expression circled here.

EXAMPLE: (continued)

$\widehat{HR}(t)$ is time-dependent, i.e., its value at time t depends on TMS_i and AGE_i at time t

TMS range: (0–3.05)
AGE range: (12–64)

The resulting formula for the hazard ratio is time-dependent in that its value depends on the *TMS* and *AGE* values of the *i*th patient at the time of transplant. That is, different patients can have different values for *TMS* and *AGE* at time of transplant. Note that in the dataset, *TMS* ranged between 0 and 3.05 and *AGE* ranged between 12 and 64.

We end our discussion of the Stanford Heart Transplant Study at this point. For further insight into the analysis of this dataset, we refer the reader to the 1977 paper by Crowley and Hu (*J. Amer. Statist. Assoc.*).

IX. The Extended Cox Likelihood

- Cox PH likelilhood (L) described in Chapter 3, Section VIII
- L now extended for extended Cox model

At the end of the presentation from Chapter 3 (Section VIII), we illustrated the Cox likelihood using the dataset shown on the left. In this section, we extend that discussion to illustrate the Cox likelihood with a time-dependent variable.

ID	TIME	STATUS	SMOKE
Barry	2	1	1
Gary	3	1	0
Harry	5	0	0
Larry	8	1	1

TIME = Survival time (in years)
STATUS = 1 for event, 0 for censorship
SMOKE = 1 for a smoker, 0 for a nonsmoker

To review: The data indicate that Barry got the event at TIME = 2 years. Gary got the event at 3 years, Harry was censored at 5 years and Larry got the event at 8 years. Furthermore, Barry and Larry were smokers while Gary and Harry were non-smokers.

Cox PH model: $h(t)=h_0(t)e^{\beta_1 SMOKE}$

In Chapter 3, we constructed the Cox likelihood with one predictor, SMOKE, in the model. The model and the likelihood are shown on the left.

Cox PH Likelihood:

$$L=\left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1}+h_0(t)e^{0}+h_0(t)e^{0}+h_0(t)e^{\beta_1}}\right]\times\left[\frac{h_0(t)e^{0}}{h_0(t)e^{0}+h_0(t)e^{0}+h_0(t)e^{\beta_1}}\right]\times\left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1}}\right]$$

The likelihood is a product of three terms, one term for each event time t_f (TIME = 2, 3, and 8). The denominator of each term is the sum of the hazards from the subjects still in the risk set at time t_f, including the censored subject, Harry. The numerator of each term is the hazard of the subject who got the event at t_f. The reader may wish to reread Section VIII of Chapter 3.

Cox extended model:

$$h(t) = h_0(t)e^{\beta_1 SMOKE+\beta_2 SMOKE\times TIME}$$

Time-dependent covariate
(its value changes over time)

Now consider an extended Cox model, which contains the predictor SMOKE and a time-dependent variable SMOKE × TIME. For this model, it is not only the baseline hazard that may change over time but also the value of the predictor variables. This can be illustrated by examining Larry's hazard at each event time.

Larry got the event at TIME = 8

Larry's hazard at each event time

TIME	Larry's Hazard
2	$h_0(t)e^{\beta_1+2\beta_2}$
3	$h_0(t)e^{\beta_1+3\beta_2}$
8	$h_0(t)e^{\beta_1+8\beta_2}$

Larry, a smoker, got the event at TIME = 8. However at TIME = 2, 3, and 8, the covariate SMOKE × TIME changes values, thus impacting Larry's hazard at each event time (see left). Understanding how the expression for an individual's hazard changes over time is the key addition toward understanding how the Cox extended likelihood differs from the Cox PH likelihood.

Cox extended model:

$$h(t) = h_0(t)e^{\beta_1 SMOKE + \beta_2 SMOKE \times TIME}$$

$$L = \left[\frac{h_0(t)e^{\beta_1+2\beta_2}}{h_0(t)e^{\beta_1+2\beta_2} + h_0(t)e^0 + h_0(t)e^0 + h_0(t)e^{\beta_1+2\beta_2}}\right] \times \left[\frac{h_0(t)e^0}{h_0(t)e^0 + h_0(t)e^0 + h_0(t)e^{\beta_1+3\beta_2}}\right] \times \left[\frac{h_0(t)e^{\beta_1+8\beta_2}}{h_0(t)e^{\beta_1+8\beta_2}}\right]$$

The extended Cox model for these data is shown again on the left.

The extend cox likehood (L) for these data is shown on the left. This likehood is constructed in a similar mannar as the likehood for the Cox PH model. Yhe difference is that the expression for the subject's hazard is allowed to vary over time.

Likelihood is product of three terms:

$$L = L_1 \times L_2 \times L_3$$

Barry ($t = 2$) Gary ($t = 3$) Larry ($t = 8$)

Just as with the Cox PH likelihood shown on the previous page, the extended Cox likelihood is also a product of three terms, corresponding to the three event times ($L = L_1 \times L_2 \times L_3$). Barry got the event first at t = 2, then Gary at t = 3, and finally Larry at t = 8. Harry, who was censored at t = 5, was still at risk when Barry and Garry got the event. Therefore, Harry's hazard is still in the denominator of L_1 and L_2.

SMOKE × TIME = 0 for nonsmokers

SMOKE × TIME changes over time for smokers

Larry's hazard changes over L_1, L_2, L_3.

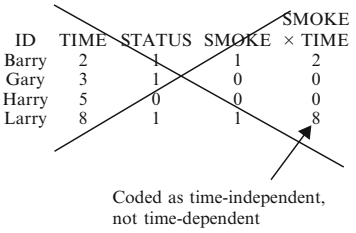
The inclusion of the time-varying covariate SMOKE × TIME does not change the expression for the hazard for the nonsmokers (Gary and Harry) since SMOKE is coded 0 for nonsmokers. However, for smokers (Barry and Larry), the expression for the hazard changes with time. Notice how Larry's hazard changes in the denominator of L_1, L_2, and L_3.

$h_0(t)$ cancels in L

$$L = \left[\frac{e^{\beta_1+2\beta_2}}{e^{\beta_1+2\beta_2} + e^0 + e^0 + e^{\beta_1+2\beta_2}}\right] \times \left[\frac{e^0}{e^0 + e^0 + e^{\beta_1+3\beta_2}}\right] \times \left[\frac{e^{\beta_1+8\beta_2}}{e^{\beta_1+8\beta_2}}\right]$$

The baseline hazard cancels in the extended Cox likelihood as it does with the Cox PH likelihood. Thus, the form of the baseline hazard need not be specified, as it plays no role in the estimation of the regression parameters.

Caution: Incorrect Coding of SMOKE × TIME



A word of caution for those planning to run a model with a time-varying covariate: It is incorrect to create a product term with TIME in the data step by multiplying each individual's value for SMOKE with their survival time. In other words, SMOKE×TIME should not be coded like the typical interaction term. In fact, if SMOKE×TIME was coded as it is on the left, then SMOKE×TIME would be a time-independent variable. Larry's value for SMOKE×TIME is incorrectly coded at a constant value of 8 even though Larry's value for SMOKE×TIME changes in the likelihood over L_1, L_2, and L_3.

Incorrectly coded SMOKE×TIME
- Time independent
- Probably highly significant
- Survival time should predict survival time
- But not meaningful

If the incorrectly coded time independent SMOKE×TIME was included in a Cox model, it would not be surprising if the coefficient estimate was highly significant even if the PH assumption was not violated. It would be expected that a product term with each individual's survival time would predict the outcome (their survival time), but it would not be meaningful. Nevertheless, this is a common mistake.

Correctly coding SMOKE×TIME
- Time dependent
- Computer packages typically allow definition in the analytic procedure

To obtain a correctly defined SMOKE×TIME time-dependent variable, computer packages typically allow the variable to be defined within the analytic procedure.

Alternatively can code using **CP format**

Alternatively, **the (start, stop) or counting process (CP) data layout** can be used, to explicitly define a time-dependent variable in the data.

- CP format introduced in Chapter 1
- Multiple observations per subject
- Time intervals at risk subdivided
- Covariate value can change from interval to interval for the same subject

The CP data format was introduced in Chapter 1. This data layout provides a straightforward approach for expressing a time-dependent covariate by allowing multiple observations to correspond to the same individual. With this format, an individual's total at-risk follow-up time is sub-divided into smaller time intervals providing a way for values of variables to change from time interval-to-interval for the same individual.

CP format with time-dependent variable SMOKE × TIME:

ID	START	STOP	STATUS	SMOKE	SMOKE × TIME
Barry	0	2	1	1	2
Gary	0	2	0	0	0
Gary	2	3	1	0	0
Harry	0	2	0	0	0
Harry	2	3	0	0	0
Harry	3	5	0	0	0
Larry	0	2	0	1	2
Larry	2	3	0	1	3
Larry	3	8	1	1	8

Coded as time dependent

START = Beginning of interval (in months)
STOP = End of interval (in months)
STATUS = 1 for event, 0 for censorship

The data layout on the left illustrates the CP approach with the (start, stop) time intervals defined by each event time (t=2, t=3, and t=8) from Barry, Gary, and Larry.

If we look at the final three observations, we can see that Larry's total time at risk is subdivided into three time intervals. Larry got the event at t=8 (STOP=8 and STATUS=1). The two previous observations indicate that Larry did not get an event (STATUS=0) over the time intervals (0, 2) or (2, 3). Over the last three observations, the time-dependent variable SMOKE×TIME changes values for Larry from 2 to 3 to 8.

Alternative CP format:
Gary and Harry do not need multiple observations since SMOKE × TIME does not vary for them (same info as above)

ID	START	STOP	STATUS	SMOKE	SMOKE × TIME
Barry	0	2	1	1	2
Gary	0	3	1	0	0
Harry	0	5	0	0	0
Larry	0	2	0	1	2
Larry	2	3	0	1	3
Larry	3	8	1	1	8

An alternative CP data layout is shown below on the left. Since Gary and Harry are nonsmokers and the coding for nonsmokers is SMOKE=0, their values for the SMOKE×TIME variable stays at zero throughout their time at risk. Therefore, it is not necessary to have multiple observations for Gary and Harry (although it is not incorrect to do so).

2 reasons to include time-varying covariate:
1) To account for PH violation
2) The values actually change over time regardless of the PH assumption

There are two main reasons why a time-varying covariate might be included in a Cox model: (1) To account for a violation of the proportional hazards assumption (usually formulated as a product term with some function of time) and (2) The covariate may actually change its values over time regardless of the PH assumption.

SMOKE×TIME
defined to account for PH violation

The use of the SMOKE×TIME variable in the last example was of the first type (to account for a PH violation). An example of the second type could be the changing of an individual's dosage level of medication over time as illustrated in the next example.

DOSE changes at 3 time points for Jane

ID	MONTHS	STATUS	DOSE1	TIME1	DOSE2	TIME2	DOSE3	TIME3
Jane	49	1	60	0	120	12	150	30

The data on the left contains one observation for Jane who had an event at 49 months (MONTHS=49 and STATUS=1). Her dose of medication at the beginning of follow-up was 60 milligrams (DOSE1=60 and TIME1=0). At the 12th month of follow-up, her dose was changed to 120 milligrams (DOSE2=120 and TIME2=12). At the 30th month of follow-up, her dose was changed to 150 mg (DOSE3=120 and TIME3=30).

MONTHS = Survival time (in months)
STATUS = 1 for event, 0 for censorship

Same info as above using CP format

(3 observations instead of 1)

ID	START	STOP	STATUS	DOSE
Jane	0	12	0	60
Jane	12	30	0	120
Jane	30	49	1	150

The same information can be expressed using the counting process data layout. On the left, the data is transposed to contain three observations for Jane allowing DOSE to be represented as a time-dependent variable. For the first time interval (START=0, STOP=12), Jane's dose was 60 mg. For the second time interval (12–30 months), Jane's dose was 120 milligrams. For the third time interval (30–49 months), Jane's dose was 150 milligrams. The data indicates that Jane had an event at 49 months (STOP=49 and STATUS=1).

START = Beginning of interval (in months)
STOP = End of interval (in months)
STATUS = 1 for event, 0 for censorship
DOSE = Dose in milligrams

Multiple observations per subject: revisited in Chapter 8 (recurrent events)

The counting process data layout is further discussed in Chapter 8 on recurrent events. With recurrent event data, subjects may remain at risk for subsequent events after getting an event.

Coding SMOKE × TIME as time-dependent

Multiple Observations per Subject

ID	TIME	STATUS	SMOKE	SMOKE × TIME
Barry	2	1	1	2
Gary	2	0	0	0
Gary	3	1	0	0
Harry	2	0	0	0
Harry	3	0	0	0
Harry	5	0	0	0
Larry	2	0	1	2
Larry	3	0	1	3
Larry	5	0	1	5
Larry	8	1	1	8

↑
Coded as time-dependent

When a time-dependent variable is defined within the Cox analytic procedure, the variable is defined internally such that the user may not see the time-dependent variable in the dataset. However, the dataset on the left will provide a clearer idea of the correct definition of SMOKE × TIME. The dataset contains multiple observations per subject. Barry was at risk at t = 2 and got the event at that time. Gary was at risk at t = 2 and t = 3. Gary didn't get the event at t = 2 but did get the event at t = 3. Harry was at risk at t = 2, t = 3, t = 5 and didn't get the event. Larry was at risk at t = 2, t = 3, t = 5, t = 8 and got the event at t = 8. Notice how the SMOKE × TIME variable changes values for Larry over time.

Multiple observations per subject: revisited in Chapter 8 (recurrent events)

Survival analysis datasets containing multiple observations per subject are further discussed in Chapter 8 on recurrent events. With recurrent event data, subjects may remain at risk for subsequent events after getting an event.

X. Summary

Review Cox PH model.

Define time-dependent variable: defined, internal, ancillary.

A summary of this presentation on time-dependent variables is now provided. We began by reviewing the main features of the Cox PH model. We then defined a time-dependent variable and illustrated three types of these variables–defined, internal, and ancillary.

Extended Cox model:

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p_1} \beta_i \mathbf{X}_i + \sum_{j=1}^{p_2} \delta_j X_j(t)\right]$$

Next, we gave the form of the "extended Cox model," shown here again, which allows for time-dependent as well as time-independent variables.

$$\widehat{HR}(t) = \exp\left[\sum_{i=1}^{p_1} \hat{\beta}_i \left[X_i^* - X_i\right] + \sum_{j=1}^{p_2} \hat{\delta}_j \left[X_j^*(t) - X_j(t)\right]\right]$$

↑
Function of time

We then described various characteristics of this extended Cox model, including the formula for the hazard ratio. The latter formula is time-dependent so that the PH assumption is not satisfied.

Model for assessing PH assumption:

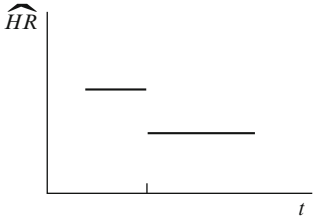
$$h(t,\mathbf{X}(t)) = h_0(t)\exp\left[\sum_{i=1}^{p}\beta_i\mathbf{X}_i + \sum_{i=1}^{p}\delta_i X_i g_i(t)\right]$$

We also showed how to use time-dependent variables to assess the PH assumption for time-independent variables. A general formula for an extended Cox model that simultaneously considers all time-independent variables of interest is shown here.

Examples of $g_i(t)$:
t, $\log t$, heaviside function

The functions $g_i(t)$ denote functions of time for the ith variable that are to be determined by the investigator. Examples of such functions are $g_i(t) = t$, $\log t$, or a heaviside function.

Heaviside functions:



The use of heaviside functions were described and illustrated. Such functions allow for the hazard ratio to be constant within different time intervals.

$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta E + \delta E g(t)]$

where

$$g(t) = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t < t_0 \end{cases}$$

$h(t,\mathbf{X}(t))$
$= h_0(t)\exp[\beta_1 E g_1(t) + \beta_2 E g_2(t)]$
where

$$g_1(t) = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } < t_0 \end{cases}$$

$$g_2(t) = \begin{cases} 1 & \text{if } t < t_0 \\ 0 & \text{if } t \geq t_0 \end{cases}$$

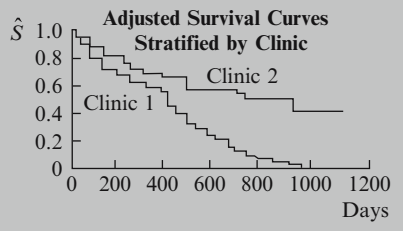
For two time intervals, the model can take either one of two equivalent forms as shown here. The first model contains a main effect of exposure and only one heaviside function. The second model contains two heaviside functions without a main effect of exposure. Both models yield two distinct and equivalent values for the hazard ratio.

EXAMPLE 1

1991 Australian study of heroin addicts
- two methadone maintenance clinics
- *addicts* dataset file
- clinic variable did not satisfy PH assumption

We illustrated the use of time-dependent variables through two examples. The first example considered the comparison of two methadone maintenance clinics for heroin addicts. The dataset file was called *addicts*. In this example, the clinic variable, which was a dichotomous exposure variable, did not satisfy the PH assumption.

EXAMPLE: (continued)



Two distinct HR's:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_2(\text{prison}) + \beta_3(\text{dose}) + \delta_1(\text{clinic})g_1(t) + \delta_2(\text{clinic})g_2(t)]$$

Heaviside functions

Diverging HR's:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_2(\text{prison}) + \beta_3(\text{dose}) + \beta_1(\text{clinic}) + \delta(\text{clinic} \times t)]$$

Adjusted survival curves stratified by clinic showed clinic 2 to have consistently higher survival probabilities than clinic 1, with a more pronounced difference in clinics after one year of follow-up. However, this stratification did not allow us to obtain a hazard ratio estimate for clinic. Such an estimate was possible using an extended Cox model containing interaction terms involving clinic with time.

Two extended Cox models were considered. The first used heaviside functions to obtain two distinct hazard ratios, one for the first year of follow-up and the other for greater than one year of follow-up. The model is shown here.

The second extended Cox model used a time-dependent variable that allowed for the two survival curves to diverge over time. This model is shown here.

Both models yielded hazard ratio estimates that agreed reasonably well with the graph of adjusted survival curves stratified by clinic.

EXAMPLE 2: Stanford Heart Transplant Study

Goals: Do patients receiving transplants survive longer than patients not receiving transplants?

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta_1 HT(t) + \delta_2 TMS(t) + \delta_3 AGE(t)]$$

Exposure variable

The second example considered results obtained in the Stanford Heart Transplant Study. The goal of the study was to assess whether patients receiving transplants survived longer than patients not receiving transplants.

The analysis of these data involved an extended Cox model containing three time-dependent variables. One of these, the exposure variable, and called $HT(t)$, was an indicator of transplant status at time t. The other two variables, $TMS(t)$ and $AGE(t)$, gave tissue mismatch scores and age for transplant patients when time t occurred after receiving a transplant. The value of each of these variables was 0 at times prior to receiving a transplant.

EXAMPLE: (continued)

Results: $HT(t)$ highly significant, i.e., transplants have better prognosis than nontransplants.
Hazard ratio estimate problematic:

$$\widehat{HR} = e^{\hat{\delta}_1} = \frac{1}{23.98}$$

More appropriate formula:

$$\widehat{HR} = \exp[-3.1718 + 0.4442\ TMS_i + 0.0552\ AGE_i]$$

The results from fitting the above extended Cox model yielded a highly significant effect of the exposure variable, thus indicating that survival prognosis was better for transplants than for nontransplants.

From these data, we first presented an inappropriate formula for the estimated hazard ratio. This formula used the exponential of the coefficient of the exposure variable, which gave an estimate of 1 over 23.98. A more appropriate formula considered the values of the covariates $TMS(t)$ and $AGE(t)$ at time t. Using the latter, the hazard ratio estimate varied with the tissue mismatch scores and age of each transplant patient.

Chapters

1. Introduction to Survival Analysis
2. Kaplan–Meier Curves and the Log–Rank Test
3. The Cox Proportional Hazards Model
4. Evaluating the Proportional Hazards Assumption
5. The Stratified Cox Procedure
6. ✓ Extension of the Cox Proportional Hazards Model for Time-Dependent Variables

Next:

7. Parametric models

This presentation is now complete. We suggest that the reader review the detailed outline that follows and then answer the practice exercises and test that follow the outline.

A key property of Cox models is that the distribution of the outcome, survival time, is unspecified. In the next chapter, parametric models are presented in which the underlying distribution of the outcome is specified. The exponential, Weibull, and log-logistic models are examples of parametric models.

Detailed Outline

I. **Preview** (page 244)

II. **Review of the Cox PH Model** (pages 244–246)

A. The formula for the Cox PH model:

$$\boxed{h(t,\mathbf{X}) = h_0(t)\exp\left[\sum_{i=1}^{p}\beta_i X_i\right]}$$

B. Formula for hazard ratio comparing two individuals:

$$\mathbf{X}^* = \left(X_1^*, X_2^*, \ldots, X_p^*\right) \text{ and } \mathbf{X} = \left(X_1, X_2, \ldots, X_p\right):$$

$$\boxed{\frac{h(t,\mathbf{X}^*)}{h(t,\mathbf{X})} = \exp\left[\sum_{i=1}^{p}\beta_i\left(X_i^* - X_i\right)\right]}$$

C. The meaning of the PH assumption:

- Hazard ratio formula shows that the hazard ratio is independent of time:

$$\frac{h(t,\mathbf{X}^*)}{h(t,\mathbf{X})} = \theta$$

- Hazard ratio for two X's are proportional:

$$h(t,\mathbf{X}^*) = \theta h(t,\mathbf{X})$$

D. Three methods for checking the PH assumption:

i. *Graphical:* Compare ln–ln survival curves or observed versus predicted curves

ii. *Time-dependent covariates:* Use product (i.e., interaction) terms of the form $X \times g(t)$.

iii. *Goodness-of-fit test:* Use a large sample Z statistic.

E. Options when the PH assumption is not met:

i. Use a stratified Cox procedure.

ii. Use an extended Cox model containing a time-dependent variable of the form $X \times g(t)$.

III. **Definition and Examples of Time-Dependent Variables** (pages 246–249)

A. Definition: any variable whose values differ over time

B. Examples of defined, internal, and ancillary time-dependent variables

IV. The Extended Cox Model for Time-Dependent Variables (pages 249–251)

A. $$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p_2} \delta_j X_j(t)\right]$$

where $\mathbf{X}(t) = (X_1, X_2, \ldots, X_{p_1}, X_1(t), X_2(t), \ldots, X_{p_2}(t))$ denotes the entire collection of predictors at time t, X_i denotes the ith time-independent variable, and $X_j(t)$ denotes the jth time-dependent variable.

B. ML procedure used to estimate regression coefficients.

C. List of computer programs for the extended Cox model.

D. Model assumes that the hazard at time t depends on the value of $X_j(t)$ at the *same* time.

E. Can modify model for lag-time effect.

V. The Hazard Ratio Formula for the Extended Cox Model (pages 251–253)

A. $$HR(t) = \exp\left[\sum_{i=1}^{p_1} \beta_i \left[X_i^* - X_i\right] + \sum_{j=1}^{p_2} \delta_j \left[X_j^*(t) - X_j(t)\right]\right]$$

B. Because $HR(t)$ is a function of time, the PH assumption is not satisfied.

C. The estimated coefficient of $X_j(t)$ is time-independent, and represents an "overall" effect of $X_j(t)$.

VI. Assessing Time-Independent Variables That Do Not Satisfy the PH Assumption (pages 254–259)

A. General formula for assessing PH assumption:

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i + \sum_{i=1}^{p} \delta_i X_i g_i(t)\right]$$

B. $g_i(t)$ is a function of time corresponding to X_i

C. Test H_0: $\delta_1 = \delta_2 = \ldots = \delta_p = 0$

D. Heaviside function:

$$g(t) = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t < t_0 \end{cases}$$

E. The model with a single heaviside function:

$$\boxed{h(t, \mathbf{X}(t)) = h_0(t) \exp[\beta E + \delta E g(t)]}$$

F. The model with two heaviside functions:

$$\boxed{h(t, \mathbf{X}(t)) = h_0(t) \exp[\delta_1 E g_1(t) + \delta_2 E g_2(t)]}$$

where

$$g_1(t) = \begin{cases} 1 \text{ if } t \geq t_0 \\ 0 \text{ if } t < t_0 \end{cases} \text{ and } g_2(t) = \begin{cases} 1 \text{ if } t < t_0 \\ 0 \text{ if } t \geq t_0 \end{cases}$$

G. The hazard ratios:

$$t \geq t_0 : \widehat{HR} = \exp\left[\hat{\beta} + \hat{\delta}\right] = \exp\left[\hat{\delta}_1\right]$$
$$t < t_0 : \widehat{HR} = \exp\left[\hat{\beta}\right] = \exp\left[\hat{\delta}_2\right]$$

H. Several heaviside functions: examples given with four time-intervals:

- Extended Cox model contains either $\{E, E \times g_1(t), E \times g_2(t), E \times g_3(t)\}$ or $\{E \times g_1(t), E \times g_2(t), E \times g_3(t), E \times g_4(t)\}$
- The model using four product terms and no main effect of E:

$$\boxed{\begin{aligned} h(t, \mathbf{X}(t)) = h_0(t) \exp[&\delta_1 E g_1(t) + \delta_2 E g_2(t) \\ &+ \delta_3 E g_3(t) + \delta_4 E g_4(t)] \end{aligned}}$$

where

$$g_i(t) = \begin{cases} 1 & \text{if } t \text{ is within interval } i \\ 0 & \text{if otherwise} \end{cases}$$

VII. An Application of the Extended Cox Model to An Epidemiologic Study on the Treatment of Heroin Addiction (pages 260–264)

A. 1991 Australian study of heroin addicts

- two methadone maintenance clinics
- *addicts* dataset file
- clinic variable did not satisfy PH assumption

B. Clinic 2 has consistently higher retention probabilities than clinic 1, with a more pronounced difference in clinics after 1 year of treatment.

C. Two extended Cox models were considered:

- Use heaviside functions to obtain two distinct hazard ratios, one for less than 1 year and the other for greater than 1 year.
- Use a time-dependent variable that allows for the two survival curves to diverge over time.

VIII. An Application of the Extended Cox Model to the Analysis of the Stanford Heart Transplant Data (pages 265–269)

A. The goal of the study was to assess whether patients receiving transplants survived longer than patients not receiving transplants.

B. We described an extended Cox model containing three time-dependent variables:

$$h(t,\mathbf{X}(t)) = h_0(t)\exp[\delta_1 HT(t) + \delta_2 TMS(t) + \delta_3 AGE(t)]$$

C. The exposure variable, called $HT(t)$, was an indicator of transplant status at time t. The other two variables, $TMS(t)$ and AGE(t), gave tissue mismatch scores and age for transplant patients when time t occurred after receiving a transplant.

D. The results yielded a highly significant effect of the exposure variable.

E. The use of a hazard ratio estimate for this data was problematical.

- An inappropriate formula is the exponential of the coefficient of $HT(t)$, which yields 1/23.98.
- An alternative formula considers the values of the covariates $TMS(t)$ and $AGE(t)$ at time t.

IX. The Extended Cox Likelihood (pages 269–274)

A. Review of PH likelihood (Chapter 3).

B. Barry, Gary, Larry, example of Cox likelihood.

X. Summary (pages 274–277)

Practice Exercises

The following dataset called "anderson.dat" consists of remission survival times on 42 leukemia patients, half of whom receive a new therapy and the other half of whom get a standard therapy (Freireich et al., *Blood*, 1963). The exposure variable of interest is treatment status (Rx = 0 if new treatment, Rx = 1 if standard treatment). Two other variables for control are log white blood cell count (i.e., log WBC) and sex. Failure status is defined by the relapse variable (0 if censored, 1 if failure). The dataset is listed as follows:

Subj	Surv	Relapse	Sex	log WBC	*Rx*
1	35	0	1	1.45	0
2	34	0	1	1.47	0
3	32	0	1	2.2	0
4	32	0	1	2.53	0
5	25	0	1	1.78	0
6	23	1	1	2.57	0

(Continued on next page)

(Continued)

Subj	Surv	Relapse	Sex	log WBC	*Rx*
7	22	1	1	2.32	0
8	20	0	1	2.01	0
9	19	0	0	2.05	0
10	17	0	0	2.16	0
11	16	1	1	3.6	0
12	13	1	0	2.88	0
13	11	0	0	2.6	0
14	10	0	0	2.7	0
15	10	1	0	2.96	0
16	9	0	0	2.8	0
17	7	1	0	4.43	0
18	6	0	0	3.2	0
19	6	1	0	2.31	0
20	6	1	1	4.06	0
21	6	1	0	3.28	0
22	23	1	1	1.97	1
23	22	1	0	2.73	1
24	17	1	0	2.95	1
25	15	1	0	2.3	1
26	12	1	0	1.5	1
27	12	1	0	3.06	1
28	11	1	0	3.49	1
29	11	1	0	2.12	1
30	8	1	0	3.52	1
31	8	1	0	3.05	1
32	8	1	0	2.32	1
33	8	1	1	3.26	1
34	5	1	1	3.49	1
35	5	1	0	3.97	1
36	4	1	1	4.36	1
37	4	1	1	2.42	1
38	3	1	1	4.01	1
39	2	1	1	4.91	1
40	2	1	1	4.48	1
41	1	1	1	2.8	1
42	1	1	1	5	1

The following edited printout gives computer results for fitting a Cox PH model containing the three predictives *Rx*, log WBC, and Sex.

Cox regression Analysis time_t: survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf. Interval]		*P*(*PH*)
Sex	0.263	0.449	0.558	1.301	0.539	3.139	0.042
log WBC	1.594	0.330	0.000	4.922	2.578	9.397	0.714
Rx	1.391	0.457	0.002	4.018	1.642	9.834	0.500

No. of subjects = 42 Log likelihood = −72.109

1. Which of the variables in the model fitted above are time-independent and which are time-dependent?
2. Based on this printout, is the PH assumption satisfied for the model being fit? Explain briefly.
3. Suppose you want to use an extended Cox model to assess the PH assumption for all three variables in the above model. State the general form of an extended Cox model that will allow for this assessment.
4. Suppose you wish to assess the PH assumption for the Sex variable using a heaviside function approach designed to yield a constant hazard ratio for less than 15 weeks of follow-up and a constant hazard ratio for 15 weeks or more of follow-up. State two equivalent alternative extended Cox models that will carry out this approach, one model containing one heaviside function and the other model containing two heaviside functions.
5. The following is an edited printout of the results obtained by fitting an extended Cox model containing two heaviside functions:

   Time-Dependent Cox Regression Analysis

Analysis time_t: survt	Coef.	Std. Err.	p > |z|	Haz. Ratio		[95% Conf. Interval]
log WBC	1.567	0.333	0.000	4.794	2.498	9.202
Rx	1.341	0.466	0.004	3.822	1.533	9.526
0–15 wks	0.358	0.483	0.459	1.430	0.555	3.682
15+ wks	−0.182	0.992	0.855	0.834	0.119	5.831

   No. of subjects = 42 Log likelihood = −71.980

   Using the above computer results, carry out a test of hypothesis, estimate the hazard ratio, and obtain 95% confidence interval for the treatment effect adjusted for log WBC and the time-dependent Sex variables. What conclusions do you draw about the treatment effect?
6. We now consider an alternative approach to controlling for Sex using an extended Cox model. We define an interaction term between sex and time that allows for diverging survival curves over time.

   For the situation just described, write down the extended Cox model, which contains *Rx*, log WBC, and Sex as main effects plus the product term Sex × time.

7. Using the model described in question 6, express the hazard ratio for the effect of Sex adjusted for *Rx* and log WBC at 8 and 16 weeks.

8. The following is an edited printout of computer results obtained by fitting the model described in question 6.

Time-Dependent Cox Regression Analysis

Analysis time_t: survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
Sex	1.820	1.012	0.072	6.174	0.849	44.896
log WBC	1.464	0.336	0.000	4.322	2.236	8.351
Rx	1.093	0.479	0.022	2.984	1.167	7.626
Sex × Time	−0.345	0.199	0.083	0.708	0.479	1.046

No. of subjects = 42 Log likelihood = −70.416

Based on the above results, describe the hazard ratio estimate for the treatment effect adjusted for the other variables in the model, and summarize the results of the significance test and interval estimate for this hazard ratio. How do these results compare with the results previously obtained when a heaviside function approach was used? What does this comparison suggest about the drawbacks of using an extended Cox model to adjust for variables not satisfying the PH assumption?

9. The following gives an edited printout of computer results using a stratified Cox procedure that stratifies on the Sex variable but keeps *Rx* and log WBC in the model.

Stratified Cox regression

Analysis time_t: survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.390	0.338	0.000	4.016	2.072	7.783
Rx	0.931	0.472	0.048	2.537	1.006	6.396

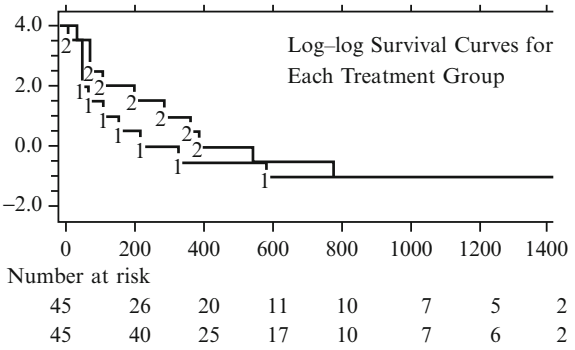
No. of subjects = 42 Log likelihood = −57.560 Stratified by sex

Compare the results of the above printout with previously provided results regarding the hazard ratio for the effect of *Rx*. Is there any way to determine which set of results is more appropriate? Explain.

Test

The following questions consider the analysis of data from a clinical trial concerning gastric carcinoma, in which 90 patients were randomized to either chemotherapy (coded as 2) alone or to a combination of chemotherapy and radiation (coded as 1). See Stablein et al., "Analysis of Survival Data with Nonproportional Hazard Functions," *Controlled Clinical Trials*, vol. 2, pp. 149–159 (1981).

1. A plot of the log–log Kaplan–Meier curves for each treatment group is shown below. Based on this plot, what would you conclude about the PH assumption regarding the treatment group variable? Explain.



2. The following is an edited printout of computer results obtained when fitting the PH model containing only the treatment group variable. Based on these results, what would you conclude about the PH assumption regarding the treatment group variable? Explain.

Cox regression Analysis time_t: survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf. Interval]		*P*(*PH*)
Tx	−0.267	0.233	0.253	0.766	0.485	1.21	0
No. of subjects 90				Log likelihood = −282.744			

3. The following printout shows the results from using a heaviside function approach with an extended Cox model to fit these data. The model used product terms of the treatment variable (*Tx*) with each of three heaviside functions. The first product term (called Time1) involves a heaviside function for the period from 0 to 250 days, the second product term (i.e., Time2) involves the period from 250 to 500 days, and the third product term (i.e., Time3) involves the open-ended period from 500 days and beyond.

Time-Dependent Cox Regression Analysis

Analysis time_t: survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Time1	−1.511	0.461	0.001	0.221	0.089	0.545
Time2	0.488	0.450	0.278	1.629	0.675	3.934
Time3	0.365	0.444	0.411	1.441	0.604	3.440

No. of subjects = 90 Log likelihood = −275.745

Write down the hazard function formula for the extended Cox model being used, making sure to explicitly define the heaviside functions involved.

4. Based on the printout, describe the hazard ratios in each of the three time intervals, evaluate each hazard ratio for significance, and draw conclusions about the extent of the treatment effect in each of the three time intervals considered.
5. Inspection of the printout provided in question 3 indicates that the treatment effect in the second and third intervals appears quite similar. Consequently, another analysis was considered that uses only two intervals, from 0 to 250 days versus 250 days and beyond. Write down the hazard function formula for the extended Cox model that considers this situation (i.e., containing two heaviside functions). Also, write down an equivalent alternative hazard function formula which contains the main effect of treatment group plus one heaviside function variable.
6. For the situation described in question 5, the computer results are provided below. Based on these results, describe the hazard ratios for the treatment effect below and above 250 days, summarize the inference results for each hazard ratio, and draw conclusions about the treatment effect within each time interval.

Time-Dependent Cox Regression Analysis

Analysis time_t: survt Column name	Coeff	StErr	p-value	*HR*	0.95	CI
Time1	−1.511	0.461	0.001	0.221	0.089	0.545
Time2	0.427	0.315	0.176	1.532	0.826	2.842

No. of subjects = 90 Log likelihood = −275.764

Answers to Practice Exercises

1. All three variables in the model are time-independent variables.
2. The computer results indicate that the Sex variables do not satisfy the PH assumption because the $P(PH)$ value is 0.042, which is significant at the 0.05 level.
3. $$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{sex}) + \beta_2(\log\ \text{WBC}) + \beta_3(Rx) + \delta_1(\text{sex})g_1(t) + \delta_2(\log\ \text{WBC})g_2(t) + \delta_3(Rx)g_3(t)]$$
   where the $g_i(t)$ are functions of time.
4. Model 1 (one heaviside function)
   $$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{sex}) + \beta_2(\log\ \text{WBC}) + \beta_3(Rx) + \delta_1(\text{sex})g_1(t)]$$
   where
   $$g_1(t) = \begin{cases} 1 \text{ if } 0 \le t < 15 \text{ weeks} \\ 0 \text{ if } t \ge 15 \text{ weeks} \end{cases}$$
   Model 2 (two heaviside functions):
   $$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_2(\log\ \text{WBC}) + \beta_3(Rx) + \delta_1(\text{sex})g_1(t) + \delta_2(\text{sex})g_2(t)]$$
   where
   $$g_1(t) = \begin{cases} 1 \text{ if } 0 \le t < 15 \text{ weeks} \\ 0 \text{ if } t \ge 15 \text{ weeks} \end{cases}$$
   and
   $$g_2(t) = \begin{cases} 0 \text{ if } t \ge 15 \text{ weeks} \\ 1 \text{ if } 0 \le t < 15 \text{ weeks} \end{cases}$$
5. The estimated hazard ratio for the effect of Rx is 3.822; this estimate is adjusted for log WBC and for the Sex variable considered as two time-dependent variables involving heaviside functions. The Wald test for significance of Rx has a p-value of 0.004, which is highly significant. The 95% confidence interval for the treatment effect ranges between 1.533 and 9.526, which is quite wide, indicating considerable unreliability of the 3.822 point estimate. Nevertheless, the results estimate a statistically significant treatment effect of around 3.8.
6. $$h(t,\mathbf{X}(t)) = h_0(t)\exp[\beta_1(\text{sex}) + \beta_2(\log\ \text{WBC}) + \beta_3(Rx) + \delta_1(\text{sex} \times \text{t})]$$

7. The hazard ratio for the effect of Sex in each time interval, controlling for *Rx* and log WBC is given as follows:

$$t = 8 \text{ weeks} \quad \widehat{HR} = \exp\left[\hat{\beta}_1 + 8\hat{\delta}_1\right]$$
$$t = 16 \text{ weeks} \quad \widehat{HR} = \exp\left[\hat{\beta}_1 + 16\hat{\delta}_1\right]$$

8. Using the model containing Sex, log WBC, *Rx*, and Sex $\times$ Time, the estimated hazard ratio for the treatment effect is given by 2.984, with a p-value of 0.022 and a 95% confidence interval ranging between 1.167 and 7.626. The point estimate of 2.984 is quite different from the point estimate of 3.822 for the heaviside function model, although the confidence intervals for both models are wide enough to include both estimates. The discrepancy between point estimates demonstrates that when a time-dependent variable approach is to be used to account for a variable not satisfying the PH assumption, different results may be obtained from different choices of time-dependent variables.

9. The stratified Cox analysis yields a hazard ratio of 2.537 with a p-value of 0.048 and a 95% CI ranging between 1.006 and 6.396. The point estimate is much closer to the 2.984 for the model containing the Sex $\times$ Time product term than to the 3.822 for the model containing two heaviside functions. One way to choose between models would be to compare goodness-of-fit test statistics for each model; another way is to compare graphs of the adjusted survival curves for each model and determine by eye which set of survival curves fits the data better.