A stochastic-covariate failure model with an application to case-control analysis

Simeon M. Berman *

Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

Received 7 July 1999; received in revised form 18 April 2000; accepted 16 May 2000

Abstract

A stochastic process \( X(t) \) is periodically stationary (and ergodic) if, for every \( k \geq 1 \) and every \((t_1, \ldots, t_k)\) in \( R^k \), the sequence of random vectors \((X(t_1 + n), \ldots, X(t_k + n))n = 0, 1, \ldots, \) is stationary (and ergodic). For such an ergodic process, let \( T \) be a positive random variable defined on the sample space of the process, representing a time of failure. The local failure-rate function is assumed to be of the form \( up(x), -\infty < x < \infty, \) where \( p(x) \) is a non-negative continuous function, and \( u > 0 \) is a small number, tending to 0; and, for each \( u, T = Tu \) is the corresponding failure-time. It is shown that \( X(Tu) \) and \( uTu \) have, for \( u \to 0 \), a limiting joint distribution and are, in fact, asymptotically independent. The marginal distributions are explicitly given. Let \( Y \) be a random variable whose distribution is the limit of that of \( X(Tu) \). Under the hypothesis that \( p(x) \) is unknown or of known functional form but with unknown parameters, it is shown how \( p(x) \) can be estimated on the basis of independent copies of the random variable \( Y \). The results are applied to the analysis of a case-control study featuring a ‘marker’ process \( X(t) \) and an ‘event-time’ \( T \). The event in the study is considered to be particularly rare, and this is reflected in the assumption \( u \to 0 \). The control-distribution is identified with the average marginal distribution of the (periodically stationary) marker process \( X(t) \), and the case-distribution is identified with that of \( Y \). The particular application is a biomedical trial to determine the risk of stroke in terms of the level of an anticoagulant in the blood of the patient. © 2000 Elsevier Science Inc. All rights reserved.

MSC: 62P10; 62N06

Keywords: Case-distribution; Control-distribution; Failure-rate function; Hermite polynomial; Method of moments; Periodic stationarity; Prothrombin time; Stochastic-covariate process

* Tel.: +1-212 998 3182; fax: +1-212 995 4121.
E-mail address: sberman@cims.nyu.edu (S.M. Berman).

0025-5564/00/$ - see front matter © 2000 Elsevier Science Inc. All rights reserved.
PII: S0025-5564(00)00029-8
1. Introduction

Let \( X(t), -\infty < t < \infty \), be a real periodically stationary stochastic process on some probability space, and let \( T \) be a positive random variable on the same space. The conditional distribution of \( T \), given \( X(s), -\infty < s < \infty \), is specified as follows. Let \( F(x) \) be some distribution function with support \([0, \infty)\), and \( Q(x) \) a non-negative continuous function on \( \mathbb{R}^1 \); then it is stipulated that, for \( t > 0 \),

\[
P(T \leq t \mid X(s), -\infty < s < \infty) = F\left( \int_0^t Q(X(s)) \, ds \right).
\]

The conditional hazard function of \( T \), given \( X(t) \), is equal to

\[
\frac{Q(X(t))}{1 - F\left( \int_0^t Q(X(s)) \, ds \right)}.
\]

if \( F \) has a density \( F' \). In the particular case where \( F \) is the standard exponential distribution, \( Q(X(t)) \) itself is the conditional hazard function, and (1.1) has been extensively studied as a failure-time distribution with a stochastic covariate hazard function. \( T \) plays the role of the failure time. See [1] and the references therein, and the monograph of Andersen et al. [2].

The particular hypothesis in this study concerning the form of \( Q \) and the formulation of the probabilistic results were motivated by a biomedical study which was originally analysed by a conventional case-control approach (see [3]). The relation of our statistical results to the biomedical problem is discussed in detail in Section 5. The mathematical setting of our results is as follows. The function \( Q \) is taken to be of the form \( Q_u(x) = up(x) \), where \( p(x) \) is a non-negative continuous function, and \( u > 0 \) is a real index that converges to 0. Put \( T = T_u \), and so, by (1.1)

\[
P(T_u \leq t \mid X(s), -\infty < s < \infty) = F\left( u \int_0^t p(X(s)) \, ds \right).
\]

The hypothesis of periodic stationarity of \( X(t) \) is a weakened version of stationarity that is needed for our particular application. It is equivalent to the condition that the finite-dimensional distribution of each vector \((X(t_1), \ldots, X(t_k))\) is invariant under the shift \( t_i \rightarrow t_i + 1, \ i = 1, \ldots, k \). Put

\[
G(x) = \int_0^1 P(X(t) \leq x) \, dt,
\]

and let \( X \) be a random variable with the distribution function \( G \). Note that \( G \) is the average of the marginal distribution of \( X(t) \) for \( t > 0 \) because \( P(X(t) \leq x) \) is a periodic function.

Theorem 2.1 furnishes the limiting joint distribution of \( T_u \) and \( X(T_u) \) for \( u \to 0 \). It states that \( uT_u \) (the scaled failure-time) and \( X(T_u) \) are asymptotically independent for \( u \to 0 \). The limiting distribution of \( uT_u \) is \( F(tEp(X)), \ t > 0 \), and that of \( X(T_u) \) is

\[
\frac{\int_{-\infty}^{x} p(y) \, dP(X \leq y)}{\int_{-\infty}^{\infty} p(y) \, dP(X \leq y)}.
\]

The paper is concerned with the estimation of the function \( p(x) \) on the basis of observed data, and the application of such an estimate to the case-control analysis. The function \( p(x) \) is the
asymptotic failure-rate function, and will be referred to simply as the failure-function. The particular problems that we will consider depend only on the distribution \((1.2)\) and the limiting distribution of \(X(T_u)\), for \(u \to 0\). The form of the distribution function \(F\) in (1.1) is irrelevant for our purposes. The only assumption on the \(X(t)\)-process is that it is periodically stationary and ergodic.

The statistical data are presumed to consist of two components. The first is the knowledge of the distribution \(G\) either by assumption or a good estimate based on a large sample. In nearly all of the statistical analysis in this study we take \(G\) to be standard normal. The second component consists of a sample of \(n\) observations on the random variable \(X(T_u)\), the \(X\)-value at the time of failure, where we use the limiting distribution for \(u \to 0\) as the sampled distribution. This work falls in the same area as other recent work on the estimation of failure-rates on the basis of the observation of so-called marker processes with stochastic covariates. In addition to the references already cited we mention recent joint work of Berman and Frydman [4], Fusaro et al. [5], Jarrow et al. [6], Jewell and Kalbfleisch [7], Tsistias et al. [8], Self and Pawitan [9], and Yashin [10]. Our particular work involves observation of the path of the marker \(X\)-process only at the time of failure.

The proposed methods of estimation for \(p(x)\) are two types: the semi-parametric and the parametric. In the former, \(p(x)\) is taken to be an arbitrary non-increasing positive function; and, in the latter, \(p(x)\) is assumed to be of a known parametric form \(p(x; \theta)\) with an unknown parameter \(\theta\). The term semi-parametric in the context of a failure model might suggest to the reader some relation to the fundamental paper of Cox [11]. However, it will become evident that any such connection is at most superficial.

While much of the work in hazard analysis with stochastic covariates, and this paper in particular, has been motivated by applications to clinical biomedical data, the results are sufficiently general to apply to other areas as well. For example, Jarrow et al. [6] apply such analysis to a financial problem where failure is associated with the default of a bond and the marker is the record of successive credit ratings.

A natural assumption for the application in our work as well as other work is that the failure-function is monotonic. Here we take it to be non-increasing. In such a case the methods proposed here introduce a new approach to the analysis of what has long been called ‘extreme-value data’. This is discussed in Section 2. In addition to the assumptions under which Theorem 2.1 holds, other technical conditions will be imposed for the purpose of applying – not proving – the theorem. These are:

(i) \(\int_0^\infty t dF(t)\) is finite (see (2.11)).
(ii) \(E\hat{p}^2(X)\) is finite (see (2.12)).
(iii) \(p(x)\) is non-increasing.

The outline of the paper is as follows. The main probabilistic result, Theorem 2.1, is in Section 2. In Section 3, we introduce semi-parametric models and estimation procedures, and define the expected failure-time function and describe its significance and methods for estimating it. In Section 4, we introduce specific parametric models for the failure-function and estimation procedures by the method of moments. Section 5 describes the biomedical problem that motivated the study, the statistical study of Hylek et al. [3], and then identifies the control distribution with the average marginal distribution of \(X(t)\) and the case-distribution with the limiting distribution of \(X(T_u)\). Our failure-function is then identified as the odds-ratio function in the work of Hylek
et al., which they estimated simply as the quotient of unspecified non-parametric estimators of the case- and control-densities, respectively. The contributions of our work to the biomedical problem of Hylek et al. are: (i) the introduction of a class of parametric and semi-parametric models for the failure-function, and corresponding statistical methods, and (ii) the introduction of the expected failure-time as a measure of risk and the role of the failure-function in its computation.

In Section 6, we use the empirical histograms of log INR for cases and controls in the paper of Hylek et al. [3] to fit a normal control density and to estimate the class frequencies and moments of the case-density. These are used to illustrate our proposed statistical analysis.

2. The stochastic model

Let \( p(x) \) be a non-negative measurable function on \( \mathbb{R}^1 \). For \( u > 0 \), put \( Q(x) = up(x) \), and write \( T = T_u \) in (1.1),

\[
P(T_u \leq t | X(s), -\infty < s < \infty) = F \left( u \int_0^t p(X(s)) \, ds \right).
\]

Assume that \( F \) has a continuous bounded density \( F' \). Let \( X \) have the distribution (1.2), and assume

\[ Ep(X) < \infty. \]

From (2.1) we obtain

\[
\frac{d}{dt} P(T_u \leq t | X(s), -\infty < s < \infty) = up(X(t))F' \left( u \int_0^t p(X(s)) \, ds \right),
\]

and, for every Borel set \( B \),

\[
P(T_u \leq s, X(t) \in B) = E \left\{ 1_B(X(t))F \left( u \int_0^s p(X(r)) \, dr \right) \right\}.
\]

It follows that

\[
P(X(t) \in B | T_u = s) = \frac{E \{ 1_B(X(t))p(X(s))F'(u \int_0^s p(X(r)) \, dr) \}}{E \{ p(X(s))F'(u \int_0^s p(X(r)) \, dr) \}}.
\]

From the latter we obtain

\[
P(X(T_u) \in B, T_u \leq t) = \int_0^t P(X(s) \in B | T_u = s) \, dP(T_u \leq s)
\]

\[
= \int_0^t E \left\{ 1_B(X(s))up(X(s))F' \left( u \int_0^s p(X(r)) \, dr \right) \right\} \, ds.
\]

**Theorem 2.1.** Let \( X(t), -\infty < t < \infty, \) be a real periodically stationary ergodic stochastic process. For \( u > 0 \), let \( T_u \) be a positive random variable satisfying (2.1), where \( F \) is a distribution with support on \([0, \infty)\) having a continuous bounded density \( F' \), and where \( p \) is a non-negative continuous function satisfying (2.2). For \( u \to 0 \), the pair \((X(T_u), uT_u)\) has a limiting joint distribution equal to the product of marginal distributions.
\[
\lim_{u \to 0} P(X(T_u) \leq x) = \frac{\int_{-\infty}^{x} p(y) \, dG(y)}{Ep(X)}
\]  
(2.7)

and
\[
\lim_{u \to 0} P(uT \leq t) = F(tEp(X)),
\]  
(2.8)

at all continuity points \( x \) and \( t > 0 \), respectively.

**Proof.** By (2.6), \( P(X(T_u) \in B, uT_u \leq t) \) is equal to the expected value of
\[
u \int_{0}^{\tau/u} 1_B(X(s))p(X(s))F'\left(u \int_{0}^{s} p(X(r)) \, dr\right) \, ds.
\]  
(2.9)

Without loss of generality we take \( B \) to be a bounded Borel set, so that
\[g(x) = 1_B(x)p(x)
\]
is bounded because \( p \) is locally bounded.

For any non-negative measurable function \( p \) such that \( Ep(X) < \infty \), we have
\[
\lim_{n \to \infty} \frac{1}{n} \int_{0}^{n} p(X(s)) \, ds = Ep(X)
\]
with probability 1 because \( \int_{n-1}^{n} p(X(s)) \, ds \), \( n = 0, \pm 1, \ldots \), is an ergodic stationary sequence. This result extends immediately to the limit where the discrete index \( n \) is replaced by the real index \( t \), with \( t \to \infty \).

Since, the integrand in (2.9) is bounded by \( C = \sup_B p(x) \cdot \max_x F'(x) \), the limit of (2.9) is altered, for any \( \varepsilon > 0 \), by at most \( \varepsilon C \) if the lower limit of integration is replaced by \( \varepsilon/u \). The resulting integral, with lower limit \( \varepsilon/u \), is equal to
\[
u \int_{\varepsilon/u}^{\tau/u} g(X(s))F'(usEp(X)) \, ds
\]  
(2.10)

plus a term that is at most equal to
\[(t - \varepsilon) \sup g(x) \cdot \sup_{s \geq \varepsilon/u, y \leq y \leq t} \left| F'(y) \left(1/s \int_{0}^{s} p(X(r)) \, dr\right) - F'(yEp(X)) \right|.
\]

By the continuity of \( F' \) and the ergodic theorem for periodically stationary processes, the latter expression converges to 0 with probability 1, for \( u \to 0 \).

We now take the limit of (2.10). Since \( \varepsilon > 0 \) was taken arbitrarily small, the lower limit \( \varepsilon/u \) may, by the previous argument, be changed back to 0. For arbitrary \( k > 1 \), the resulting expression is equal to
\[
u \sum_{j=1}^{k} \int_{(j-1)\varepsilon/u}^{j\varepsilon/u} g(X(s))F'(usEp(X)) \, ds,
\]
which is at most equal to
\[
\sum_{j=1}^{k} u \int_{(j-1)t/k}^{j/t/k} g(X(s)) \, ds \cdot \max_{(j-1)t/k \leq y \leq j/t/k} F'(yEp(X)).
\]

By the ergodic theorem, the latter expression has the limit, with probability 1,

\[
Eg(X)(t/k) \sum_{j=1}^{k} \max_{(j-1)t/k \leq y \leq j/t/k} F'(yEp(X)).
\]

Since \(k\) is arbitrary, we let \(k \to \infty\) in the latter expression and, by the continuity of \(F'\), we obtain

\[
Eg(X) \int_{0}^{t} F'(yEp(X)) \, dy.
\]

or, equivalently,

\[
\frac{Eg(X)}{Ep(X)} F(tEp(X)).
\]

This is the limit of the asymptotic upper bound of (2.9). By similar reasoning, with the maximum over \([(j-1)t/k, j/t/k]\) replaced by the minimum, we obtain the same expression for the asymptotic lower bound of (2.9). Thus (2.9) has the (constant) limit \(Eg(X)F(tEp(X))/Ep(X)\), with probability 1. The same limit holds for the expected value because (2.9) is bounded by a fixed constant. \(\square\)

Suppose that \(\int_{0}^{\infty} t \, dF(t) < \infty\); then the mean of the limiting distribution of \(uT_u\) in (2.8) is

\[
\int_{0}^{\infty} t \, dF(t)/Ep(X).
\]

We will refer to (2.11) as the asymptotic expected failure-time.

As we stated in Section 1, the observed data consist of a sample of \(n\) independent copies of \(X(T_u)\) for small \(u\). Letting \(u \to 0\), we replace the distribution of \(X(T_u)\) by its limit (2.7), and consider the sample to be drawn from the latter distribution. For simplicity, we will refer to \(X(T)\) as a random variable having the distribution given by the limit (2.7).

We think of the distribution of \(X(T)\) as an extreme-value distribution (for minima) associated with \(X(\cdot)\) through the failure-rate function \(p\) when \(p\) is non-increasing. It is intuitively clear that the observed values of \(X(T)\) tend to be smaller than the values of \(X(t)\) for any fixed \(t \geq 0\) because the process is more likely to fail at smaller values of \(X(\cdot)\). This is rigorously confirmed by the fact that \(X(T)\) is stochastically smaller than \(X(t)\), for any fixed \(t \geq 0\). Indeed, as a weighted average of the function \(p(x), -\infty < x < \infty\), the ratio

\[
\int_{x}^{\infty} p(y) \, dP(X(t) \leq y) / P(X(t) \leq x)
\]

is non-increasing in \(x\); hence, it is at least equal to its limit for \(x \to \infty\):

\[
\int_{-\infty}^{\infty} p(y) \, dP(X(t) \leq y) \leq \frac{\int_{-\infty}^{x} p(y) \, dP(X(t) \leq y)}{P(X(t) \leq x)}.
\]

This is equivalent to
\[ P(X(t) \leq x) \leq \frac{\int_{-\infty}^{x} p(y) \, dP(X(t) \leq y)}{EP(X)}, \]

or

\[ P(X(t) \leq x) \leq P(X(T) \leq x). \]

Suppose, in addition to (2.2), we assume

\[ EP^2(X) < 1; \]

that is,

\[ \int_{-\infty}^{\infty} p^2(x) \, dG(x) < \infty. \]

Let \( h_n(x), n = 0, 1, \ldots, \) be an orthonormal basis for \( L_2(dP). \) Then \( p(x) \) has the \( L_2 \)-expansion

\[ \sum_n h_n(x) \int_{-\infty}^{\infty} h_n(y)p(y) \, dG(y), \]

and so

\[ \frac{p(x)}{EP(X)} = \sum_{n=0}^{\infty} h_n(x)Eh_n(X(T)). \]  

(2.13)

3. Statistical results: the semi-parametric approach

In our main application and in most of our examples, we take \( G(x) \) as \( \Phi(x) \), the standard normal distribution, so that \( X(T) \) has the density function,

\[ \frac{p(x)\phi(x)}{\int_{-\infty}^{\infty} p(y)\phi(y) \, dy}, \]

where \( \phi(x) = \Phi'(x) \). The only assumptions on \( p(x) \) are non-negativity, continuity, and non-increase. The function \( p(x) \) is statistically identifiable on the basis of observations from the density (3.1) up to a constant positive multiple of \( p(x) \): For any \( c > 0 \), (3.1) is obviously unchanged if \( p(x) \) is replaced by \( cp(x) \).

The density (3.1) has the formal appearance of a posterior density in a Bayesian context, where \( p \) is the prior density. This might suggest the potential use of a Bayesian method to estimate the prior from the posterior density. The problem with this is that \( p \), as a non-increasing non-negative function, is the density of a \( \sigma \)-finite infinite measure on the line, and so is not a proper prior density.

A typical density of the form (3.1) is where \( p(x) \) is constant for \( x \leq b_1 \), constant for \( x \geq b_2 \), linear for \( b_1 \leq x \leq b_2 \) and continuous for all \( x \), where \( b_1 < b_2 < 0 \). (See Example 3.1.) Fig. 1 displays the graphs of the density (3.1) and the standard normal density in the particular case \( b_1 = -2, b_2 = -1, p(b_1) = 25, \) and \( p(b_2) = 1 \).

Suppose now that \( p \) is an arbitrary non-negative Borel function satisfying \( \int_{-\infty}^{\infty} p^2(x)\phi(x) \, dx < \infty \) and let \( H_n(x), n = 0, 1, \ldots \) be the classical Hermite polynomials defined by \( (d/dx)^n\phi(x) = (-1)^nH_n(x)\phi(x), n = 0, 1, \ldots \) (see [12, p. 133]) Then, by (2.13),
If $Y_1, \ldots, Y_n$ are independent copies of $X(T)$, then the right-hand member of (3.2) can be estimated by using the sample Hermite moments $(1/n) \sum_{j=1}^n H_i(Y_j)$ in the place of true Hermite moments $E H_i(X(T))$.

A major aim of the research in the medical application, described in Section 5, is achievable by considering the asymptotic expected failure time (2.11) as a functional of the underlying distribution function $G$. For a given distribution function $F$ (in (1.1)) and failure-function $p(x)$, we introduce the family of normal distributions with means $t$ and unit variance, $\Phi(x - t)$, in the role of $G(x)$, and write the corresponding asymptotic expected failure time (2.11) as

$$\frac{\int_0^\infty s dF(s)}{\int_{-\infty}^\infty p(x) \phi(x - t) \, dx}. (3.3)$$

This is increasing in $t$ because $p(x)$ is non-increasing. Put

$$\beta(t) = \int_{-\infty}^\infty p(x) \phi(x - t) \, dx; (3.4)$$

then, by the elementary relation $\phi(x - t) = \phi(x) \exp(tx - \frac{1}{2}t^2)$ and Formula (12.6.7) in [12], we obtain the representations

$$\beta(t)/\beta(0) = e^{-t^2}E(e^{\alpha(T)}), (3.5)$$

$$\beta(t)/\beta(0) = \sum_{n=0}^\infty \frac{t^n}{n!} E H_n(X(T)). (3.6)$$
These suggest two potential estimation methods for the function $\beta(t)/\beta(0)$: (3.5) suggests using the sample moment generating function $(1/n) \sum_{j=1}^{n} \exp(tY_j)$ and (3.6) suggests using the sample Hermite moments.

Next we consider the estimation of the function $p(x)$ in the particular case where the failure-rate has a threshold point $b_1$ and a saturation point $b_2$, with $b_1 < b_2$. This means that the failure-rate is a constant equal to $p(b_1)$ for all $x \leq b_1$, a constant equal to $p(b_2)$ for all $x \geq b_2$, and is monotonic and satisfies $p(b_2) < p(x) < p(b_1)$ for all $b_1 < x < b_2$. We present a semi-parametric method for the estimation of the unknown constants $b_1, b_2$, and $p(b_1)/p(b_2)$.

We begin with $b_1$ and $b_2$. Let $y_k$, $k = 0, 1, \ldots, m$, be an increasing set of real numbers. Then, for any $k$, $y_k \leq b_1 < y_{k+1}$ if and only if
\[ p(-\infty) = p(y_0) = \cdots = p(y_k) > p(y_{k+1}). \]
Similarly, $y_k < b_2 \leq y_{k+1}$ if and only if
\[ p(\infty) = p(y_m) = \cdots = p(y_{k+1}) < p(y_k). \]

Put
\[ P_k = \frac{\int_{y_k}^{y_{k+1}} p(x)\phi(x) \, dx}{\int_{y_k}^{y_{k+1}} \phi(x) \, dx}; \quad (3.7) \]
then $p(y_{k+1}) \leq P_k \leq p(y_k)$, and it follows that $y_k \leq b_1 < y_{k+1}$ if and only if $p(-\infty) = P_0 = \cdots = P_{k-1} > P_k$ and $y_k < b_2 \leq y_{k+1}$ if and only if $p(\infty) = P_m = \cdots = P_{k+1} < P_k$. Since $p(x)$ is non-increasing, it follows that $y_k \leq b_1 < y_{k+1}$ if and only if $k$ is the smallest integer for which
\[ \min_{0 \leq j \leq k-1} P_j > \max_{1 \leq j \leq m} P_j, \quad (3.8) \]
and $y_k < b_2 \leq y_{k+1}$ if and only if $k$ is the largest integer for which
\[ \max_{k+1 \leq j \leq m} P_j < \min_{0 \leq j \leq k} P_j. \quad (3.9) \]

Inequalities (3.8) and (3.9) are obviously undisturbed if the numbers $(P_j)$ are multiplied by a common positive constant. In particular, if $P_k$ is divided by $\int_{-\infty}^{\infty} p(x)\phi(x) \, dx$, then it becomes
\[ Q_k = \frac{p(y_k Y \leq y_{k+1})}{\int_{y_k}^{y_{k+1}} \phi(x) \, dx}, \quad (3.10) \]
where $Y$ has the density (3.1). It follows that the integers $k$ defined by (3.8) and (3.9) are the same as when $(P_j)$ is replaced by $(Q_j)$.

The non-parametric estimators that we propose for $b_1$ and $b_2$ are the interval estimators obtained by replacing $(P_j)$ by the sample estimators of $(Q_j)$, and then picking the smallest and largest integers $k$ for which (3.8) and (3.9), respectively, hold. Put
\[ \hat{Q}_j = \frac{(1/n) \sum_{h=1}^{n} 1[y_j < Y_h \leq y_{j+1}]}{\int_{y_j}^{y_{j+1}} \phi(x) \, dx}; \quad (3.11) \]
and

- \text{...}

- \text{...}
\[\gamma = \min \left\{ k : 1 \leq k \leq m, \min_{0 \leq j \leq k-1} \hat{Q}_j > \max_{k \leq j \leq m} \hat{Q}_j \right\}\] (3.12)

and

\[\delta = \max \left\{ k : 0 \leq k \leq m - 1, \max_{k+1 \leq j \leq m} \hat{Q}_j < \min_{0 \leq j < k} \hat{Q}_j \right\}.\] (3.13)

Then we take \([y_{\gamma}, y_{\gamma+1}]\) and \([y_{\delta}, y_{\delta+1}]\) as estimators of the intervals containing \(b_1\) and \(b_2\), respectively. In the assumed absence of additional information about the parameters, we arbitrarily define their point estimators as the midpoints of their interval estimators:

\[\hat{b}_1 = \frac{1}{2} (y_{\gamma} + y_{\gamma+1}), \quad \hat{b}_2 = \frac{1}{2} (y_{\delta} + y_{\delta+1}).\] (3.14)

Next we estimate \(p(b_1)/p(b_2)\). Since, by definition, \(y_{\gamma} \leq b_1\) and \(b_2 \leq y_{\delta+1}\), it follows that

\[P(Y \leq y_{\gamma}) = p(b_1) \Phi(y_{\gamma}),\]
\[P(Y \geq y_{\delta+1}) = p(b_2)[1 - \Phi(y_{\delta+1})]\]

and so

\[\frac{p(b_1)}{p(b_2)} = \frac{P(Y \leq y_{\gamma})[1 - \Phi(y_{\delta+1})]}{P(Y \geq y_{\delta+1}) \Phi(y_{\gamma})}.\]

Put \(F_n(x) = (1/n) \sum_{j+1}^n 1[Y_j \leq x]\); then \(P(Y \leq y_{\gamma})/P(Y \geq y_{\delta+1})\) has the consistent estimator \(F_n(y_{\gamma})/1 - F_n(y_{\delta+1})\), and so

\[\frac{F_n(y_{\gamma})}{1 - F_n(y_{\delta+1})} = \frac{1 - \Phi(y_{\delta+1})}{\Phi(y_{\gamma})}\]

is a consistent estimator of \(p(b_1)/p(b_2)\).

**Example 3.1.** Suppose that \(p(x)\) is linear on \([b_1, b_2]: p(x) = Ax + B, b_1 \leq x \leq b_2\), with \(A < 0\), and \(Ab_2 + B > 0\). Then, by integration we find that the density (3.1) takes the form

\[\frac{b_1 1_{[x \leq b_1]} + x 1_{[b_1 \leq x < b_2]} + b_2 1_{[x \geq b_2]} + r}{b_2 + \int_{b_1}^{b_2} \Phi(u) \, du + r} \phi(x),\]

where \(r = B/A\) and \(b_2 + r < 0\), and where we have used the integration formula

\[\int \phi(x) \, dx = x \Phi(x) + \phi(x) + C.\] (3.17)

The parameters in (3.16) are \(b_1, b_2\) and \(r\). The first two are estimated by the method just described, and \(r\) is estimated by noting that

\[r = \frac{b_1 - b_2 p(b_1)/p(b_2)}{p(b_1)/p(b_2) - 1},\] (3.18)

and then using the estimator of \(p(b_1)/p(b_2)\).

The function \(\beta(t)/\beta(0)\) (see (3.4)) is, in the particular case (3.16), equal to
\[
\frac{b_2 + \int_{b_1}^{b_2} \Phi(u - t) \, du + r}{b_2 + \int_{b_1}^{b_2} \Phi(u) \, du + r}.
\]  

(3.19)

Example 3.2. Consider the function \( p(x) \) that is exponential and decreasing on \([b_1, b_2] \):

\[
p(x) = \exp \left\{ - c \left[ b_1 1_{x \leq b_1} + x 1_{b_1 < x < b_2} + b_2 1_{x \geq b_2} \right] \right\},
\]

where \( c > 0 \). Then \( p(b_1)/p(b_2) = \exp \left[ c(b_2 - b_1) \right] \), and so

\[
c = \frac{\log p(b_1)/p(b_2)}{b_2 - b_1}.
\]

(3.20)

The parameters \( b_1 \) and \( b_2 \) are estimated by the described method, and then \( c \) is estimated in the form (3.21) by using the estimator of \( p(b_1)/p(b_2) \). The function \( \beta(t) \) in (3.4) takes the form

\[
e^{-cb_1} \Phi(b_1 - t) + \exp \left( -ct + \frac{1}{2}c^2 \right) \left[ \Phi(b_2 + c - t) - \Phi(b_1 + c - t) \right] + e^{-cb_2} \Phi(t - b_2).
\]

(3.21)

4. Statistical results: the parametric approach

In this section we assume that \( p(x) \) is of the form \( p(x; \theta) \), where the latter is a known function with an unknown real or vector parameter \( \theta \). The density (3.1) is of the form

\[
\frac{p(x; \theta) \phi(x)}{\int_{-\infty}^{\infty} p(y; \theta) \phi(y) \, dy}.
\]

(4.1)

The immediate difficulty in finding a good estimator of \( \theta \) is that in the cases of interest there do not exist non-trivial sufficient statistics. (See Examples 3.1, 3.2, 4.1–4.3.) The standard procedures in this situation are maximum likelihood estimation and the method of moments. Maximum likelihood is unsuitable. The formal likelihood equation takes the form

\[
\sum_{j=1}^{n} \frac{\partial}{\partial \theta} \log p(Y_j; \theta) - nE \left( \frac{\partial}{\partial \theta} \log p(Y; \theta) \right) = 0.
\]

However, in the important case where there are unknown threshold or saturation points the regularity conditions required for the derivation of the likelihood equation and for the verification of the asymptotic properties of the estimator are not satisfied. Even if one is willing to overlook these serious deficiencies, there is still the problem that the likelihood equation displayed above is not explicitly solvable.

For the reason just stated, I have resorted to the old method of moments, which generally yields consistent and asymptotically normal estimators. Perhaps there exists another non-classical method of estimation that is more efficient than the method of moments in this case; however, I have not found it.

Example 4.1. In this example we present a function \( p(x) \) with a saturation point \( b \) and threshold point \(-\infty\), so that there does not exist a finite threshold:
\[ p(x) = A \min(x, b) + B, \]  
where  
\[ A < 0, \quad Ab + B > 0. \]  
Thus \( p(x) \) is linear and decreasing for \( x \leq b \), and constant for \( x \geq b \). By direct integration we obtain  
\[ \int_{-\infty}^{\infty} p(x) \phi(x) \, dx = Ab + B - A[b \Phi(b) + \phi(b)], \]  
and the density (4.1) takes the form  
\[ \min(x, b) + r \quad \frac{\phi(x)}{b[1 - \Phi(b)] - \phi(b) + r}, \]  
where \( r = B/A \), and \( b + r < 0 \). The function (4.4) represents a two-parameter family of densities with the parameter pair \( (b, r) \), with \( b + r < 0 \).

Next we show how the parameters are expressible in terms of the first two Hermite moments of the density (4.4), denoted \( EH_1(Y) = E(Y) \) and \( EH_2(Y) = E(Y^2) - 1 \). Direct integration yields  
\[ E(Y) = \frac{\Phi(b)}{b[1 - \Phi(b)] - \phi(b) + r} \]  
and  
\[ E(Y^2) - 1 = \frac{-\phi(b)}{b[1 - \Phi(b)] - \phi(b) + r}, \]  
and so  
\[ \frac{-E(Y)}{E(Y^2) - 1} = \frac{\Phi(b)}{\phi(b)}. \]  
We claim: For every \( y > 0 \), the equation  
\[ y = \frac{\Phi(x)}{\phi(x)} \]  
has a unique solution \( x \). Indeed, \( \Phi(x)/\phi(x) \) has the derivative \([\phi(x) + x\Phi(x)]/\phi(x)\), which, by the formula (3.17) is equal to \( \int_{-\infty}^{\infty} \Phi(u) \, du/\phi(x) > 0 \); hence \( \Phi/\phi \) is strictly increasing. Furthermore, \( \lim_{x \to -\infty} \Phi(x)/\phi(x) = 0 \) and \( \lim_{x \to \infty} \Phi(x)/\phi(x) = \infty \); hence, the statement concerning (4.8) holds. It follows from the hypothesis \( b + r < 0 \) that \( E(Y) \) in (4.5) is negative, and so the left-hand member of (4.7) is positive, and, consequently, Eq. (4.7) has a unique solution \( b \). Having determined \( b \), the corresponding value of \( r \) is obtained by solving either (4.5) for \( r \) or (4.6) for \( r \).

Estimation of the parameter pair \( (b, r) \) on the basis of a sample of observations \( Y_1, \ldots, Y_n \) from the density (4.4) can now be done by the method of moments by simply replacing the left-hand members of (4.5) and (4.6) by their sample analogues \( (1/n) \sum_{j=1}^{n} Y_j \) and \( (1/n) \sum_{j=1}^{n} Y_j^2 - 1 \), respectively, and then solving (4.7) for \( b \) and (4.5) or (4.6) for \( r \). The solution of (4.7) is, of course, approximate.
Standard integration yields

\[
\frac{\beta(t)}{\beta(0)} = \frac{r + b - \int_{-\infty}^{b-t} \Phi(u) \, du}{r + b - \int_{-\infty}^{b} \Phi(u) \, du}
\]

for the function \(\beta(t)\) defined by (3.4).

**Example 4.2.** Define

\[
p(x) = \alpha + \beta \Phi\left(\frac{\mu - x}{\sigma}\right),
\]

where \(\alpha\), \(\beta\), and \(\sigma\) are positive constants, and \(\mu\) is a real constant. Then the density (4.1) is with \(\gamma = \alpha/\beta\),

\[
\left[\gamma + \Phi\left(\frac{\mu - \gamma}{\sigma}\right)\right] \phi(x)
\]

\[
\frac{\gamma + \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)}{\gamma + \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)}.
\]

For the proof of (4.11), note that

\[
\int_{-\infty}^{\infty} p(x) \phi(x) \, dx = \alpha + \beta \int_{-\infty}^{\infty} \Phi\left(\frac{\mu - x}{\sigma}\right) \phi(x) \, dx
\]

\[
= \alpha + \beta \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right),
\]

where the second equality is a consequence of the fact that the integral in the middle member is the convolution, evaluated at the point \(\mu\), of the \(N(0,1)\) and the \(N(0,\sigma^2)\) distributions. The density (4.11) has the parameters \(\gamma > 0, \sigma > 0, \text{ and } -\infty < \mu < \infty\).

By a calculation similar to that following (4.11), we find that \(\beta(t)\), defined by (3.4), satisfies

\[
\frac{\beta(t)}{\beta(0)} = \frac{\gamma + \Phi\left(\frac{\mu - \gamma}{\sqrt{1 + \sigma^2}}\right)}{\gamma + \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)}.
\]

In this particular case it follows that

\[
p(x) = \beta\left(\mu - (\mu - x)\sqrt{1 + \sigma^2}/\sigma\right).
\]

By direct but tedious calculations involving integration by parts and completion of the square, we find the first three Hermite moments of the density (4.11):

\[
EH_1(Y) = -\frac{1}{\sqrt{1 + \sigma^2}} \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)/\left[\gamma + \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)\right];
\]

\[
EH_2(Y) = -\frac{\mu}{(1 + \sigma^2)^{3/2}} \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)/\left[\gamma + \Phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right)\right] ;
\]
From (4.13) and (4.14) we obtain
\[ \frac{EH_1(Y)}{EH_2(Y)} = \frac{1 + \sigma^2}{\mu}. \]  
(4.16)

From (4.14) and (4.15) we obtain
\[ \frac{EH_2(Y)}{EH_3(Y)} = \frac{\mu}{\frac{\mu^2}{1+\sigma^2} - 1}, \]  
(4.17)

which, by (4.16), implies
\[ \frac{EH_2(Y)}{EH_3(Y)} = \frac{\mu}{\mu EH_2(Y)/EH_1(Y) - 1}. \]

Finally we obtain \( \gamma \) in terms of \( EH_i(Y), \ i = 1, 2, 3 \) by substituting \( \mu \) and \( \sigma^2 \) in (4.13) and solving the resulting equation for \( \gamma \):
\[ \gamma = \frac{1}{\sqrt{1+\sigma^2}} \phi \left( \frac{\mu}{\sqrt{1+\sigma^2}} \right) - \phi \left( \frac{\mu}{\sqrt{1+\sigma^2}} \right). \]  
(4.20)

The method-of-moments estimators of \( \mu, \sigma^2 \) and \( \gamma \) are obtained by substituting the sample Hermite moments for the theoretical Hermite moments in (4.18)–(4.20).

Example 4.3. Let \( B(t) \) be a distribution function with support \([0, \infty)\) and define
\[ p(x) = \int_{0}^{\infty} e^{-xt - \frac{t^2}{2}} dB(t). \]  
(4.21)

Then
\[ p(x) \phi(x) = \int_{0}^{\infty} \phi(x + t) dB(t), \]
and \( \int_{-\infty}^{\infty} p(x) \phi(x) dx = 1 \); hence the density (4.1) is
\[
\int_0^{\infty} \phi(x + t) \, dB(t),
\]

(4.22)

which is the density of \(\xi - \eta\), where \(\xi\) and \(\eta\) are independent random variables, with \(\xi\) the standard normal, and \(\eta\) having the distribution function \(B\). From the relation \(E H_m(\xi - \eta) = (-1)^m E \eta^m\) (see [13, Lemma 3.1]), it follows that the moments of \(B\), when they exist, can be estimated by the corresponding sample Hermite moments of the density (4.22).

A standard but long calculation shows that \(\beta(t)\) in (3.4) is, when \(p(x)\) is given by (4.21), represented as

\[
\beta(t) = \int_0^{\infty} e^{-tu} \, dB(u).
\]

5. Application to a case-control study

The paper of Hylek et al. [3] describes the investigation to determine safe and effective blood-levels of the anticoagulant warfarin for the prevention of stroke in a high risk population of patients. This medication is taken daily in a regular dose. The effective blood-level of the patient, called ‘prothrombin-time’, is measured by periodic blood-sample tests and is expressed numerically as the 'International Normalized Ratio' or INR. By its increasing the clotting-time of the patient’s blood, warfarin reduces the likelihood of the formation of an internal clot and the corresponding chance of stroke. This medication has been shown to be effective and safe at proper INR-levels; however, it is ineffective at improperly low levels and unsafe on account of bleeding problems at improperly high levels.

The challenge in the monitoring of the INR of a patient is the volatility of the INR for a fixed dose of medication: the prothrombin-time is sensitive to fluctuations in diet and other factors, known and unknown. The purpose of the study was to determine a lower-bound on the interval of effective INR-values. It has been established in relatively recent research that patients undertaking anticoagulation rarely suffer strokes [14]. For this reason it is impractical to execute a prospective study of the relation of the INR- level to the risk of stroke, and so the investigators did an historical case-control study. The population of cases consisted of patients with nonrheumatic atrial fibrillation (an established risk-factor for stroke) who had been taking anticoagulation and who had been admitted to a particular medical facility during the period 1989–1994 as a consequence of the occurrence of a stroke. The population of controls consisted of a much larger set of patients who were similar to the case-patients except that they had been admitted at the time to that medical facility for a reason other than stroke. The data collected by the investigators included the INR-levels of the case-patients just following their strokes (at the times of their admission) and the INR-levels of the control-patients at some time during their confinement.

Now we show that INR-level (or log INR-level) is appropriately represented as an ergodic periodically stationary process with a continuous time parameter \(t\). The \(t\)-units are days, and the integers represent the time points at which the patient takes the daily dose of warfarin. Hence the INR-level, as a function of time, is the ‘output-level’ of a stochastic system with periodic ‘inputs’ at integer times where the system has attained equilibrium. Such a system has the property that for any time points \(t_1, \ldots, t_k\), the joint distribution of the corresponding output levels is invariant
under a one-unit shift $t_i \rightarrow t_i + 1$. Similar systems have long been used in electrical engineering, where the inputs are electrons and the outputs are the energy-levels.

A particular stochastic model that has been used in the latter area is also appropriate for the INR-level; it is a variation of the 'shot-noise' model. (See, for example, [15, p. 433].) Let \( \{W_n : n = 0, \pm 1, \ldots\} \) be a sequence of independent identically distributed (i.i.d.) non-negative random variables, and let \( f(t), -\infty < t < \infty, \) be a non-negative measurable function such that \( f(t) = 0 \) for \( t \leq 0 \) and for \( t \geq T_0 \) for some constant \( T_0 > 0. \) Define the stochastic process

\[
X(t) = \sum_{n=-\infty}^{\infty} W_n f(t - n), \quad -\infty < t < \infty.
\]

Here \( W_n f(t - n) \) represents the contribution to the INR-level at time \( t \) due to the dose at time \( n, \) and so \( X(t) \) is the sum of all the contributions. (Here there is a tacit simplifying assumption of the additivity of the contributions.) It is easily verified that \( X(t) \) is periodically stationary. Ergodicity is a consequence of the boundedness of the support of \( f \) and the independence of \( \{W_n\}. \)

The random variable \( T \) represents the time of occurrence of an ischemic stroke. The time-origin of \( T \) is irrelevant in our analysis because \( T \) is not observed. For the case population, the data for each patient are \( X(T) \); and, for the control population, it is \( X(t) \), where \( t \) is a non-random time-point that depends on the patient.

Now we will rigorously define the control-distribution and the case-distribution in the context of Theorem 2.1. We want the stochastic model to incorporate the fact that failures are known to be very rare. We do this by assuming the failure function to be \( u_p(x) \) and then applying the conclusion of Theorem 2.1 about the limiting distribution of \( (uT_u, X(T_u)) \) for \( u \rightarrow 0 \) to the case where \( u \) is fixed but small.

First we show why the control distribution is identified with the average marginal distribution of the process. For fixed \( t > 0, \) the distribution of \( X(t) \) among patients who have not failed during therapy of duration \( t \) is \( P(X(t) \in B | T_u > t), \) for any Borel set \( B, \) which, by (2.4) with \( s = t, \) is equal to

\[
\frac{E\{1_B(X(t)) \left[ 1 - F(u \int_0^t p(X(s)) \, ds) \right] \}}{E\{1 - F(u \int_0^t p(X(s)) \, ds) \}}
\]

which converges to \( P(X(t) \in B) \) for \( u \rightarrow 0. \) Under the reasonable assumption that the time durations (in days) since the initiation of anticoagulation for the selected controls are homogeneously distributed in calendar time, the distribution of \( X(t) \) for the controls is the average over \( t \) of the marginal distributions. The average exists (because \( P(X(t) \leq x) \) is a periodic function of \( t \)) and is equal to the integral (1.2). For this reason we identify the latter distribution as the control distribution.

The case-distribution is defined as the limit for \( u \rightarrow 0 \) of the distribution of \( X(T_u), \) the actual level measured at the time of failure. If \( G(x) \) is the control distribution function, then it follows from (2.7) that \( X(T_u) \) has, for \( u \rightarrow 0, \) the limiting distribution function,

\[
\int_{-\infty}^{x} p(y) \, dG(y) \bigg/ \int_{-\infty}^{\infty} p(y) \, dG(y), \quad (5.1)
\]

which we identify now as the case-distribution.
We can think of (5.1), the case distribution, as the distribution derived from the control distribution by the operation (5.1) involving the failure function \( p \). If \( G \) has a density function \( g = G' \), then the case-distribution also has a density, and it is of the form

\[
p(x)g(x) \int_{-\infty}^{\infty} p(y)g(y)\,dy.
\]

The ratio of the case-density to the control-density,

\[
p(x) \left/ \int_{-\infty}^{\infty} p(y)g(y)\,dy, \right.
\]

is known as the ‘odds-ratio’ in case-control analysis. Thus, the odds-ratio is equivalent to the failure-rate \( p(x) \) up to a constant positive multiple. Hylek et al. estimated the odds-ratio by fitting non-parametric density estimators to the case and control densities, and then numerically computing their ratio. On the basis of the sudden decrease in the estimated value of \( p(x) \) for \( 1.7 \leq \text{INR} \leq 2.0 \), they concluded that the lower bound on effective INR-values is somewhere in that interval. The authors recommended that the INR should be maintained in the interval from 2 to 3.

The primary reason for the interest in the lowest effective INR-level is the avoidance of bleeding problems. Since the INR fluctuates in a seemingly random manner about a targeted level, there is a positive probability that at some times the actual INR will briefly go below the lowest effective level or above the highest safe level. Hence there is an interest in the relation between changes in the target INR-level and corresponding changes in risk of stroke. This relation is not describable directly in terms of changes in the value of \( p(x) \) for target-points \( x \); indeed, a point \( x \) at which failure occurs is generally lower than the target-point. In particular a change in the target from \( x \) to \( x + t \) does not generally imply a change in relative risk of \( (p(x + t) - p(x))/p(x) \).

As a measure of the risk of failure associated with a given target level, we propose the reciprocal of the asymptotic expected failure-time (2.5):

\[
Ep(X) \left/ \int_{0}^{\infty} t\,dF(t). \right.
\]

(A longer time to failure is equivalent to a smaller risk.)

This ratio depends on the target level only through the distribution of \( X \) because \( p \) and \( F \) are functions determined by nature alone. Thus, for a specified target level \( x \), the numerator above is written as \( E_x p(X) \).

Suppose that the addition of a constant \( t \) to the process \( X(.) \) has the effect of changing the target from \( x \) to \( x + t \) but leaving the distribution of the process otherwise unchanged; then \( E_x p(X + t) = E_{x+t} p(X) \), by (1.2). This holds in the application considered here where \( X \) has a normal distribution with a mean equal to the target level. For fixed \( x \), put

\[
\beta(t) = Ep_x(X + t).
\]

Then \( \beta(t)/\beta(0) \) represents the risk for target \( x + t \) relative to the risk for target \( x \). This is the basis of our interest in the function \( \beta(t) \) introduced in (3.4) in the special case where \( X \) is standard normal.

The formula for \( \beta(t) \) rests on the assumption that the process \( X(s) \), representing random fluctuations relative to a fixed target level, is changed by exactly a constant \( t \) when \( t \) units are
added to the target-level. This linearity assumption may be true for small changes; however, larger changes may introduce non-linearity. For example, the volatility of the actual level may increase with the target level; thus the variance of $X(s) + t$ may exceed the variance of $X(s)$ itself. To mathematically describe such a model, one might define a real function $f(x, t)$ such that the process $X(s)$ is transformed into the process $f(X(s), t)$ by the addition of $t$ units to the target level; then the function $E_p(X + t)$ is replaced by $E_p(f(X, t))$, where $f(X, 0) = X$. For example, to represent a change in variance, $f(x, t)$ can be taken to be $q(t)(x + t)$, for some positive increasing function $q$, with $q(0) = 1$. We will not study such possibilities in this paper.

6. Numerical illustrations

The numerical data in the form of class frequencies presented in this section were visually inferred from Fig. 1, parts A and B, of the paper of Hylek et al.; they did not publish the underlying data. The numerical inferences are from histograms of the logarithms of the INR-measurements for 222 controls and 74 cases, respectively. The histogram of the controls appears to represent an underlying normal control distribution, for which we estimated the mean to be 0.840526 and the standard deviation to be 0.368695. We transformed the horizontal axes in their histograms by replacing log INR by the standardized variable

$$x = \frac{\text{log INR} - 0.840526}{0.368695}.$$  \hspace{1cm} (6.1)

With these standardized units, the control density is approximated by a standard normal density. The histogram of the cases was similarly transformed, and the following frequency distribution was obtained:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.06</td>
<td>2</td>
</tr>
<tr>
<td>-2.79</td>
<td>2</td>
</tr>
<tr>
<td>-2.52</td>
<td>4</td>
</tr>
<tr>
<td>-2.26</td>
<td>13</td>
</tr>
<tr>
<td>-1.99</td>
<td>7</td>
</tr>
<tr>
<td>-1.72</td>
<td>9</td>
</tr>
<tr>
<td>-1.45</td>
<td>6</td>
</tr>
<tr>
<td>-1.18</td>
<td>7</td>
</tr>
<tr>
<td>-0.91</td>
<td>7</td>
</tr>
<tr>
<td>-0.65</td>
<td>3</td>
</tr>
<tr>
<td>-0.38</td>
<td>2</td>
</tr>
<tr>
<td>-0.11</td>
<td>3</td>
</tr>
<tr>
<td>0.16</td>
<td>4</td>
</tr>
<tr>
<td>0.43</td>
<td>1</td>
</tr>
<tr>
<td>0.70</td>
<td>3</td>
</tr>
<tr>
<td>0.98</td>
<td>1</td>
</tr>
</tbody>
</table>

**Total** 74
The control and case histograms are in Figs. 2 and 3, respectively.

We estimated the first five moments and Hermite moments of the case-distribution by the standard procedure of replacing the grouped observations in an interval by the midpoint of the interval:

<table>
<thead>
<tr>
<th>Moments</th>
<th>Hermite moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>−1.255</td>
<td>−1.255</td>
</tr>
<tr>
<td>2.525</td>
<td>1.525</td>
</tr>
<tr>
<td>−5.001</td>
<td>−1.236</td>
</tr>
<tr>
<td>11.031</td>
<td>−1.118</td>
</tr>
<tr>
<td>−25.049</td>
<td>−6.137</td>
</tr>
</tbody>
</table>

By the statement following (3.2), our semi-parametric estimator of \( p(x) / \int_{-\infty}^{\infty} p(y) \phi(y) \, dy \) based on the first five sample Hermite moments is

\[
0.097 + 0.128x + 1.044x^2 - 0.716x^3 - 0.047x^4 + 0.051x^5.
\]

(see Fig. 4). Here the linear combination of Hermite polynomials has been expressed as a single polynomial. Similarly, by the method described as a consequence of formula (3.6), the semi-parametric estimator of \( \beta(t)/\beta(0) \) is

\[
1 - 1.255t + 0.762t^2 - 0.206t^3 - 0.047t^4 + 0.051t^5.
\]

(see Fig. 5).

According to (6.1), the variable \( x \) in (6.3) and the variable \( t \) in (6.4) are in units of standardized deviation of log INR from its mean 0.840526. By inverting (6.1) we obtain the value of INR corresponding to a given \( x \) (or \( t \)):

\[
\text{INR} = \exp(0.368695x + 0.840526).
\]

---

Fig. 2. Histogram of controls with standard normal units of log INR.
Fig. 3. Histogram of cases with standard normal units of log INR.

Fig. 4. Estimated failure-function (divided by its value at INR = 2).
Selected INR-values are indicated on the horizontal axes in Figs. 4 and 5.

The INR-value corresponding to the target log INR (0.840526) for the control-distribution is, by (6.5) with $x = 0$, equal to 2.3176. A reduction of 0.5000 to the INR-level 1.8176 yields a reduced target $-0.6591$. The corresponding value of $\beta(t)/\beta(0)$, the risk function, is estimated by substituting $t = -0.6591$ in (6.4): $\beta(-0.6591)/\beta(0) = 2.20202$; hence, the reduction of the target INR by 0.5000 units increases the risk by a factor of 2.2020. Similar calculations may be done for the failure-function $p(x)$; for example the $p$-values for INR equal to 1.7 and 2.0 are 1.1073 and 0.2566, respectively.

Under the assumption that there exist threshold and saturation points $b_1$ and $b_2$, respectively, we find, by the method of Section 3, the interval estimators

$$[-2.26, -1.99] \quad \text{and} \quad (-0.91, -0.65). \quad (6.6)$$

These are based on the following values of $\hat{Q}_j$ defined in (3.11) and calculated from the frequency data (6.2):

<table>
<thead>
<tr>
<th>Interval</th>
<th>$\hat{Q}_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>833</td>
<td></td>
</tr>
<tr>
<td>606</td>
<td></td>
</tr>
<tr>
<td>667</td>
<td></td>
</tr>
<tr>
<td>$-2.26, -1.99$</td>
<td>1140</td>
</tr>
<tr>
<td></td>
<td>361</td>
</tr>
<tr>
<td></td>
<td>292</td>
</tr>
<tr>
<td></td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>112</td>
</tr>
<tr>
<td>$-0.91, -0.65$</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

The corresponding point estimators are

$$\hat{b}_1 = -2.125, \quad \hat{b}_2 = 0.78. \quad (6.7)$$

We apply (3.15) to estimate $p(b_1)/p(b_2)$. From (6.2), we obtain $y_j = -2.26, y_{\delta+1} = -0.65, F_n(y_j) = 8/74, 1 - F_n(y_{\delta+1}) = 17/74$. From this and the standard normal table we obtain

$$\text{Estimate of } p(b_1)/p(b_2) = 29.35. \quad (6.8)$$

From (6.5), we find that the estimated INR threshold and saturation points in (6.7) are 1.059 and 1.738, respectively. If one makes the assumption that the failure-rate function is linear on [1.059,
1.738], then it is an exponential function of log INR on the interval $[-2.125, -0.78]$. Hence, since the data is given as logarithmic, we take $p(x)$ as an exponential function on the latter interval, in accordance with the model of Example 3.2.

Applying (3.21), (6.7) and (6.8), we obtain the estimate

$$\hat{c} = 2.512.$$  \hfill (6.9)
Upon the substitution of the estimated values of $b_1$, $b_2$ and $c$ in (3.20), the integral $\int_{-\infty}^{\infty} p(x)\phi(x)\,dx$ is equal to 16.2564. The integration employs the identity $\phi(x+c) = \phi(x)e^{-xc-\frac{1}{2}c^2}$. By this relation and some elementary calculation, one finds the case-density

$$
\begin{align*}
12.8009\phi(x), & \quad x \leq -2.125, \\
1.4428\phi(x+2.512), & \quad -2.125 \leq x \leq -0.78, \\
0.4364\phi(x), & \quad x \geq -0.78.
\end{align*}
$$

See Fig. 6.

The observed frequencies for the class intervals in (6.2) are compared to the expected frequencies calculated by integration of density (6.10) over the respective intervals. We have added the class interval $x \geq 1.24$ because it has a significant expected frequency.

<table>
<thead>
<tr>
<th>Observed</th>
<th>Expected</th>
<th>Observed</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.46</td>
<td>3</td>
<td>3.04</td>
</tr>
<tr>
<td>2</td>
<td>3.12</td>
<td>2</td>
<td>3.37</td>
</tr>
<tr>
<td>4</td>
<td>5.68</td>
<td>3</td>
<td>3.47</td>
</tr>
<tr>
<td>13</td>
<td>9.87</td>
<td>4</td>
<td>3.32</td>
</tr>
<tr>
<td>7</td>
<td>9.26</td>
<td>1</td>
<td>2.96</td>
</tr>
<tr>
<td>9</td>
<td>7.50</td>
<td>3</td>
<td>2.54</td>
</tr>
<tr>
<td>6</td>
<td>5.64</td>
<td>1</td>
<td>1.81</td>
</tr>
<tr>
<td>7</td>
<td>3.95</td>
<td>0</td>
<td>3.47</td>
</tr>
<tr>
<td>7</td>
<td>2.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The chi-square statistic for goodness-of-fit is $\chi^2 = 18.087$. Since there are 17 classes and three estimated parameters, it is reasonable to take 13 degrees of freedom. The $P$-value of the observed $\chi^2$, 18.087, is approximately 0.15, so that, by conventional standards, the fit is acceptable.

Having shown that the model of Example 3.2 furnishes a satisfactory fit to the data, we found that the other models did not fit as well. Since the latter may be of interest in other applications, we will illustrate the estimation procedure by means of data of this application.

For the model of Example 3.1, we use estimators (6.7) and (6.8) in (3.18) to obtain $\hat{r} = 0.733$.

For the model of Example 4.1, Eq. (4.7) is solved for $b$ after substituting the estimated Hermite moments on the left-hand side to obtain the equation $0.8230 = \Phi(b)/\phi(b)$; the approximate solution is $\hat{b} = -0.61$. After the substitution of $\hat{b}$, and the sample mean for $E(Y)$ in (4.5), the solution for $r$ is $\hat{r} = 0.56$.

For the model of Example 4.2, we obtain $\hat{\mu} = -2.4713$, $\hat{\sigma}^2 = 1.0337$, and $\hat{\gamma} = 0.0081$ from (4.18)–(4.20), respectively.

References


