4 Basics of Survival Analysis

Survival analysis deals with the statistical modelling of time periods between a selected moment and an expected event. This event is called the result or end point.

Data for survival analysis can also be treated as time to event occurrence, survival time, time to failure, reliability time, duration and so on. The analysis of such data is important in medicine,[1] the social sciences[2] and engineering.[3]

4.1 SYSTEM RESPONSE

The time interval between the starting point and the expected events can be treated as a *random variable*, which is the response of the system – called *survival time*.[4] It should be noted that the starting point must be well defined as there are usually several ways of calculating it.[5]

Time is a continuous variable, therefore survival time T is usually also treated as a continuous random variable. In practice, however, it is observed with a predefined period accuracy[6] and is often expressed on a discrete scale.[7]

Survival analysis is based on statistical inferences regarding the *cumulative distribution* F of survival time T. Most commonly it refers to its simple estimation based on a *single random homogeneous sample*, comparison of survival time T between two attempts[8] and modelling of cumulative distribution F, as a potential function of several *explanatory variables*.

These issues do not differ from the typical inferencing and statistical modelling, but the rationale for special treatment of survival analysis is as follows:

- data for survival analysis are often cut[9]
- standard random variable distributions[10] are often not adequate to model survival cumulative distribution F of survival time T

4.2 CENSORED AND TRUNCATED DATA

When examining an exemplary problematic situation regarding the assessment of the patient's reaction to the applied treatment method, the starting point is usually determined as the patient's inclusion in the group after admission to hospital. Then the time T to the expected event is estimated – such as for example death of the patient.[11]

In practice we have full[12] and cutoff[13] data. For patients who did not die, it is known that their survival time T must be larger than observation time t. This type of data is known as cutoff on the right side – the real value of survival time T is to the right. This data type is the most common among cutoff data and may occur for many reasons.[14]

Another type of incomplete data is the subset cutoff data or left-hand side cutoff data. For section cutoff data survival time T is not exactly known, but it is known

that it is within time interval $T \in (t_1, t_2)$. This means that the expected events did not happen before time t_1, but they occurred before time t_2. Therefore, we can only say that time T is within interval (t_1, t_2). This type of data often occurs in sociological studies carried out for a specific time interval. If for the subset cutoff data the initial value is zero $t_1 = 0$, this type of data is known as a left-hand cut. When the research has to determine the time a specific event in the life of someone, it is always left- or right-hand cut.

An important feature distinguishing the survival analysis from many other mathematical statistical methods is that it can use cutoff data, which may contain important information about the nature of the phenomenon. But this is not the rule, and in this respect one should approach it carefully. As mentioned before, when time t is defined as the closing time for research, than for $T > t$ data will be cut off and for $T > t$ they will be full. If time t is determined in the course of the research, or we decide in advance that the study will be aborted when there is an adequate number of expected events, such a cutting mechanism does not bring any significant data for survival analysis[15] and they cannot be included in the analysis, although survival time T is not, in this case, completely independent of cutoff time t.[16] The data cut mechanism which is significant for survival analysis occurs when the data are functionally connected with survival, e.g. a patient is withdrawn from tests due to effect on survival.[17]

The cutting mechanisms are also important if the patient's response to the applied treatment is negative.

Cutoff data cannot be excluded from survival analysis as they can potentially cause potentially serious deviations when inferences are made,[18] but the presence of certain types of such data means the necessity to use special analytical methods. Occasionally, you can analyze a situation where neither the cutoff data nor the full data can be recorded. Such a situation may occur when you have to make an inference regarding the time to machine failure, was stopped to be used due to expiry[19] of the certificate allowing its further operation or which was sent for a mandatory overhaul. Such data are called *cut off* and require special methods of analysis, which will not be discussed here.

4.3 THE CUMULATIVE DISTRIBUTION OF SURVIVAL TIME

Cumulative distribution F of survival time, with random variable T, should be continuous and positively defined. These conditions, for instance, are met by the cumulative distribution of *gamma function* Γ:[20]

$$\forall_{\Gamma(a)=\int_0^\infty X^{a-1}\cdot e^{-X}dX} \quad F(t) = \frac{f^a}{\Gamma(a)} \cdot t^{a-1} \cdot exp(-f \cdot t) \tag{4.1}$$

where: $F(t)$ – cumulative distribution, $L(t)$ – density of probability distribution, f – frequency, t – time.

Distributions frequently used in survival analysis are collected in Tables 4.1 and 4.2.

The most popular distribution in modelling dependencies in survival analysis is the *Weibull distribution*, for which probability density $L(t)$ can be written as:

TABLE 4.1
Generalized gamma distribution.

$$\forall_{x,a,b,c>0}\ L = \frac{a}{b^{c/a}\cdot\Gamma\left(\frac{c}{a}\right)} \cdot X^{c-1} \cdot \exp\left(-\frac{X^a}{b}\right)$$

Parameter a	Parameter C	Distribution	Density of distribution $L(X)$
1	1	Exponential	$\forall_{x,b>0}\ L = \frac{1}{b}\cdot\exp\left(-\frac{X}{b}\right)$
1	c	Gamma	$\forall_{x,b,c>0}\ L = \frac{1}{b^c\cdot\Gamma(c)} \cdot X^{c-1}\cdot\exp\left(-\frac{X}{b}\right)$
2	2	Rayleigh	$\forall_{x,b>0}\ L = \frac{2}{b}\cdot X\cdot\exp\left(-\frac{X^2}{b}\right)$
a	a	Weibull	$\forall_{x,a,b>0}\ L = \frac{a}{b^a}\cdot X^{a-1}\cdot\exp\left[-\left(\frac{X}{b}\right)^a\right]$
2	3	Maxwell	$\forall_{x,b>0}\ L = \frac{4}{\sqrt{\pi}\cdot b^{3/4}}\cdot X^2\cdot\exp\left(-\frac{X^2}{b}\right)$

$$\Gamma(a) = \int_0^\infty X^{a-1}\cdot e^{-X}dX;\ b = \frac{1}{f}$$

$$L(t) = a\cdot f^a\cdot t^a\cdot t^{a-1}\cdot\exp(-f\cdot t)^a \tag{4.2}$$

where: a – constant.

The parameter a in equation (4.2) is responsible for the shape and frequency f for the scale of the probability density for the Weibull distribution (Figure 4.1).

Average $\bar{x}$ for the Weibull distribution is:

$$\bar{x} = \frac{1}{f}\cdot\Gamma\left(1+\frac{1}{a}\right) \tag{4.3}$$

and the variance σ^2:

$$\sigma^2 = \frac{1}{f^2}\cdot\left[\Gamma\left(1+\frac{2}{a}\right) - \Gamma^2\left(1+\frac{1}{a}\right)\right] \tag{4.4}$$

For $a = 1$, the Weibull distribution density L takes the form of an exponential function with an average: $\bar{x} = \frac{1}{f}$.

TABLE 4.2
Some statistical distributions used in the survival analysis (Evans M., Hastings N., Peacock B., 1993).

Distribution	Cumulative distribution	Density	Average	Variance	Survival function	Hazard function	Cumulative hazard function
	$F(X)$	$L(X)$	$\overline{X}$	σ^2	$S(X)=1-F(X)$	$h(X)=\frac{L(X)}{S(X)}$	$H(X)=\int_0^t h(X)\cdot dt$
Exponential	$1-exp\left(-\frac{X}{b}\right)$	$\frac{1}{b}\cdot exp\left(-\frac{X}{b}\right)$	b	b^2	$exp\left(-\frac{X}{b}\right)$	$\frac{1}{b}$	$\frac{X}{b}$
Logistical	$1-\frac{1}{1+exp\left(-\frac{X-a}{b}\right)}$	$\frac{exp\left(-\frac{X-a}{b}\right)}{b\cdot\left[1-exp\left(-\frac{X-a}{b}\right)\right]^2}$	a	$\frac{\pi^2\cdot b^2}{3}$	$\frac{1}{1+exp\left(\frac{X-a}{b}\right)}$	$\frac{1}{1+exp\left(-\frac{X-a}{b}\right)}$	$log\left[1+exp\left(\frac{X-a}{b}\right)\right]$
Rayleigh	$1-exp\left(-\frac{X^2}{2\cdot b^2}\right)$	$\frac{X}{b^2}\cdot exp\left(-\frac{X^2}{2\cdot b^2}\right)$	$b\cdot\sqrt{\frac{\pi}{2}}$	$b^2\cdot\left(2-\frac{\pi}{2}\right)$	$exp\left(-\frac{X^2}{2\cdot b^2}\right)$	$\frac{X}{b^2}$	$\frac{X^2}{2\cdot b^2}$
Weibull	$1-exp\left[-\left(\frac{X}{b}\right)^a\right]$	$\frac{a\cdot X^{a-1}}{b^a}\cdot exp\left[-\left(\frac{X}{b}\right)^a\right]$	$b\cdot\Gamma\left(\frac{a+1}{a}\right)$	$b^2\cdot\left[\Gamma\left(\frac{a+1}{a}\right)-\Gamma^2\left(\frac{a+1}{a}\right)\right]$	$exp\left[-\left(\frac{X}{b}\right)^a\right]$	$\frac{a\cdot X^{a-1}}{b^a}$	$\left(\frac{X}{b}\right)^a$

$$\Gamma(a)=\int_0^\infty X^{a-1}\cdot e^{-X}dX;b=\frac{1}{f}$$

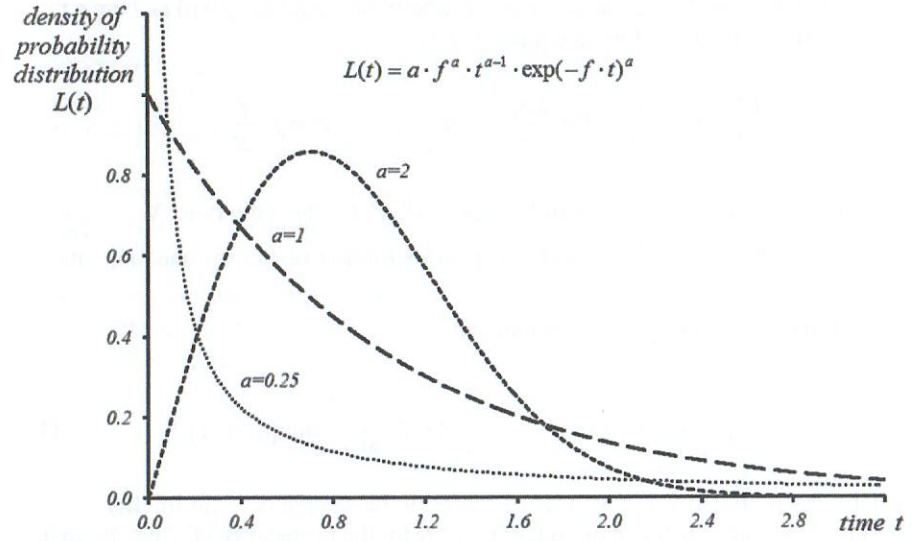

FIGURE 4.1 Density of probability distribution $L(t)$ for the Weibull distribution for different values of parameter a and frequency $f = 1$.

4.4 EXPONENTIAL DISTRIBUTION

For binominal distribution, the probability of the event P with the absence of $n = 0$ undesirable case,[21] with events[22] N, can be written as:[23] $\forall_{n=0}\ P(n=0,N)=(1-R)^N$, where R is the probability of occurrence of an adverse event. In order to express the probability as a function of time $P(n = 0,N) = f(t)$ period of observation t can be divided into a number of sections Δt. Assuming that the value of risk R[24] does not change $R(t) = idem$ for the entire duration of observation t, then with tending to infinity $N \to \infty$ of sections Δt, the probability $P(n = 0,N)$ will be equal to zero:[25] $\forall_{\rho\in(0;1]}\lim_{\Delta t\to 0}P(n=0,N)=\lim_{N\to\infty}(1-R)^n\equiv 0$. Assuming that the risk R is constant $R(t) = idem$, the frequency f of occurrence of adverse events is also constant $f = idem$. According to the frequency definition of probability and the earlier assumption that the risk R does not change $R = idem$, one can write (Kłos R., 2007):

$$R=idem=\frac{ES_N}{N}\to R=\frac{f\cdot t}{N}=f\cdot\Delta t \tag{4.5}$$

where: R – risk, ES_N – value of expected E average number of events S_N at population number N, Δt – time section representing the accuracy of counting the time elapsed, N – population number.

Discrete dependent variable ES_N of the expected average number of events n in discrete function of the independent variable of number of observations N can be replaced with continuous dependent variable of risk R as a function of the independent

variable of number N of equal sections Δt, where the frequency f will be here a factor of proportionality according to equation (4.5):

$$\frac{ES_N}{N} = f \cdot \frac{t}{N} \rightarrow f \equiv \frac{ES_N}{t} \Rightarrow \forall_{\Delta t = \frac{t}{N} = const} R = f \cdot \frac{t}{N} \tag{4.6}$$

Substituting equation (4.6) defining the risk R to the equation: $\forall_{\rho \in (0;1]} \lim_{\Delta t \to 0} P(n=0,N) = \lim_{N \to \infty} (1-R)^N \equiv 0$ and using the definition of the exponential function $\forall_{k \in \mathbb{N}} \lim_{k \to \infty} \left(1+\frac{-a}{k}\right)^k \equiv \exp(-a)$, one can get:

$$\forall_{\rho \in (0;1]} \lim_{\Delta t \to 0} P(n=0,N) = \lim_{N \to \infty} \left(1 - f \cdot \frac{t}{N}\right)^N \equiv \exp(-f \cdot t) \tag{4.7}$$

where: P – probability, n – number of events in the sample of N cardinality.

The cumulative distribution function $F(t, f)$ for the probability (4.7) can be written by using the definition of inverse probability:

$$\forall_{f>0} F(t,f) = P(0 \le T \le t \,|\, f) = 1 - \exp(-f \cdot t) \tag{4.8}$$

and the density L of the exponential distribution can be found by differentiating the cumulative distribution F in equation (4.8): $\forall_{f>0} L(t,f) = \frac{dF(t,f)}{dt} = f \cdot \exp(-f \cdot t)$. Integrating the density of L in the range from zero to infinity, one can show that the exponential distribution is normalized.[26]

4.5 THE WEIBULL DISTRIBUTION

The probability of a survival additional time Δt, when up to now the life time is t, is the conditional probability:

$$P(\Delta t \,|\, t) = \frac{P(\Delta t \cap t)}{P(t)} \tag{4.9}$$

The numerator of equation (4.9) is the probability of survival of cumulative time $t + \Delta t$; therefore, equation (4.9) can be rewritten into the form:

$$P(\Delta t \,|\, t) = \frac{P(\Delta t + t)}{P(t)} = \frac{exp\left[-f(\Delta t + t)\right]}{exp(-f \cdot t)} = exp(-f \cdot \Delta t) \tag{4.10}$$

From equation (4.10) we can see that the exponential distribution conditional probability of survival of extra time Δt[27] is not a function of the current survival time t: $P(\Delta t|t) \neq f(t)$. This is referred to as the independence of the exponential distribution

from current age:[28] $\forall_{P=\exp(-f \cdot \Delta t)} P(\Delta t + t) = P(t) \cdot P(\Delta t)$, but this property is in contradiction with experience.[29]

Modifying this approach by introducing the formula for the density of the probability distribution L, the properties of the exponential distribution $\forall_{f>0} L(t,f) = f \cdot \exp(-f \cdot t)$ can be improved, assuming that the probability of an event for a *single Bernoulli trial*[30] is proportional to the duration of the trial Δt and is independent from the number of the trial (Figure 4.1). Assuming that the probability of an event in an i – trial can be written as: $\forall_{f_i < f_{i+1};\, i>1} P_i = f_i \cdot \Delta t$, where frequency f_i of event occurrence is increased $f_i < f_{i+1}$ with the aging of the organism or device. Thus, assuming the independence of trials, probability of $n = 0$ adverse events in N trials will be reflected by the ratio:

$$\forall_{n=0} P(n=0,N) = \prod_{i=1}^{N} (1 - f_i \cdot \Delta t) \tag{4.11}$$

As before, you can calculate the limit for $N \to \infty \Leftrightarrow \Delta t \to 0$. It is convenient to do so after taking the logarithm of expression (4.11): $\ln P(n=0,N) = \sum_{i=1}^{N} \ln(1 - f_i \cdot \Delta t)$, where the probability can be adopted approximately as follows: $\forall_{\Delta t \to 0} \ln P(n=0,N) \equiv -\sum_{i=1}^{N} f_i \cdot \Delta t$. For infinitesimally small Δt we can calculate the limit: $-\lim_{\substack{N \to \infty \\ \Delta t \to 0}} \sum_{i=1}^{N} f_i \cdot \Delta t = -\int_0^t f(t)\,dt \equiv -\Lambda(t)$. We can therefore write that $\ln P(n=0,N) = -\int_0^t f(t)\,dt$, and then converting, we can obtain dependence of the probability $P(n = 0, N)$ of the absence of adverse events as a function of time t: $\ln P(n=0,N) = -\int_0^t f(t)\,dt$. Thus, similarly to equation (4.8) the cumulative distribution $F(t)$ can be written using inverse probability definition: $F(t,f) = 1 - \exp\left[-\int_0^t f(t)\,dt\right]$. Density L of distribution in this case will be:

$$\forall_{f>0} L(t,f) = \frac{dF(t)}{dt} = f(t) \cdot exp\left[-\int_0^t f(t)\,dt\right] \equiv f(t) \cdot exp\left[-\Lambda(t)\right] \tag{4.12}$$

In particular, for the density of probability distribution L in equation (4.12), when the frequency $f = \frac{n}{N}$ of the occurrence of the adverse event is not a function of time $f \neq f(t)$, equation (4.12) expresses the density for the *exponential distribution*. For $\Lambda(t) \triangleq t^a$, equation (4.12) expresses the density of the Weibull distribution:

$$\forall_{f>0} L(t,f) = f(t) \cdot exp\left[-t^a\right] \tag{4.13}$$

4.6 THE DENSITY DISTRIBUTION OF SURVIVAL TIME

For the continuous random variable T, i.e. the survival time, from the probability density function L one can determine the distribution of the probability of time T occurrence within the interval (t_1, t_2): $P(t_1 \le T \le t_2) = \int_{t_1}^{t_2} L(t)dt$. Cumulative distribution F of time T is given by the formula:

$$F(t) = P(T \le t) = \int_0^t L(t)dt \tag{4.14}$$

In survival analysis it is often preferred to use three alternative functions of probability defining distribution of the random variable T: *survival function* $S(t)$, *hazard function* $h(t)$ and *cumulative hazard function* $H(t)$.

4.7 THE SURVIVAL FUNCTION

Survival function $S(t)$ defines the probability of survival[31] for more than a certain average time t:[32]

$$\forall_{t \ge 0}\, S(t) \equiv P(T > t) = 1 - F(t) \tag{4.15}$$

where: $S(t)$ – survival function.

Survival function[33] $S(t)$ is a nongrowing continuous function, for which (0) = 1. For the Weibull distribution the survival function $S(t)$ is given by the equation: $S(t) = \exp[-(f \cdot t)^a]$. In Figure 4.2, the survival function $S(t)$ is shown for a

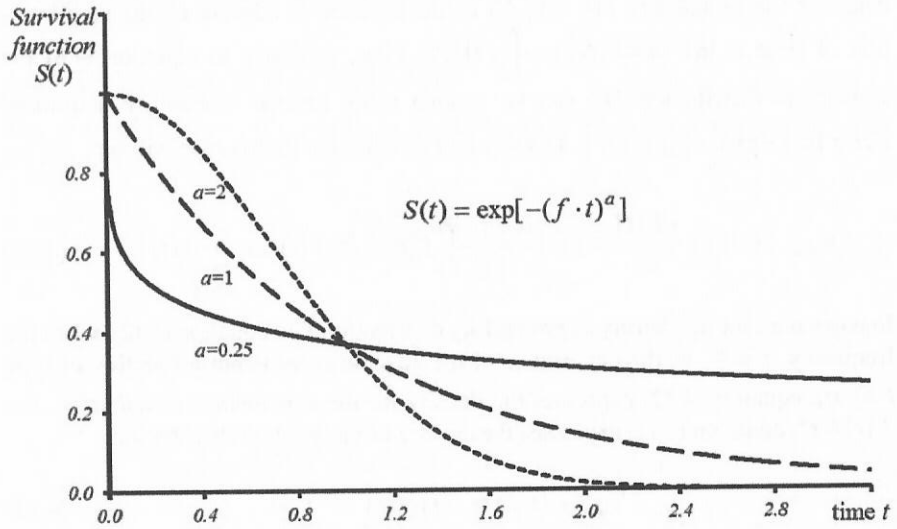

FIGURE 4.2 Survival function $S(t)$ for the Weibull distribution for different values of parameter a = {0.25; 1; 2} and frequency f = 1.

Weibull distribution for different values of the parameter a = {0.25; 1; 2} and frequency f = 1.

The expected value of survival time ET is associated with survival function $S(t)$ by the following formula: $ET = \int_0^\infty S(t)dt$, hence this value is represented by the field under the survival function $S(t)$.

4.8 THE HAZARD FUNCTION

Following the definition of conditional probability, we can write:

$$P(T \le t+\Delta t \mid T > t) \equiv \frac{P(t < T \le t+\Delta t)}{P(T > t)} = \frac{P(T > t \cap T \le t+\Delta t)}{P(T > t)} = \frac{L(t)\cdot\Delta t}{S(t)} = h(t)\cdot\Delta t \tag{4.16}$$

where: $h(t)$ – hazard function.

In equation (4.16) the function $h(t)$ is defined as the *hazard function*: $h(t) = \frac{L(t)}{S(t)}$.

From equation (4.16) the hazard function $h(t)$ can be shown as the limit of the conditional probability per time unit:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(T \le t+\Delta t \mid T > t)}{\Delta t} \tag{4.17}$$

Hazard function $h(t)$ represents the probability of survival time T to be in the vicinity of the selected time t, but it will not occur before that time.[34] It describes the intensity of failures for the selected time[35] t. Hazard function $h(t)$ gives the value of probability per unit of time (4.17), therefore, a case[36] in which its value is greater than 1 may occur. For the Weibull distribution, this function is expressed by the equation: $h(t) = f \cdot a \cdot (f \cdot t)^{a-1}$.

Figure 4.3 shows the selected shapes of the hazard function $h(t)$ for the Weibull distribution for different values of the parameter a for frequency f = 1.

4.9 THE CUMULATIVE HAZARD FUNCTION

The cumulative hazard function $H(t)$ can be defined as an integral of the hazard function $h(t)$:

$$H(t) = \int_0^t h(t)dt \tag{4.18}$$

where: $H(t)$ – cumulative hazard function.

For the Weibull distribution, it is expressed by the equation: $H(t) = (f \cdot t)^a$.

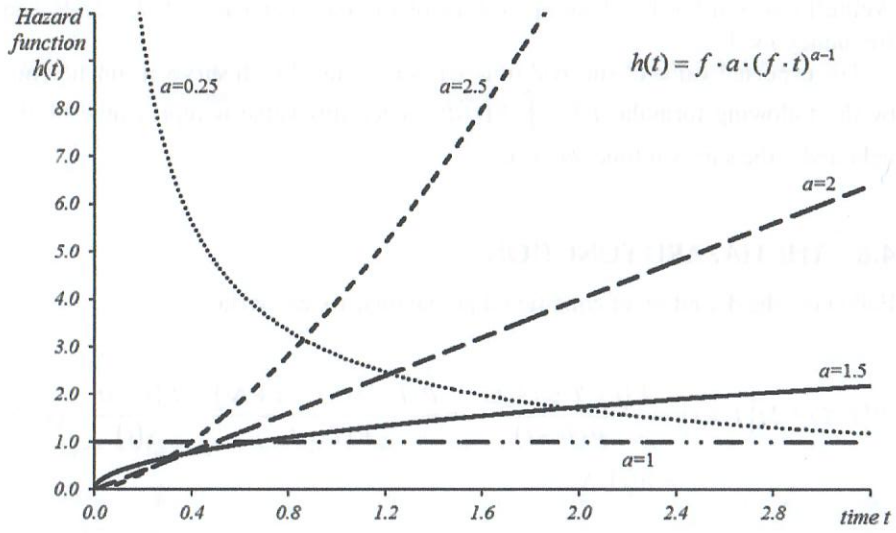

FIGURE 4.3 Hazard function $h(t)$ for the Weibull distribution for different values of parameter a with frequency $f = 1$.

4.10 THE FUNCTION OF RISK

In engineering the hazard function $h(t)$ is sometimes called a *risk function* $R(t)$ or the *failure intensity* $\lambda(t)$ and is defined as the quotient of work time probability density of component $L(t)$ at point t by the probability, for which operating time of component is at least equal t:

$$\forall_{t\geq 0;S(t)\geq 0;\frac{dF}{dt}=\dot{F}(t)\equiv L(t)} h(t)\equiv R(t)\equiv^{\text{def}} \frac{L(t)}{S(t)}\equiv\frac{\dot{F}(t)}{S(t)} \tag{4.19}$$

where: $R(t)$ – risk function.

For discrete distribution of probability of intensity of failure $R(t)$, we can write the equation: $\exists_{P_k=P(k\leftarrow X)} R(t)=\frac{P_k}{\sum_{i=k}^{\infty}P_i}$. Differentiating the survival function $\forall_{t\geq 0} S(t)=P(T>t)=1-F(t)$ an interesting relationship can be shown: $\dot{F}(t)=\frac{d[1-S(t)]}{dt}=-\dot{S}(t)$, which according to (4.19) gives: $\forall_{S(t)>0} R(t)=-\frac{\dot{S}(t)}{S(t)}=-\frac{1}{S(t)}\cdot\frac{dS(t)}{dt}\Rightarrow\int_0^t R(t)dt=-\int_0^{S(t)}\frac{dS(t)}{S(t)}=-\ln|S(t)|=-\ln S(t)$. It results in the following relations:

$$\forall_{S(t)>0} S(t)=exp\left[-\int_0^t R(t)dt\right]=\exp\left[-H(t)\right] \tag{4.20}$$

Equation (4.20) is called the *Wiener relation*. In engineering, cumulative hazard function $H(t)$ is referred to as distribution *function of safety unreliability.*

4.11 INTERFUNCTION DEPENDENCIES

Summarizing the review of the basic functions used in survival analysis, we can conclude that it is sufficient to provide one of them to describe the other since they are related to each other by the dependencies collected in Table 4.3.

In survival analysis most often survival function $S(t)$ and hazard function $h(t)$[37] are used. In practice it is known that a reliable estimate of the specific functions of survival analysis can be done if the sample size is larger than $N = 30$, otherwise the estimation results are burdened. Despite the convergence of different formulas to calculate the characteristic parameters considered here as timelines, there are differences of interpretation between the theory of reliability, safety analysis, and other applications of survival analysis.[38]

4.12 HAZARD

Risk function $R(t)$ can represent the probability of occurrence of *CNSyn* or *DCS* symptoms in function of time t. The formulas $\forall_{t\geq 0} S(t)=P(T>t)=1-F(t)$ and $\forall_{S(t)>0} S(t)=exp\left[-\int_0^t R(t)dt\right]$ can be used to count the cumulative distribution

TABLE 4.3
Useful dependencies for more frequent functions in survival analysis.

$$S(t)=1-F(t)=\int_t^{\infty}L(t)dt$$

$$L(t)=-\frac{d}{dt}\cdot S(t)$$

$$R(t)\equiv h(t)=-\frac{d}{dt}\cdot\ln S(t)$$

$$H(t)=-\ln S(t)=\int_0^t R(t)\cdot dt\equiv\int_0^t h(t)\cdot dt$$

$$S(t)=\exp\left[-H(t)\right]=exp\left[-\int_0^t h(t)\cdot dt\right]$$

F – cumulative distribution
h – hazard function
H – cumulated hazard function
L – probability density
R – risk function
S – survival function

$F(t)$ of the probability of *DCS* or *CNSyn* symptoms occurrence in function of time expressed by the value of risk function $R(t)$:

$$\forall_{\xi(t)\triangleq F(t)} F(t) \equiv 1 - S(t) = 1 - exp\left[-\int_0^t R(t)\cdot dt\right] \tag{4.21}$$

where: $\xi(t)$ – function of *DCS* or *CNSyn* risk identically equal to cumulative distribution of survival time $F(t)$.

Integral of risk function $R(t)$ from time $t = 0$ to t defines the integral of risk occurrence of *CNSyn* or *DCS* during this period of time. Therefore the value of the risk function $R(t)$ from equation (4.21) determines the hazard function $\xi(t)$ of the onset of symptoms of *DCS* or *CNSyn*. The values of risk function $R(t)$ can be determined by matching them to the experimental data. The limits of integration can be also extended to several hours after the dive.

Using survival analysis to mathematical modelling of the risk function $R(t)$ and of hazard function $\xi(t)$ of the onset of *DCS* or *CNSyn* symptoms, these two issues must be distinguished. The risk $R(t)$ is here identified with the probability of *DCS* or *CNSyn*, and the hazard of *DCS* or *CNSyn* is identified with completing the survival function $S(t)$, which is a cumulative distribution of survival: $\xi(t) = 1 - S(t)$. Hazard $\xi(t)$ is the probability of *DCS* or *CNSyn* occurrence provided the level of risk $R(t)$ of *DCS* or *CNSyn* occurrence is accepted.

4.13 RISK OF DECOMPRESSION SICKNESS

Decompression sickness (*DCS*) can take the classic model of tissue supersaturation as a value of risk function $R(t)$. The risk of *DCS* can be calculated for a set of theoretical tissues. Denoting $R_i(t)$ as a function of risk for the ith theoretical tissue, the theoretical survival function $\forall_{S(t)>0} S(t) = \exp\left[-\int_0^t R(t)dt\right]$ can be written as:[39]

$$\forall_{R(t)=\sum_i R_i(t)} S(t) \equiv \prod_i S_i(t) = \prod_i exp\left[-\int_0^t R_i(t)\cdot dt\right] = exp\left[-\int_0^t \sum_i R_i(t)\cdot dt\right] \tag{4.22}$$

A simple model of the risk function $R(t)$ is shown here as an example of matching experimental data using the method of maximum likelihood.

According to the practice with the lapse of time from the end of the decompression, the likelihood of *DCS* occurrence tends to zero. As an algebraic model of *DCS* risk, the decreasing exponential function of time can be accepted: $R(t) = exp(-c \cdot t)$.

There is a moment of time $t = \tau$ after which there will no longer be *DCS* symptoms related to the earlier hyperbaric exposure. As regards the time τ there are different opinions, but most often it is assumed that it is a period not longer than three

days. In the adopted mathematical model we can assume that $\xi(t = \tau) \equiv \xi(t \to \infty)$. Using this assumption and relation (4.20), the survival function $S(t)$ for the risk $R(t) = \exp(-c \cdot t)$ can be written as: $S(t) = exp\left[-\int_0^\infty \exp(-c\cdot t)\cdot dt\right]$. The value of the integral can be calculated[40] as: $\int_0^\infty e^{-c\cdot t}\cdot dt = -\frac{1}{c}\cdot e^{-c\cdot t}\Big|_0^\infty = \frac{1}{c}$. Hence, the value of the survival function is: $S = \exp(-c^{-1})$ and the hazard of $\xi = 1 - \exp(-c^{-1})$.

DCS or *CNSyn* occurrence should be treated as an independent event, having cardinality of n for a selected random sample having cardinality of N. Similarly, absence of *DCS* or *CNSyn* symptoms is regarded as events independent of each other having cardinality of $N - n$ for the selected random sample having cardinality of N. Thus, the probability function Φ defining cumulative probability of an event can be written as the quotient of probabilities of partial events: $\Phi = \xi^n \cdot S^{N-n}$. Because the probabilities are numbers smaller than 1, their quotient is generally much smaller than 1. Therefore, it is more convenient to use logarithmic probability function $\psi \equiv \ln \Phi = n \cdot \ln(1 - S) + (N - n)$. On the basis of experimentally performed dives, the exact value of the parameter c cannot be calculated because the general population can never be examined, only a sample selected from it at random can. However, its most reliable value can be calculated. To this end the maximum of probability function Φ[41] must be found. The necessary condition for the extreme of the function to exist is zeroing out its first derivative $\frac{d}{dS}\Phi \equiv 0$. This condition is equivalent to zeroing out the logarithmic probability function $\frac{d}{dS}\Psi \equiv 0 = -\frac{n}{1-S} + \frac{N-n}{S} \Rightarrow S = \frac{N-n}{N}$. For $S = \exp(-c^{-1})$, $\exp\left(-c^{-1}\right) = \frac{N-n}{N} \Rightarrow c = \left[\ln\left(\frac{N}{N-n}\right)\right]^{-1}$ can be written.

If these criteria were applied to an experiment in which 100 dives were made and where 20 cases of *DCS* were recorded and in 80 cases no *DCS* occurred, it can be written that hazard ξ is[42] $\xi = 1 - \exp(-c^{-1}) = 0.2 \to c \cong 4.48$ and hence the risk function $R(t)$ will be (Wienke B.R., 2003) $R(t) \cong \exp(-4.48 \cdot t)$.

Some of the first attempts to apply survival analysis to evaluating the *DCS* risk were made by Hills, who showed that the survival function $S(\rho) = f(H)$ dependent on risk $R = \frac{n}{N}$ in the coordinates of depth H for eight-hour saturation has Weibull distribution. He did this by showing a good linear dependence of the logarithm of survival function $\ln S$ on the logarithm of shifted saturation depth H (Hills B.A., 1997):

$\forall_{H>0} S\left(R = \frac{n}{N}\right) = \exp\left[-\left(\frac{H-a}{b}\right)^c\right] \to \ln[-\ln S(\rho)] = c\cdot \ln\left(\frac{H-a}{b}\right)$, where: a, b, c – constants of proportionality, H – depth for an eight-hour-long exposure, n – number of cases of *DCS*, N – number of replications for depth H.

Survival analysis is one of the most promising methods for predicting risk of *DCS*, which is bridging the statistical and deterministic methods (Brubakk A.O., Neuman T.S., 2003).

4.14 RISK OF CENTRAL NERVOUS SYNDROME

Survival analysis has been also applied to predict the *central nervous syndrome* (*CNSyn*) risk. One of the proposed algebraic models of the hazard function $h\left(t, p_{O_2}\right)$ is the equation[43] (Harabin A.L., Survanshi S.S., Homer L.D., 1994)

$$\forall_{p_{O_2} \geq p_g} R\left(t, p_{O_2}\right) = a_0 \cdot a_2 \cdot \left(p_{O_2} - p_g\right)^{a_1} \cdot t^{a_2 - 1} \tag{4.23}$$

where: $R\left(t, p_{O_2}\right)$ – risk function of *CNSyn* occurrence, p_{O_2} – oxygen tension, p_g – boundary value of partial oxygen pressure, $a_0..a_2$ – constants.

The parameter a_0 in an algebraic model of the risk function $R\left(t, p_{O_2}\right)$ of *CNSyn* occurrence serves as a scaling factor. The boundary value of oxygen partial pressure p_g expressed in absolute atmospheres is the level below which $p_{O_2} < p_g$ the risk of *CNSyn* occurrence is not accumulated. Constant a_1 and a_2 being exponents enable modelling nonlinearity of the risk function $R\left(t, p_{O_2}\right)$. For $R\left(t, p_{O_2}\right)$ and $a_2 = 1$ the risk function is constant $R = const$. For $a_1 > 0$ the risk function $R\left(t, p_{O_2}\right)$ will grow linearly with the increase in oxygen partial pressure p_{O_2} or more than linearly if $a_1 > 1$. For $a_2 > 1$, the risk function $R\left(t, p_{O_2}\right)$ will increase more than linearly with the increase in exposure time t.

Hazard function[44] ξ, expressing the probability $P\left(t, p_{O_2}\right)$ of *CNSyn* symptom occurrence can be determined from Wiener relation (4.20) and equation (4.21) and written as: $\forall_{R\left(t, p_{O_2}\right) = a_0 \cdot a_2 \cdot \left(p_{O_2} - p_g\right)^{a_1} \cdot t^{a_2 - 1}} \xi \equiv F\left(t, p_{O_2}\right) = P\left(t, p_{O_2}\right)$ $= 1 - \exp\left[-\int_0^t R\left(t, p_{O_2}\right) dt\right]$. Since $\int b \cdot x^{a-1} dx \equiv \frac{b}{a} \cdot x^a$, the hazard ξ of *CNSyn* symptom occurrence for the single exposure procedure can be converted to the form:

$$\xi\left(t, p_{O_2}\right) \equiv F\left(t, p_{O_2}\right) = 1 - exp\left[-a_0 \cdot \left(p_{O_2} - p_g\right)^{a_1} \cdot t^{a_2}\right] \tag{4.24}$$

Parameters $a_0..a_2$ and p_g for equation (4.24) were determined from experimental data (Harabin A.L., 1993) using the method of maximum credibility (Kłos R., 2007) – Figure 4.4, Figure 4.5 and Table 4.4.

Analyzing the data contained in Table 4.3, it can be noted that the model is not satisfactory, because for many parameters the uncertainty of their designation is greater than their set values. We found that the algebraic mathematical model simplified in relation to equation (4.23) of risk R of onset of *CNSyn* symptoms:

$$R(p_{O_2}) = a_0 \cdot (p_{O_2} - p_g)^{a_1} \tag{4.25}$$

gives a better approximation for the experimental data collected in Table 4.4 (Harabin A.L., Survanshi S.S., Homer L.D., 1995).

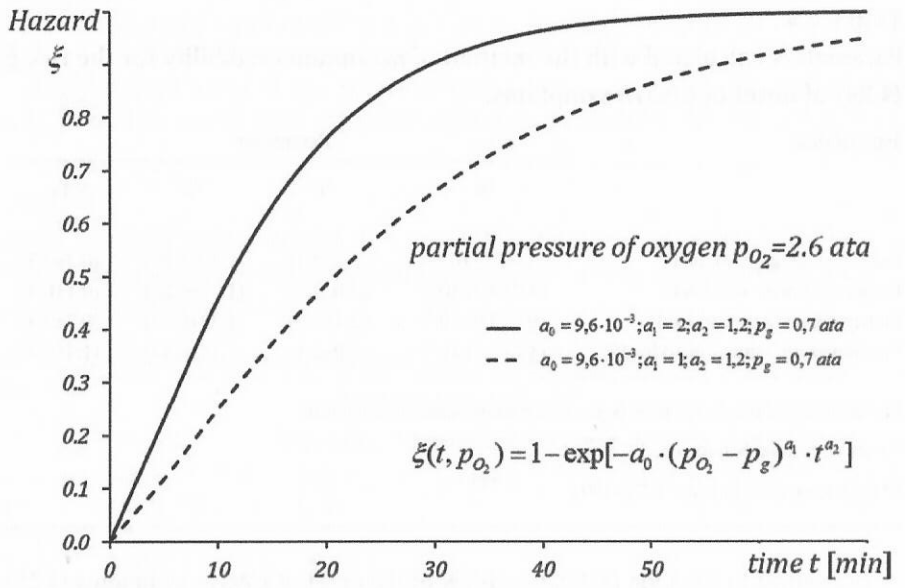

FIGURE 4.4 Hazard function x for dives of the single exposure type rest for the oxygen partial pressure $p_{O_2} = 0.26\,MPa$ and values of a_0, a_1 and p_{ig} estimated from the data for equation (4.24) – solid line, dotted line for $a_1 = 1$ (Yarbrough O.D., Welham W., Brinton E.S., Behnke A.R., 1947; Donald K., 1992).

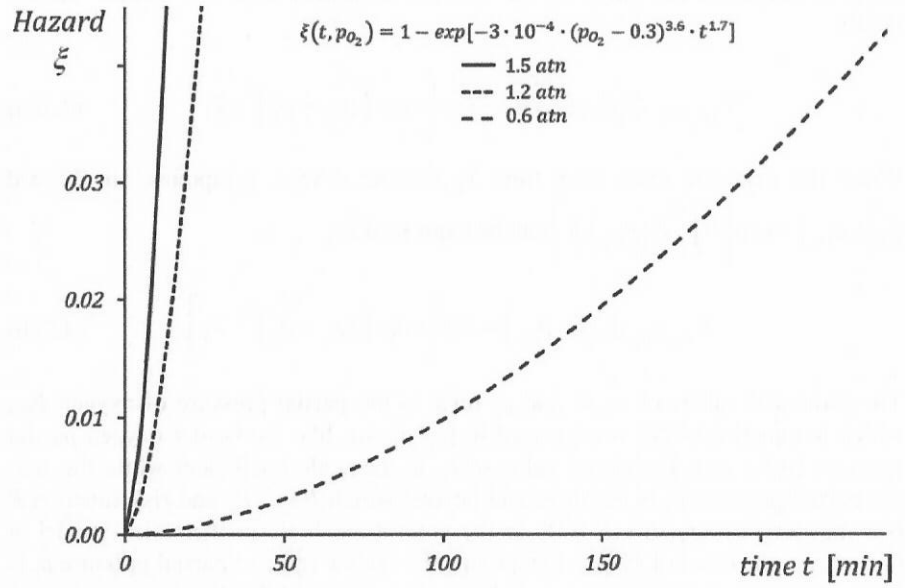

FIGURE 4.5 Hazard x resulting from model (4.24) of single exposure type dives performed in 1979–1986 (Harabin A.L., Survanshi S.S., Homer L.D., 1994).

TABLE 4.4
Parameters calculated with the method of maximum credibility for the risk ξ (4.24) of onset of *CNSyn* symptoms.

Type of data	Parameter			
	a_0	a_1	a_2	p_g [*ata*]
Immersed, under workload**	$(3\pm27)\cdot10^{-4}$	(3.6±5.1)	(1.70±0.33)	(0.3±0.7)
Immersed, under workload*	$(4.7\pm4.7)\cdot10^{-3}$	(3.0±1.8)	(1.27±0.22)	(0.1±0.4)
Immersed, without workload*	$(9.6\pm3.9)\cdot10^{-3}$	(2.1±1.6)	(1.22±0.21)	(0.7±0.4)
Not immersed, without workload*	$(1.8\pm1.8)\cdot10^{-3}$	(3.0±1.1)	(1.75±0.14)	(1.4±0.3)

The accuracy of the designation is given as a single standard deviation.

* tests were done before year 1972

** tests were done in years 1972–1994

In contrast to equation (4.23), the risk R of the onset of *CNSyn* symptoms (4.25) is not a function of time $R\left(p_{O_2}\right) \neq f(t)$.[45] Risk function $R(p_{O_2})$ in equation (4.25) is used to calculate the hazard ξ of the onset of *CNSyn* symptoms from all the observations contained in Table 4.4. When *CNSyn* causes abortion of exposure after the time T_1, the hazard $\xi_1\left(t, p_{O_2}\right) = F\left(t, p_{O_2}\right) = 1 - \exp\left[-\int_0^{T_1} R\left(p_{O_2}\right) dt\right]$ of *CNSyn* symptom onset is calculated according to the formula consistent with the Wiener relation (4.20):

$$\forall_{p_{O_2} \geq p_g}\ \xi_1\left(T_1, p_{O_2}\right) = 1 - exp\left[-a_0 \cdot \left(p_{O_2} - p_g\right)^{a_1} \cdot T_1\right] \tag{4.26a}$$

When the exposure ends after time T_0 without *CNSyn* symptoms, the hazard $\xi_0\left(t, p_{O_2}\right) = \exp\left[-\int_0^{T_0} R\left(p_{O_2}\right) dt\right]$ can be expressed as:

$$\forall_{p_{O_2} \geq p_g}\ \xi_0\left(T_0, p_{O_2}\right) = exp\left[-a_0 \cdot \left(p_{O_2} - p_g\right)^{a_1} \cdot T_0\right] \tag{4.26b}$$

The estimated values of $a_0..a_2$ and p_g refer to the partial pressure of oxygen p_{O_2}, which in equation (4.26) is expressed in $[p_g]$ = *ata*, like the border oxygen partial pressure $[p_g]$ = *ata*. Estimated value of a_0 is the scale coefficient while the border partial pressure p_g is the threshold beyond which $p_{O_2} > p_g$ and risk function R becomes larger than zero ($R > 0$). In the present algebraic mathematical model of hazard ξ of the onset of *CNSyn* symptoms, the border value of partial pressure p_g is always larger than or equal to 1: $p_g \geq 1$ *ata* – it is assumed that there are no signs of *CNSyn* symptoms below the partial pressure of O_2 at $p_{O_2} < 1 ata$. For the estimated coefficient $a_1 = 0$ risk function R is constant, $R = const$, independent of the value of oxygen partial pressure p_{O_2}. When the estimated coefficient $a_1 = 1$, the risk function

R increases linearly together with the increase in oxygen partial pressure p_{O_2}. For the $a_1 > 1$, this increase is faster than linear.

For the fixed value of the border oxygen partial pressure $p_g \equiv 1$ *ata*[46] (Table 4.5), parameter values from algebraic mathematical model of hazard ξ of *CNSyn*

TABLE 4.5
Experimentally tested oxygen dive profiles, which were used to determine the hazard ξ of the onset of *CNSyn* symptoms (Harabin A.L., Survanshi S.S., Homer L.D., 1995).

Profile no.	Exposure profiles with places marked where symptoms of *CNSyn* occurred						*n* /*N*
	Time [*min*]/Depth [*fsw*]						
	Exposures with the change of depth						
1	60/25	15/40†	60/25				1/14
2	120/25‡	15/40††					3/15
3	120/20†	25/35†	95/20				2/20
4	240/20††	25/35†					3/19
5	120/20	25/35†	95/20	25/35†			2/16
6	120/20	15/40†	105/20				1/39
7	240/20	15/40‡					1/16
8	120/20	10/50†	110/20				1/18
9	15/40	30/20	15/40†‡				3/24
10	15/40	45/25	15/40†				1/4
11	15/40†	60/20	15/40†‡				4/47
12	15/40	90/20†‡	15/40	90/20	15/40	15/20	4/11
13	15/40††	225/20††					4/64
Total							30/307
	Exposure to single depth						
14	154/25*						0/12
15	240/25						0/22
16	129/30*††						4/18
17	90/30‡						1/40
18	25/35						0/47
19	30/35††						5/40
20	15/40						0/70
21	20/40‡‡						2/17
22	5/50						0/57
23	10/50						0/58
Total							12/381
Cumulative total							42/688

Cases of absence of *CNSyn* symptoms described in publications by various researchers were collected, also occurrence of symptoms †, and occurrence of seizures ‡, time and depth of dive of exposure end was given

Exposure numbers 1–13 are shallow water profiles with 1–3 trips to depth of {35; 40; 50} *fsw* with the time of stay in a range of 10–25 *min* while 14–23 are exposure to single depth.

N – total number of exposures
n – number of cases of occurrence
* – approximate time of the exposure

occurrence, estimated with the maximum credibility method, were $a_0 = (1.33 \pm 0.22) \cdot 10^{-3}$ and $a_1 = (3.39 \pm 0.5)$.

Most of the research results used in the process of reaching conclusions were obtained in the course of investigations carried out in warm water 22°C.[47] The diver was in a fetal position[48] performing a cyclic task[49] with the use of cycle ergometer in shallow water with an average load – oxygen consumption averaged to $\dot{v}_0 \approx 1.3\,dm^3 \cdot min^{-1}$. Close attention was paid to the content of oxygen and carbon dioxide, trying to keep them at the appropriate levels: $C_{O_2} > 95\%$ and $C_{CO_2} < 1\%$.[50]

Exposures during which the diver dissimulated[51] and on completing them claimed that he had *CNSyn* symptoms were excluded from the analysis. The symptoms regarded as finishing the exposures are nausea, numbness, dizziness, cramps, impaired hearing and vision, loss of consciousness and convulsions.[52]

The results of matching the algebraic models of risk $R(p_{O_2})$ as expressed by equation (4.25) and the danger $\xi\left(T, p_{O_2}\right)$ of the onset of *CNSyn* symptoms, expressed by equation (4.26), are shown in Table 4.6.

Hazard $\xi\left(T, p_{O_2}\right)$ of the onset of symptoms of *CNSyn* for single exposure type dives[53] was calculated directly from the model (4.26) (Figure 4.6).

For the multilevel dives, danger was calculated separately for each level and then those values were added. Examples of calculation for profile 13 from Table 4.6 are shown in Figure 4.7.

Number n of predicted events with occurrence of *CNSyn* was calculated for each profile from Table 4.5 by multiplying the values obtained for hazard $\xi\left(T, p_{O_2}\right)$ by the number of experimental dives N (Table 4.6).

The presented method of modelling and predicting the hazard ξ of onset of *CNSyn* symptoms raises no objections as regards the profiles of single type exposure. However, for multilevel profiles, it assumes that there is no difference between the duration sequences of stay at the individual depths. For example, there is no difference between exposures 6 and 13 from Table 4.5.

According to the theory of the biochemical mechanisms of *CNSyn* symptoms occurrence, there should be a difference between the profile in which the trip to the greater depth is at the beginning of the transit and that in the middle or at the end of exposure. It is possible to conceive of a situation in which the statistical average production of harmful metabolites during multilevel oxygen exposure poses a relatively low risk, as in the case under study. Increased production of the harmful radicals during a trip is stopped, gradually purged or relaxed during a stay at a smaller depth. This effect, however, can be small when oxygen exposure continues, in comparison to the effects produced during rest under norm-baric conditions. Hence, for exposures that differ in the moments the trip starts, there should occur different concentrations of harmful radicals after a dive, giving differences in the initial state for the repetitive dive. If, however, the time until the next dive is long enough to complete full deactivation of harmful metabolites, such exposures are indistinguishable when this method of inference is used.

TABLE 4.6

Comparison of the danger ξ of onset of *CNSyn* symptoms defined using the algebraic semi-empirical mathematical models (4.26) with the binominal distribution (Harabin A.L., Survanshi S.S., Homer L.D., 1995).

Profile no. from Table 4.5	No. of observations	No. of *CNSyn* cases	Binominal distribution			Algebraic model (4.26a)	
			Confidence interval†		Predicted number of *CNSyn*	Hazard of *CNSyn*	Predicted number of *CNSyn*
	N	n	p_1	p_r	n_b	ξ	n
Exposures with the change of depth							
1	14	1	0.0018	0.3387	0.0–4.7	0.0989	1.4*
2	15	3	0.0433	0.4809	0.6–7.2	0.0979	1.5*
3	20	2	0.0123	0.3770	0.2–7.5	0.0915	1.8*
4	19	3	0.0338	0.3958	0.6–7.5	0.0966	1.8*
5	16	2	0.0155	0.3835	0.2–6.1	0.1312	2.1*
6	39	1	0.0006	0.1348	0.0–5.3	0.0916	3.6*
7	16	1	0.0016	0.3023	0.0–4.8	0.0943	1.5*
8	18	1	0.0014	0.2729	0.0–4.9	0.1082	1.9*
9	24	3	0.0266	0.3236	0.6–7.8	0.0485	1.2*
10	4	1	0.0063	0.8059	0.0–3.2	0.0982	0.4*
11	47	4	0.0237	0.2038	1.1–9.6	0.0897	4.2*
12	11	4	0.1093	0.6921	1.2–7.6	0.1597	1.8*
13	64	4	0.0173	0.1524	1.1–9.8	0.0909	5.8*
Exposure to single depth							
14	12	0	0.0000	0.3187	0.0–3.8	0.0768	0.9‡
15	22	0	0.0000	0.1889	0.0–4.2	0.1171	2.6‡
16	18	4	0.0641	0.4764	1.2–8.6	0.1168	2.1‡
17	40	1	0.0006	0.1316	0.0–5.3	0.0830	3.3‡
18	47	0	0.0000	0.0933	0.0–4.4	0.0398	1.9‡
19	40	5	0.0419	0.2680	1.7–10.7	0.0475	1.9‡
20	70	0	0.0000	0.0637	0.0–4.5	0.0376	2.6‡
21	17	2	0.0146	0.3644	0.2–6.2	0.0498	0.8‡
22	57	0	0.0000	0.0776	0.0–4.4	0.0268	1.5‡
23	58	0	0.0000	0.0763	0.0–4.4	0.0529	3.1‡

†according to (Kłos R., 2007):

$$\forall_{0<n<N} \begin{cases} P(\rho \geq p_r) = \sum_{x=0}^{n} \binom{N}{x} \cdot p_r^x \cdot (1-p_r)^{N-x} = \dfrac{\alpha_0}{2} \\ P(\rho \leq p_l) = \sum_{x=n}^{N} \binom{N}{x} \cdot p_l^x \cdot (1-p_l)^{N-x} = \dfrac{\alpha_0}{2} \end{cases};$$

TABLE 4.6 *(Continued)*
Comparison of the danger ξ of onset of *CNSyn* symptoms defined using the algebraic semi-empirical mathematical models (4.26) with the binominal distribution (Harabin A.L., Survanshi S.S., Homer L.D., 1995).

$$\forall_{n=0}\begin{cases} p_l = 0 \\ p_r = 1 - 10^{\lg\frac{\alpha_0}{N}} \end{cases};\quad \forall_{n=N}\begin{cases} p_l = 10^{\lg\frac{\alpha_0}{N}} \\ p_r = 1 \end{cases}$$

$$^{\ddagger} n = N \cdot \{1 - \exp[-1{,}33 \cdot 10^{-3} \cdot (p_{O_2} - 1)^{3.39} \cdot t]\}$$

N – general number of exposures from Table 4.5
t –exposure time from Table 4.5
n – number of *CNSyn* cases
* adding the values calculated according to formula from the note ‡

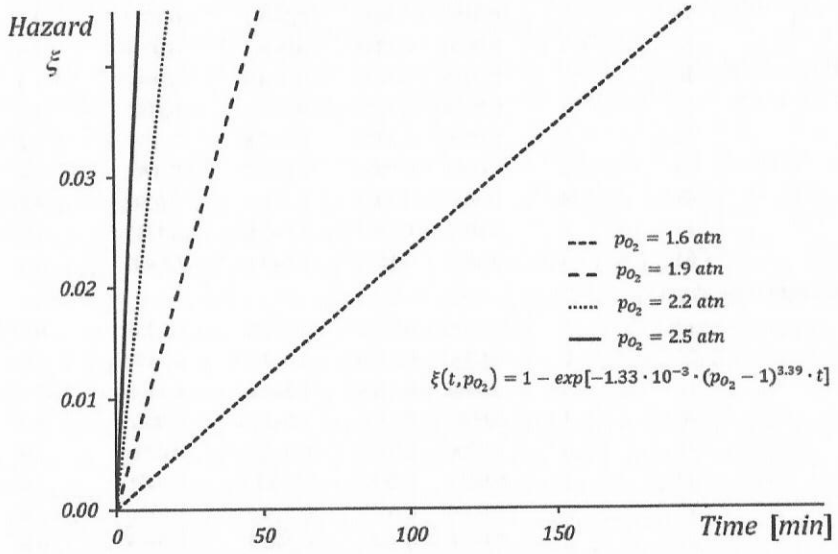

FIGURE 4.6 Hazard ξ for single type exposure dive defined according to model (4.26a).

4.15 RISK COMPARISON

According to model (4.26), there is no difference in predicting hazard ξ of *CNSyn* onset of symptoms. This is confirmed by the small difference in values of profiles 6 and 13. If there were a difference between profiles 6 and 13, it could be manifested by a range of prediction of *CNSyn* cases calculated from a binominal distribution (Table 4.6).

When comparing the risk values, a set of hypotheses can be put forward:

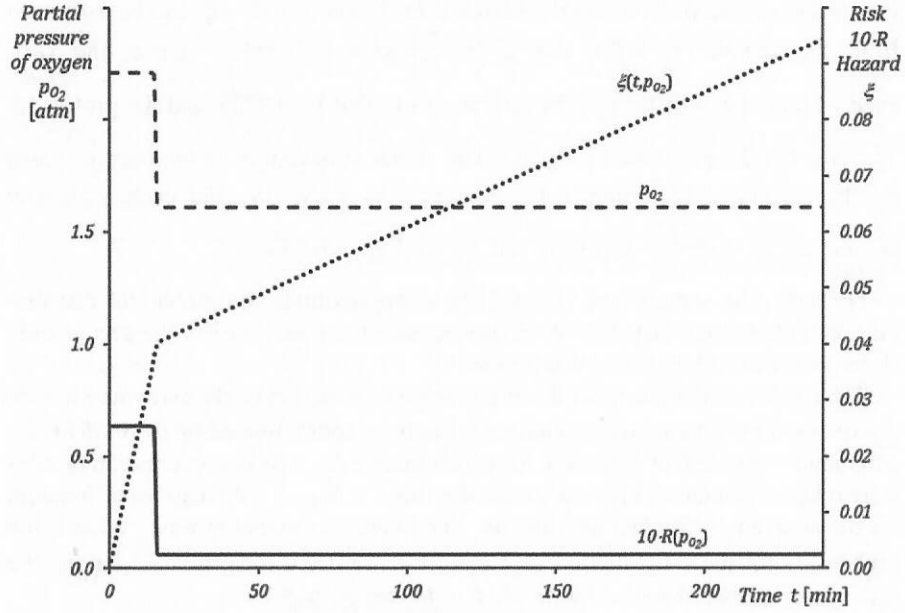

FIGURE 4.7 The dependence of the cumulative hazard ξ and risk R on time t for multilevel dive no. 13 from Tables 4.5 and 4.6 (Harabin A.L., Survanshi S.S., Homer L.D., 1995).

$$\begin{cases} H_0 : \xi_6 = \xi_{13} \\ H_0 : \xi_6 \neq \xi_{13} \end{cases} \tag{4.27}$$

From Table 4.5 we learn that for profile 6 with the number of performed tests $N_6 = 39$, there was $n_6 = 1$ – one case of *CNSyn* symptoms; with $N_{13} = 64$ performed tests there were $n_{13} = 4$ cases of *CNSyn* symptoms for profile 13. To verify whether different moments to begin trips have an influence on the risk of onset of *CNSyn* symptoms the probability of the results of experiments can be calculated, assuming that the null hypothesis H_0 is true.

The results of the experiments can be treated as samples of two populations.[54] Because the average for the two populations is not known, they should be estimated from the sample. Assuming respectively that the value of random variable is $X \leftarrow 1$ if symptoms of *CNSyn* occur and $X \leftarrow 0$ as the case of absence of *CNSyn* symptoms, according to the definition the estimation of expected value $EX_i \equiv \Sigma_i X_i \cdot P_i$ from the sample i for binominal distribution, the estimator of hazard ξ_i can be written as: $\hat{\xi}_i = \frac{1 \cdot n_i + 0 \cdot (N_i - n_i)}{N_i} = \frac{n_i}{N_i} \equiv \overline{x}_i$, where $\overline{x}_i$ is the average value from the sample (Kłos R., 2007). From here estimators of expected values for each profile can be calculated: $\hat{\xi}_6 \equiv \overline{x}_6 = \frac{n_6}{N_6} = \frac{1}{39} \cong 0.0256$ and $\hat{\xi}_{13} \equiv \overline{x}_{13} = \frac{n_{13}}{N_{13}} = \frac{4}{64} \cong 0.0625$.

For the binominal distribution the variance $D^2 X \equiv N \cdot x \cdot (1 - \xi)$ can be estimated from the sample: $D^2 X \triangleq s_i^2 = N \cdot \hat{\xi}_i \cdot \left(1 - \hat{\xi}_i\right)$(Kłos R., 2007). Hence, the variance of profile 6 will be: $s_6^2 \cong 39 \cdot 0.0256 \cdot (1 - 0.0256) \cong 0.9728$ and for profile 13: $s_{13}^2 \cong 64 \cdot 0.0625 \cdot (1 - 0.0625) \cong 3.755$. The standard deviation of the average value can be calculated according to the equation: $s_{\overline{x}_i} = \sqrt{\hat{s}_i^2 / N}$. For each trial, they amount to: $s_{\overline{x}_6} \cong \sqrt{\frac{0.9728}{39}} \cong 0.1579$ and $s_{\overline{x}_{13}} \cong \sqrt{\frac{3.75}{64}} \cong 0.2421$.

Formally, the sample size is sufficient to approximate the binominal distribution with the normal distribution. However, since the size of one of the groups only slightly exceeds 30, the *t-test* was applied.

The calculations were made for the average distribution of the estimated hazard $\hat{\xi}_6$ expressed by binominal distribution. It has been approximated by *t – Student* distribution for profile 6 of Table 4.5, for which sample $N_{13} = 64$ decompression profiles were taken at random. This resulted in adopting $\nu = N_{13} - 1 = 63$ degrees of freedom for the adopted *t – Student* distribution. The interesting question was, "What is the probability for the estimated average danger $\hat{\xi}_{13}$ to occur for profile 13 having the same value as that observed for profile 6 or higher $\hat{\xi}_{13} \geq \hat{\xi}_6$."

Using the formula $t = \left(\overline{x} - \mu\right) : \sqrt{s^2 / N - 1}$ referred to as the statistics, where: $\overline{x}$ represents average from sample, μ is the true value of binominal distribution describing the sample, s^2 represents sample variance, N the size of the sample, the probability a_0 of the statistical distribution of sample 6 can be calculated, wherein the estimated hazard x is the same or greater than the estimated hazard for distribution 13. For this purpose, the distribution for sample 6 is treated as a reference. Hence, for the statistics t value of $\mu \equiv \overline{x}_6$, $s = s_{\overline{x}_6}$, $N = 40$, while the test value will be $\overline{x} = \overline{x}_{13}$: $t = \frac{\overline{x}_{13} - \overline{x}_6}{s_{\overline{x}_6}} \cong \frac{0.0625 - 0.0256}{0.1579} \cong 0.2337$. For the *t – Student* distribution value of statistic $t \cong 0.2337$ corresponds to the value of probability $\alpha_0 \left(N_{13} - 1 = 63\right) \cong 0.40800$.

From the performed calculations we can see that of all 64 samples of the profile 6 population, which could be selected at random, approximately 41% have the same or higher value of the estimated risk as in the case of the profile $\hat{\xi}_{13}$. Based on the studies carried out and estimates made, it cannot be concluded that the estimated risks ξ of *CNSyn* symptoms for both populations and the exposure distribution applied differ significantly.[55] The result justifies the way accepted for estimating risk, i.e. without taking into account the sequence of trips made.

4.16 SUMMARY

The methods of survival analysis were introduced to the problems associated with diving by Weathersby and Thalmann (Gerth W.A., 2002). The hazard prediction model ξ of *CNSyn* symptoms, proposed by the US Navy, which derives from this theory, seems to be sufficiently precise. Its strength is that several researchers have confirmed the same border value of oxygen partial pressure $p_g \equiv 1\,ata$, beyond which

the danger of *CNSyn* is expected. This limits the safe exposure time when breathing a medium containing oxygen under partial pressure above the magnitude defined by this border value $p_{O_2} > p_g$.

In the proposed algebraic mathematical model, the dose of *CNSyn* toxicity accumulates only during the dive. The dose of *CNSyn* toxicity is independent of the sequence of the dive phases.[56]

However, it seems that despite identical durations of individual phases of the dive, there should be a difference between the profile which, for example, starts exposure with a trip to a greater depth and the profile in which the trip is undertaken at the end of dive. The results of the tests allowed for statistically comparing two profiles for differences in the risk run at different times of the trip. The statistical inference process applied to them has not given any grounds for rejecting the null hypothesis, which refers to an absence of statistically significant differences in risk carried when changing the sequence of trips.

NOTES

1 For example, the analysis of time from the start of treatment until disease recurrence, death, etc.
2 For example, time of unemployment, primiparity (a medical term used to refer to a condition or state in which a woman is bearing a child for the first time and/or has given birth to an offspring at one time).
3 For example, time of damage of the piece of equipment, uptime, etc.
4 Depending on the scheme in question, it can be called the time of trouble-free operation, duration, expectations or answers.
5 For example, in estimating survival of a heart attack, the starting point may be related to the time of onset of symptoms, admission to hospital, start of a particular treatment, etc.
6 Such as day, hour, minute, etc.
7 For example, in reliability engineering it can be expressed in a number of cycles performed by a machine until failure occurs.
8 For example, for the two alternative methods of treatment.
9 Data are called cutoff when their value cannot be accurately estimated – this will be discussed later.
10 For example, binominal, normal, F, χ^2, t, etc.
11 In typical statistical inference, one would have to wait until the patient dies, but this may take many years or decades, i.e. sometimes it is unrealistic to wait for the established endpoint of the study to occur.
12 For example, for dead patients the survival time T can be defined.
13 That is, survival time T is greater than the end point of test t – it is known that survival time T is within the range (t; ∞).
14 A patient may be withdrawn from the program of research, contact with him may be lost at the end of the test cycle, there is no economic or practical justification of patient monitoring until a planned final event occurs, etc.
15 Very frequent irregularities in practice.
16 Sufficient conditions for data not to carry significant information about the nature of the phenomenon is independence of survival time T from final time t.
17 As a result of illness or loss of communication with the cured patient, or as a result of his/her failure to turn up for appointments.

18 Similarly, taking into consideration data "incorrectly" cut off may lead to inadequate conclusions.
19 Some technical systems (e.g. pressure vessels, elevators, cranes) require a periodic certification process, after the period of permitted operation, the device is usually suitable for safe use, but despite its technical efficiency, its exploitation is prohibited by law.
20 Generalized distribution Γ – Table 4.1.
21 For example, the onset of *CNSyn* symptoms.
22 For example, at N time periods.
23 As it will be shown later, the start of thinking about the inverse event simplifies the introduction of cumulative distribution of risks.
24 For example, risk of the onset of *CNSyn* symptoms.
25 There is always a possibility of even a small risk, thus postulating risk values at zero $R \equiv 0$ is in contradiction with observed reality.
26 $\forall_{f>0} \int_0^\infty L(t,f) \cdot dt = f \cdot \int_0^\infty e^{-f \cdot t} \cdot dt = f \cdot \left. \frac{1}{-f} \right|_0^\infty = 0 - \frac{f}{-f} = 1.$
27 When the lifetime was so far t.
28 Remaining lifetime Δt does not depend on the past and has the same exponential distribution, as up to now survival time is t.
29 As a rule, people die and the machines start to deteriorate after reaching a certain age.
30 For example, during a single diving cycle.
31 The probability of failure-free operation, of survival, of any other defined event, etc.
32 For example, $S(t)$ is the probability that a given person will live up to time t.
33 In engineering an equivalent of survival function $S(t)$ is used to determine the reliability and is called safety and *reliability function*.
34 The value of hazard function $h(t)$ should be considered as a potential for occurrence of the expected event (mostly unsuccessful) showing that the analysis of the problem situation characterized by the survival function $S(t)$ has been completed; when the function $S(t)$ decreases, the $h(t)$ grows. The function $h(t)$ can be graphically compared to the operation of car speedometer. On the basis of its constant it can be concluded what distance will be covered after the selected time has elapsed; on the basis of the fixed value of the function $h(t)$, the number of expected events within the selected time can be concluded.
35 In engineering safety, the hazard function $h(t)$ is defined as *intensity of safety failure* and is often written as $\lambda_B(t)$.
36 Depending on the adopted units of time.
37 One important reason for using the hazard function $h(t)$ is that the conditional distribution of the expected survival time beyond the moment t_0 can be calculated directly from it for $h(t > t_0)$.
38 For example, a measure of reliability is the probability of meeting the system's requirements per unit of time, while a measure of hazard is the probability of an adverse situation occurring for the space-time surrounding the system.
39 Depending on the adopted explication.
40 Using the formula: .
41 Search for the dominant using the method of maximum credibility.
42 Constant c has the inverse of a unit of time; in this case it may be $[c] = h^{-1}$.
43 Dependence equivalent to the risk function , which defines the probability of *CNSyn* occurrence for the procedure of single exposure type, in conjunction with exposure to oxygen partial pressure , to which a diver is exposed for time t.
44 Which is a cumulative distribution function F.
45 Risk function R describing the intensity of the *CNSyn* occurrence at a selected time in (4.25) is constant value $R = const$ independent of the time for oxygen partial pressure ; the *CNSyn* risk dependency on time occurs when hazard is being calculated.

46 Estimated value of the border partial pressure for data from Table 4.5 was $p_g = (1.3 \pm 0.4)$, but the value $p_g \equiv 1$ *ata* was adopted as more physiological.
47 The tests carried out in water at 12°C and 4°C were not taken into consideration because they pose an additional threat identified earlier by Donald (Donald K.W., 1992).
48 A cycle ergometer was submerged in a shallow container set in a hyperbaric chamber so that the diver leaning while sitting on a cycle ergometer at any time could raise his head and after removing the mouthpiece could start to breathe the air in the chamber (Harabin A.L., Survanshi S.S., Homer L.D., 1995).
49 6 *min* work and 4 *min* rest (Harabin A.L., Survanshi S.S., Homer L.D., 1995).
50 It was not specified whether this tension was related to the atmospheric pressure.
51 This is understood here as presence manifested in an attempt to meet the norms or hide the actual situation and feelings, impulses, behaviors, etc.
52 The frequency distribution of particular symptoms is shown in Table 3.2.
53 Exposures at one value of the partial pressure of oxygen throughout the experimental diving – profiles 14–23 from Tables 4.5 and 4.6.
54 Infinitely many such samples can be taken from them, and the distribution of averages from such samples will tend, following the central limit theorem, towards a normal distribution with the population average estimated as the average of the sample and a standard deviation estimated as the standard deviation of the sample divided by the square root of the sample size (Kłos R., 2007).
55 Running a formal *t-test* for unrelated samples has the same effect, which says that there are no grounds to reject the null hypothesis H_0 (4.27) at the significance level of $a_0 = 5\%$.
56 Hazard of *CNSyn* for profiles with a trip to the greater depth does not depend on the time of the trip.

REFERENCES

Brubakk AO & Neuman TS. 2003. *Bennett and Elliott's Physiology and Medicine of Diving.* London: Saunders. ISBN 0-7020-2571-2.

Donald KW. 1992. *Oxygen and the Diver.* Harley Swan: The SPA Ltd. ISBN 1-85421-176-5.

Evans M, Hastings N & Peacock B. 1993. *Statistical Distributions.* New York: John Willey & Sons, Inc. ISBN 0-471-55951-2.

Gerth WA. 2002. Overview of survival functions and methodoology. In *Survival Analysis and Maximum Likelihood Techniques as Applied to Physiological Modeling,* ed. PK Weathersby & WA Gerth, 1–28. Kensington: Undersea and Hyperbaric Medical Society Inc.

Harabin AL. 1993. *Human Central Nervous System Oxygen Toxicity Data from 1945 to 1986.* Bethesda: Naval Medical Research Institute. NMRI 93-03; AD-A268-225.

Harabin AL, Survanshi SS & Homer LD. 1994. *A Model for Predicting Central Nervous System Toxicity from Hyperbaric Oxygen Exposure in Man: Effects of Immersion, Exercise, and Old and New Data.* Bethesda: Naval Medical Research Institute. NMRI 94-0003; AD-A278 348.

Harabin AL, Survanshi SS & Homer LD. 1995. A model for predicting central nervous system toxicity from hyperbaric oxygen exposure in humans. *Toxicol. Appl. Pharmacol.*, 132, 19–26.

Hills BA. 1997. *Decompression Sickness – The Biophysical Basis of Prevention and Treatment: Vol. 1.* Chichester: John Willey&Sons. ISBN 0-471-99457-X.

Kłos R. 2007. *Zastosowanie metod statystycznych w technice nurkowej - Skrypt.* Gdynia: Polskie Towarzystwo Medycyny i Techniki Hiperbarycznej. ISBN 978-83-924989-26.