

MATHEMATIK-ARBEITSPAPIERE

ESTIMATION OF NONPARAMETRIC RISK FUNCTIONS
IN MATCHED CASE-CONTROL STUDIES

ANGELIKA VAN DER LINDE
AND GERHARD OSIUS

PREPRINT No. 49

DECEMBER 1997



MATHEMATIK-ARBEITSPAPIERE

A: MATHEMATISCHE FORSCHUNGSPAPIERE

ESTIMATION OF NONPARAMETRIC RISK FUNCTIONS
IN MATCHED CASE-CONTROL STUDIES

ANGELIKA VAN DER LINDE
AND GERHARD OSIUS

PREPRINT NO. 49

DECEMBER 1997

FACHBEREICH MATHEMATIK UND INFORMATIK
UNIVERSITÄT BREMEN

Bibliothekstraße
D-28359 Bremen
Germany

Estimation of nonparametric risk functions in matched case-control studies

Angelika van der Linde
Gerhard Osius

University of Bremen, Germany

Abstract

In epidemiological studies one is interested in investigating the probability of disease depending on risk factors and in particular in detecting interactions of risk factors. Within the setting of parametric logistic regression interactions can be modeled only in a clumsy and limited way. Modeling the risk function nonparametrically, estimating it e.g. by a smoothing (thin plate) spline is attractive as a more explorative approach. For prospective studies this amounts to smoothing within the framework and distributional assumptions of generalized regression models (for binary observations). Case-control studies as retrospective studies with exposure to risk factors being observed do not immediately fit into this setting. In the special case of one-to-one matched studies however there is an appropriate likelihood again within the range of generalized models. Inferences will be illustrated using simulated and real data.

Keywords: nonparametric risk functions, conditional likelihood, general regression model, splines

1. Introduction

Consider the probability of a disease depending on several risk factors. Investigating this dependency we are particularly interested in the associated odds ratios or the relative risk function. Under the standard assumptions of a logistic model an exponential dependence of the odds ratios on a linear combination of the explanatory risk factors is implied. Alternative parametric models have been suggested and explored (e.g. by Thomas (1981)) especially with respect to their potential for studying interactions of risk factors. In recent years also more flexible models based on nonparametric relative risk functions have been proposed for epidemiological studies, for instance by O'Sullivan (1988). For prospective studies this approach results in nonparametric regression (smoothing) in generalized models. An example is given by O'Sullivan, Yandell and Raynor (1986), a more general introduction to the main statistical ideas and techniques is provided in the book by Green and Silverman (1994).

Here we are concerned with estimating nonparametric relative risk functions given a retrospective sampling scheme, more precisely that of a matched case-control study. A difficulty then arises with the likelihood. A smoother as (pointwise) estimator of a nonparametric regression function is obtained as a penalized likelihood estimator. What is the likelihood to refer to in a matched case-control study? We deal with this problem in section 2 and argue that it can be solved in such a way that smoothing is to be performed again within a binary regression model.

Smoothing implies a trade-off between data fit and smoothness of the estimated function quantified by a so-called smoothing parameter. How to determine the smoothing parameter? Along with the smoother a (pointwise) variance is required. As the smoother is "parameterized" using all design points, that is, the estimate's dimension is equal to the number of matched pairs, the usual asymptotics do not hold within this framework. A brief outline of the smoothing techniques and ideas how to deal with these particular problems are given in section 3.

Having provided the theoretical methods we then present two examples to study performance in section 4. With simulated data we illustrate how well the relative risk function may be discovered. With real data we contribute to an old discussion. We study the probability of lung cancer in dependence of smoking and exposure to asbestos and we explore the interactions of these two risk factors.

2. Model and Likelihood

2.1 Model

Let $Y \in \{0, 1\}$ denote the indicator variable for the disease event whose dependence on a vector $X \in \mathbb{R}^P$ of risk factors is under study, taking an additional vector $Z \in \mathbb{R}^S$ of confounders (matching variables) into account. We allow Z to contain discrete as well as continuous variables among its components and take X to be continuous. Although the joint distribution $\mathcal{L}(Y, X, Z)$ of (Y, X, Z) needs to be considered, our interest focuses on models and estimation of the conditional distribution $\mathcal{L}(Y|X, Z)$ thus leaving the “nuisance distribution” $\mathcal{L}(X, Z)$ arbitrary. To be specific, if $p(Y = y, X = x, Z = z)$ denotes the joint density of (Y, X, Z) at the point (y, x, z) , we aim to model and estimate the conditional probability of disease

$$\pi(x, z) = p(Y = 1|X = x, Z = z) = \frac{p(Y = 1, X = x, Z = z)}{p(X = x, Z = z)} \quad (1)$$

or equivalently the corresponding odds

$$\begin{aligned} \text{odds } \pi(x, z) &= \frac{\pi(x, z)}{1 - \pi(x, z)} \quad (2) \\ &= \frac{p(Y = 1|X = x, Z = z)}{p(Y = 0|X = x, Z = z)} = \frac{p(Y = 1, X = x, Z = z)}{p(Y = 0, X = x, Z = z)}. \end{aligned}$$

For comparison of different values u and v of the risk factor X (leaving the confounder fixed) one refers to the odds ratio

$$\text{OR } (u, v|z) = \frac{\text{odds } \pi(u, z)}{\text{odds } \pi(v, z)} \quad (3)$$

or the log-odds-ratio

$$\log \text{OR } (u, v | z) = \text{logit } \pi(u, z) - \text{logit } \pi(v, z). \quad (4)$$

The odds ratios are important with respect to modeling because the odds ratio function $(u, v) \mapsto \text{OR } (u, v|z)$ carries the complete information about the *association* of X and Y given $Z = z$ in the following sense. The *joint* distribution $\mathcal{L}(X, Y|Z = z)$ of (X, Y) given $Z = z$ is *uniquely* determined by

- the *marginal* distributions $\mathcal{L}(X|Z = z)$ and $\mathcal{L}(Y|Z = z)$ of X and Y given $Z = z$

and

- the odds ratio function $\text{OR}(-, -|z)$

(cf. Osius 1997a). Hence, the odds ratio function carries all information about the *joint* distribution not contained in the *marginal* distributions. The usual approach to model $\pi(x, z)$ is a logistic regression

$$\text{logit } \pi(x, z) = \alpha(z) + \beta^T w(x)$$

where the influence of the matching variable z is given by an unknown function α and the risk factor x enters via an R -dimensional vector $w(x)$ of formal covariate derived from x and an unknown parameter vector $\beta \in \mathbb{R}^R$. If $x = (x_1, x_2)$ has only two components (as in our examples), the choice

$$w(x_1, x_2) = (x_1, x_1^2, x_2, x_2^2, x_1x_2, x_1x_2^2, x_1^2x_2, x_1^2x_2^2)$$

leads to a “second-order model” for which $\text{logit } \pi(x_1, x_2, z)$ is a quadratic function in x_1 and x_2 .

The corresponding log odds ratio for the logistic regression model

$$\log \text{OR}(u, v|z) = \beta^T w(u) - \beta^T w(v)$$

does not depend on the value z .

A more general approach, particularly to the exploration of interactions (between the risk factors) is provided by

$$\text{logit } \pi(x, z) = \alpha(z) + h(x) \tag{5}$$

with an unknown smooth non-parametric risk function h (which has to be estimated). This setup includes in particular all parametric risk functions of the form $h(x) = g_\lambda(\beta^T w(x))$ for which the linear predictor $\beta^T w(x)$ enters through a family g_λ of transformations (including the identity) depending on an additional parameter $\lambda \in \mathbb{R}$ (cf. Breslow and Storer 1985, Moolgavkar and Venzon 1987). The log odds ratio for this model is

$$\log \text{OR}(u, v|z) = h(u) - h(v) \tag{6}$$

and in order to uniquely determine the function h we fix its value at an arbitrary argument u_0 (typically $u_0 = 0$) by a constraint $h(u_0) = h_0$ with given $h_0 \in \mathbb{R}$. Now the models (5) and (6) turn out to be equivalent, since (6) leads to (5) using

$$\alpha(z) = \text{logit } \pi(u_0, z) - h_0.$$

An important feature of this non-parametric model is that the odds ratio in (6) does not depend upon the matching value z , which in turn results from the absence of any interaction between x and z in (5).

So far the model (6) is formulated using the conditional distribution $\mathcal{L}(Y|X, Z)$ (i.e. the probabilities $\pi(x, z)$) and this is appropriate for (prospective) cohort studies, when sampling of Y is conditional on given values of

X and Z . In order to use the same model for (retrospective) matched case-control studies, where sampling of X is conditional on Y and Z we have to reinterpret the odds ratio. Denoting the conditional density of $\mathcal{L}(X|Y, Z)$ by

$$p(X = x | Y = y, Z = z) = \frac{p(Y = y, X = x, Z = z)}{p(Y = y, Z = z)} \quad (7)$$

we can rewrite the “odds ratio of disease” (3) as an “odds ratio of exposure”

$$\begin{aligned} \text{OR}(u, v|z) &= \frac{p(Y = 1, X = u, Z = z)}{p(Y = 0, X = u, Z = z)} \bigg/ \frac{p(Y = 1, X = v, Z = z)}{p(Y = 0, X = v, Z = z)} \quad (8) \\ &= \frac{p(X = u|Y = 1, Z = z)}{p(X = u|Y = 0, Z = z)} \bigg/ \frac{p(X = v|Y = 1, Z = z)}{p(X = v|Y = 0, Z = z)}. \end{aligned}$$

Hence the nonparametric odds ratio model (6) also provides a model for the (density of) the conditional distribution $\mathcal{L}(X|Y, Z)$.

2.2 Sampling of matched pairs and likelihood

A 1 : 1 matched pair consists of the matching variable Z and a pair X_0 and X_1 of risk factors for the control ($Y = 0$) and the case ($Y = 1$) and is typically obtained as follows. First a “case” (X_1, Z) is drawn from the conditional distribution $\mathcal{L}(X, Z|Y = 1)$ and then an independent control X_0 is drawn from the conditional distribution $\mathcal{L}(X|Y = 0, Z)$. Hence, conditional on $Z = z$ the risk factors X_0 and X_1 are independent with $\mathcal{L}(X_y|Z = z) = \mathcal{L}(X|Y = y, Z = z)$ for $y = 0, 1$, leading to

$$\mathcal{L}(X_0, X_1|Z = z) = \mathcal{L}(X|Y = 0, Z = z) \otimes \mathcal{L}(X|Y = 1, Z = z) \quad (9)$$

where \otimes denotes the product for probability measures. Since the distribution of the matching variable Z is not of interest here, we shall base the following analysis on the conditional distribution of (X_0, X_1) given $Z = z$ specified by (9). Note that the same conditional distribution may arise from other sampling procedures for (X_0, X_1, Z) , e.g. if the case X_1 is drawn from the conditional distribution $\mathcal{L}(X|Y = 1, Z = z)$.

Starting with the (conditional) sampling distribution (9) we now additionally condition upon another suitably chosen variable, in order to obtain a conditional sampling distribution which is uniquely determined by the odds ratios (8). To achieve this, we exploit the lexicographic ordering on the \mathbb{R}^P , which will be denoted by $u \leq v$. The pair (X_0, X_1) can now be uniquely described by its order statistics $U = \min(X_0, X_1)$, $V = \max(X_0, X_1)$ and the rank indicator $R = I(X_0 \leq X_1)$. The joint density of R, U and V can be evaluated as (Osious, 1997b):

$$\begin{aligned}
p(R = 0, U = u, V = v) &= I\{u < v\} \cdot p(X_0 = v, X_1 = u | Z = z) \\
p(R = 1, U = u, V = v) &= I\{u \leq v\} \cdot p(X_0 = u, X_1 = v | Z = z).
\end{aligned} \tag{10}$$

We finally obtain the conditional sampling distribution of R as a binomial

$$\mathcal{L}(R|U = u, V = v, Z = z) = B(1, p(u, v, z)).$$

Using (8) – (10) the conditional probability

$$p(u, v, z) = P\{R = 1 | U = u, V = v, Z = z\}$$

is given by

$$\text{logit } p(u, v, z) = \log \text{OR } (v, u|z) \quad \text{for } u \leq v \tag{11}$$

and only depends on the odds ratios. For $u = v$ we get $p(u, v, z) = 1$ and hence in this case the distribution of R contains no information whatsoever.

For $u < v$ we can represent the matched pair as a 2×2 -table and interpret R as a result of a “pseudo experiment”, which chooses between the two tables with

	case	control			case	control	
u	1	0		u	0	1	
v	0	1		v	1	0	
	$R = 0$			$R = 1$			

Although the use of conditional distributions is a well known practice in logistic regression for matched data (cf. Breslow and Day, 1980) our approach still holds if some or all components of X are continuous and explicitly specifies the random vector – namely the order statistic (u, v) of (x_0, x_1) – upon which the sampling distribution is conditioned.

Consider now a sample (X_{0j}, X_{1j}, Z_j) of $j = 1, \dots, J$ of J independent matched pairs, each having a conditional distribution according to (9). Using the above notations (with an additional index j for the j -th pair) and denoting the observed value of R_j by r_j we get the following conditional likelihood

$$\begin{aligned}
L &= \prod_j P\{R_j = r_j | U_j = u_j, V_j = v_j, Z_j = z_j\} \\
&= \prod_j p(u_j, v_j, z_j)^{r_j} \cdot (1 - p(u_j, v_j, z_j))^{1-r_j}
\end{aligned} \tag{12}$$

which represents an ordinary Bernoulli likelihood with respect to J independent “pseudo experiments”. We assume $u_j \neq v_j$ here (i.e. we omit all pairs with $u_j = v_j$), since pairs with $u_j = v_j$ do not contribute to the above likelihood.

Using the non-parametric odds-ratio model (6), the probabilities (11) are completely determined by the corresponding differences of the risk function

$$\text{logit } p(u_j, v_j, z_j) = h(v_j) - h(u_j) \quad \text{for } u_j < v_j . \quad (13)$$

3. Smoothing

Since our applications deal with a two-dimensional risk factor $X = (X_1, X_2)$ we restrict the following discussion for simplicity to the two-dimensional case although it similarly applies to higher dimensions.

Consider the likelihood of the “pseudo-experiment” of observing the maximum exposure $v_j = (v_{j1}, v_{j2})$ (w.r.t. lexicographic ordering) as that of the case in the j -th matched pair. As has been seen in section 2 this is the likelihood of a $B(1, p_j)$ -distributed variable R_j with

$$\text{logit } p_j = h(v_j) - h(u_j) =: \theta_j ,$$

and we assume h to be a smooth function on a two-dimensional domain. A (thin plate) smoothing spline as estimate of h is obtained maximizing the penalized log-likelihood

$$\begin{aligned} & \sum_{j=1}^J l(\theta_j | r_j) - \frac{1}{\gamma} P(\theta) \\ &= \sum_{j=1}^J (\theta_j r_j - \log(1 + \exp \theta_j)) - \frac{1}{2\gamma} \theta' B^{-1} \theta \end{aligned}$$

with $\theta = (\theta_1, \dots, \theta_J)'$. γ denotes the smoothing parameter and $P(\theta) = \frac{1}{2} \theta' B^{-1} \theta$ a roughness penalty term. Specification of B^{-1} results from the following reasoning.

3.1 Derivation of the roughness penalty

We measure roughness $J(h)$ of a function h by an integral over partial second derivatives,

$$J(h) = \iint \frac{\partial^2 h}{\partial x_1^2} + 2 \frac{\partial^2 h}{\partial x_1 \partial x_2} + \frac{\partial^2 h}{\partial x_2^2} dx_1 dx_2 .$$

Using $J(h)$ as penalty in maximum likelihood estimation we find that – the log-likelihood depending on θ only – the estimate is the smoothest function realizing $\theta_j = h(v_j) - h(u_j)$, $j = 1, \dots, J$. This function is not uniquely defined but can be identified by fixing additionally an absolute value of h : $h(0, 0) = h_0$. Thus we can reduce estimation of h to the finite dimensional estimation of θ and replace $J(h)$ by $P(\theta)$, the roughness penalty of the smoothest function realizing θ and h_0 , $P(\theta)$ actually not depending on h_0 . $P(\theta)$ can be given explicitly and computed applying the representations and formulae given in Wahba (1990), ch. 2.4). We end up with

$$B^- = \lim_{\xi \rightarrow \infty} B_\xi^{-1}$$

where

$$B_\xi = \xi \tilde{V}\tilde{V} + \tilde{M},$$

$$\tilde{V} = ((v_{ji} - u_{ji}))_{\substack{j=1, \dots, J \\ i=1, 2}},$$

$$\tilde{M} = ((K(u_k, u_j) - K(v_k, u_j) - K(u_k, v_j) + K(v_k, v_j))_{k, j=1, \dots, J},$$

and K is specified as in (Wahba, 1991, p. 34, 2.4.25).

This yields

$$P(\theta) = \theta'(\tilde{M}^{-1} - \tilde{M}^{-1}\tilde{V}(\tilde{V}'\tilde{M}^{-1}\tilde{V})^{-1}\tilde{V}\tilde{M}^{-1})\theta.$$

The penalized likelihood estimate can also be interpreted as Bayesian estimate (posterior mode) w.r.t. to a prior

$$\theta \sim N(0, \gamma B),$$

roughly speaking. More precisely, due to taking limits $\xi \rightarrow \infty$, we use a partially improper prior according to which the “ultra-smooth” (linear) functions are weighted with a uniform prior and the nonlinear (parts of) functions are the more likely the smoother (in terms of $J(h)$) they are.

For any fixed smoothing parameter γ the estimate $\hat{\theta}(\gamma)$ is obtained using Fisher scoring (for details see Green and Silverman, 1994, ch. 5). γ usually is not known but has to be estimated from the data.

3.2 Determining the smoothing parameter

Most of the criteria for determining the smoothing parameter are based on cross-validation. In Bayesian terms, applying this technique smoothing is an empirical Bayesian procedure as opposed to a fully Bayesian approach where a (hyper-)prior for γ is specified. Here we consider the empirical approach only. Cross-validation means to leave out single observations r_j in turn, to estimate $\theta(\gamma)$ without this observation by $\hat{\theta}^{-j}(\gamma)$ and to use $\hat{\theta}^{-j}(\gamma)$

or $\hat{p}^{-j}(\gamma) = \text{logit}^{-1}(\hat{\theta}^{-j}(\gamma))$ in comparison to $\hat{\theta}_j(\gamma)$ or r_j to compute a function which is then minimized w.r.t. to γ .

- (1) $\hat{\theta}^{-j}(\gamma)$ may be used for predicting the observation r_j by its estimated mean $\hat{p}_j^{-j}(\gamma)$ and comparing to r_j , for example using the Pearson distance. This yields

$$CV_P(\gamma) = \sum_{j=1}^J (r_j - \hat{p}_j^{-1}(\gamma))^2 / \hat{p}_j^{-j}(\gamma)(1 - \hat{p}_j^{-j}(\gamma))$$

to be minimized in γ .

- (2) Alternatively the probabilities $\hat{p}_j^{-1}(\gamma)$ may be predicted and the related measure on $\{0, 1\}$ compared to δ_{r_j} , for example using the quadratic loss function. This yields

$$CV_Q(\gamma) = \sum_{j:r_j=1}^J (1 - \hat{p}_j^{-j}(\gamma))^2 + \sum_{j:r_j=0}^J (\hat{p}_j^{-j}(\gamma))^2$$

to be minimized in γ . (See also Fahrmeir and Tutz, 1994, ch. 5.2.)

- (3) Or the stability of the parameter estimate leaving out single observations may be quantified in

$$CV_R(\gamma) = \frac{1}{J} \sum_{i=1}^J \left[\sum_{j=1}^J \hat{\theta}_j(\gamma) - \hat{\theta}_j^{-i}(\gamma) \right]^2 + \frac{1}{J} \sum_{i=1}^J \sum_{j=1}^J (\hat{\theta}_j(\gamma) - \hat{\theta}_j^{-i}(\gamma))^2.$$

This criterion is called “robustified cross-validation” due to Robinson and Moyeed (1989).

Computing and minimizing these functions may be quite a task because each parameter estimate requires an iteration. In case of normally distributed observations calculations can be simplified referring to a “deletion theorem”. It states that $\hat{\theta}^{-j}(\gamma)$ is obtained in the same way as $\hat{\theta}(\gamma)$, that is multiplying the vector of observations by a “hat matrix” $A(\gamma)$, if the j -th observation is replaced by its estimated mean $\hat{\theta}_j^{-j}(\gamma)$ (see Wahba, 1990, p. 50f for details). Using the deletion theorem $\hat{\theta}_j^{-j}(\gamma)$ can be calculated using a formula, given the observations, $A(\gamma)$ and $\hat{\theta}(\gamma)$. In case of binary observations the deletion theorem does not immediately hold. In order to apply it, however, an approximation of the log-likelihood by Taylor expansion in $\hat{\theta}(\gamma)$ is introduced. This approximation of the log-likelihood (also used in Fisher scoring) yields the “working observations” $z(\gamma) = (z_1(\gamma), \dots, z_J(\gamma))'$ which can be treated

as if normally distributed.

For

$$z(\gamma) = \hat{\theta}(\gamma) + [\text{diag}(\hat{p}_j(\gamma)(1 - \hat{p}_j(\gamma))]^{-1}(r - \hat{p}(\gamma))$$

with $r = (r_1, \dots, r_J)'$, $\hat{p}(\gamma) = (\hat{p}_1(\gamma), \dots, \hat{p}_J(\gamma))$ in obvious notation, the working hypothesis is

$$z(\gamma) \sim N(\theta[\text{diag}(\hat{p}_j(\gamma)(1 - \hat{p}_j(\gamma))]^{-1})$$

and all calculations can be performed as in the normal case. Note that this is a computational technique only, not a theoretically derived result on the distribution of $z(\gamma)$. Cross-validatory criteria “leaving out working observations” are then defined, for example

$$\begin{aligned} CVZ(\gamma) &= \sum_{j=1}^J (z_j(\gamma) - \hat{\theta}_j^{-z_j(\gamma)}(\gamma))^2 \\ &= \sum_{j=1}^J \frac{(z_j(\gamma) - \hat{\theta}_j(\gamma))^2}{1 - a_{jj}(\gamma)^2}, \quad A(\gamma) = ((a_{ij}(\gamma)))_{i,j=1,\dots,J} \end{aligned}$$

due to the deletion theorem.

Generalizing $a_{jj}(\gamma)$ to $\frac{1}{J} \text{tr} A(\gamma)$ the criterion of generalized cross-validation used by O’Sullivan et al (1988) and also referred to by Green and Silverman (1994, p. 109) is obtained. Similarly the criteria given above may be modified using $\hat{\theta}_j^{-z_j(\gamma)}(\gamma)$ resp. $\hat{p}_j^{-z_j(j)}(\gamma)$ instead of $\hat{\theta}_j^{-j}(\gamma)$ resp. $\hat{p}_j^{-j}(\gamma)$. Our estimates for the examples presented in section 4 are based on robustified cross-validation w.r.t. working observations.

The major question of “How do the different criteria perform?” is not yet settled. A small simulation study for binary regression is simultaneously under work and will be published elsewhere (v.d. Linde, 1997a). For our examples we do have and give information about the coherence of results using different criteria in determining the smoothing parameter.

3.3 Precision of estimation

A natural way to obtain a measure of precision for the estimates of θ_j is the Bayesian interpretation of smoothing. For a fixed smoothing parameter γ^* the posterior distribution of θ based on independent Bernoulli-distributed observations and the normal prior (in the limit) may be simulated. Note that the posterior covariance then does not describe the variation centred in the mode $\hat{\theta}(\gamma^*)$ but in the mean $E(\theta|r, \gamma^*)$ which may also be obtained using the simulations. In order to avoid the computational burden of simulation the approximation of the log-likelihood by Taylor-expansion already referred to may be used. Applying this approximation all calculations technically reduce to those of the normal case (cp. Gu, 1992a), applied now

to the working observations $z(\gamma^*)$. The mean of the approximate posterior distribution of θ is equal to its mode and furthermore equal to the mode of exact posterior distribution. Thus the covariance of the approximate posterior distribution can be used to give an idea of the variability in estimating θ . This is what we have used in our examples.

4. Examples

Presenting the examples for two-dimensional risk factors we offer two analyses: smoothing as an approach to nonparametric estimation of the relative risk function in its own right or as an explorative step in data analysis followed by maximum likelihood estimation of a suitably chosen parametric relative risk function based on second-order models. There are two reasons for using explorative smoothing for suggestion of a parametric model. The first one is parsimony in representation of the function implying easier availability and possibly interpretation of odds ratios of interest. The second one is the possibility of testing for submodels (e.g. no interaction or linearity in x_1 or x_2) and checking model fit, which is costly in smoothing and somewhat elaborated for logistic regression models. In particular, any given regression model may be formally tested against a suitably chosen enlarged regression model, which typically contains additional powers of x_1 resp. x_2 and/or products thereof. Furthermore, asymptotic goodness-of-fit measures are available which for the standard goodness-of-fit (like Pearson's statistic) however reduce for binary data again to tests with respect to an enlarged (nonlinear) logistic regression model, (cf. Osius and Rojek 1992, Sec. 3.1).

4.1 Simulated data

We consider the relative risk function on the unit square

$$h(x_1, x_2) = 0.4 \left[\left(1 - \frac{1}{0.1 + x_1}\right)(1 + 2.5 \exp(x_2^2)) + (\exp(x_2))^3 + 4 \right]$$

with $h(0, 0) = -10.6$ which is monotone in both x_1 and x_2 .

Fig. 4.1.1 displays h as a function of exposures on a grid over the unit square. Remember that h is related to the probability of disease by

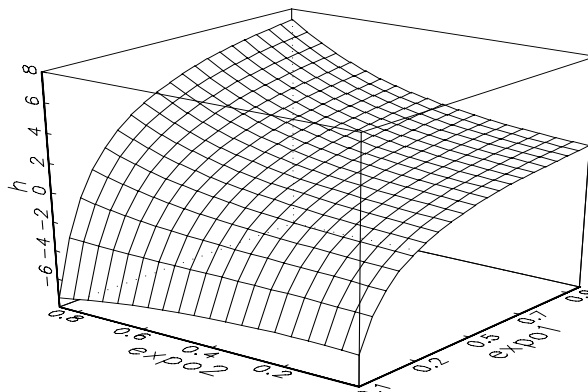
$$\text{logit } \pi(x, z) = \alpha(z) + h(x)$$

and to the probability of observing the maximum exposure as that of the case by

$$\text{logit } p_j = h(v_j) - h(u_j).$$

We generated for 100 matched pairs values of two exposures, expo 1 and expo 2, randomly distributed within the unit square. We then simulated observations r_j according to p_j and assigned the values of exposures correspondingly.

Fig. 4.1.1



The relative risk function h as function of two exposures over the unit square.

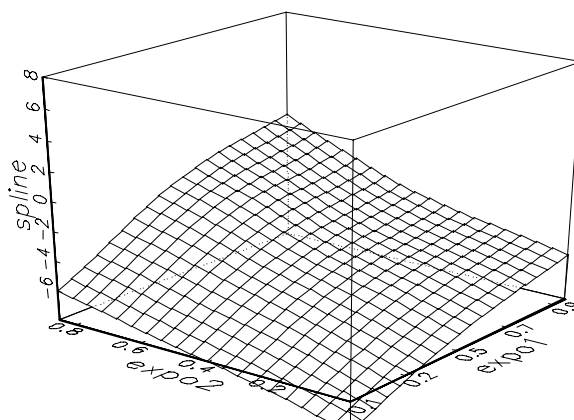
4.1.1 Smoothing

The smoothing parameter resulting from robustified cross-validation

$$\ln \gamma^* = 2.16$$

yields an overall oversmoothed estimate of h although some general features of the curvature of h are met. Fig. 4.1.2 displays the smoothing spline.

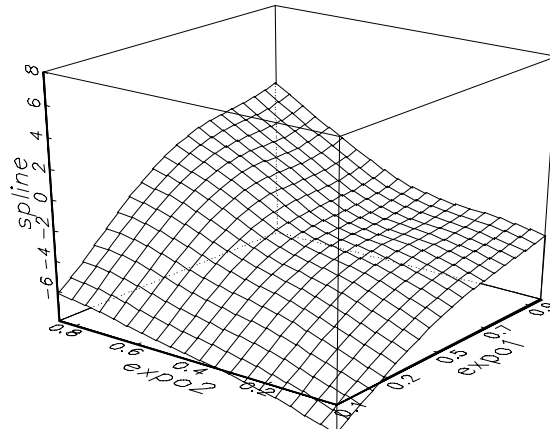
Fig. 4.1.2



The smoothing spline based on the smoothing parameter obtained from robustified cross-validation.

For comparison we also minimized different cross-validatory criteria. Robustified cross-validation for this data set is in agreement with that based on Pearson's distance. Criteria based on comparison of measures of $\{0, 1\}$ suggest less smoothing, $\ln \gamma = 3.4$. The corresponding smoothing spline is shown in Fig. 4.1.3.

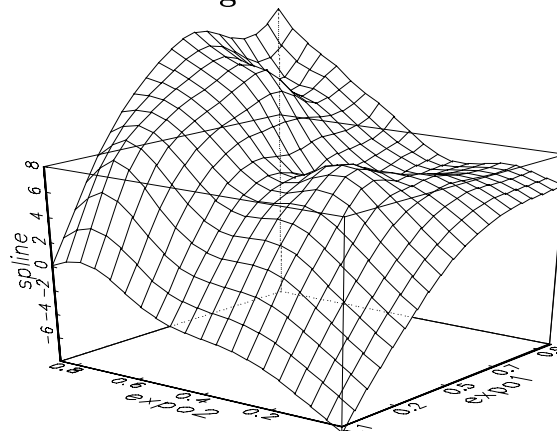
Fig. 4.1.3



The smoothing spline based on the smoothing parameter obtained using the Kullback–Leibler distance as measure of discrepancy between δ_{r_j} and $B(1, \hat{p}_j^{-z_j})$.

Some of the cross-validatory functions did not attain a minimum within the interior of the pre-specified interval $\gamma \in [1, 1000]$ but decreased with the amount of smoothing. One of these was the function based on generalized cross-validation. However, the smoothing parameter corresponding to the right margin of the interval already results in undersmoothing as shown in Fig. 4.1.4.

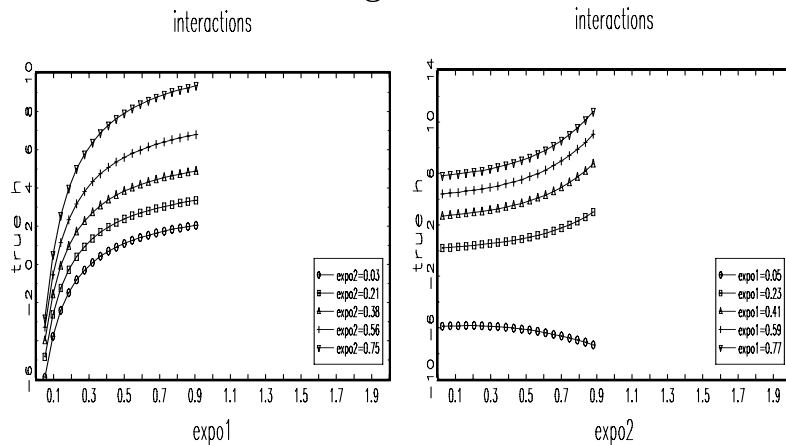
Fig. 4.1.4



The smoothing spline based on the smoothing parameter obtained from generalized cross-validation.

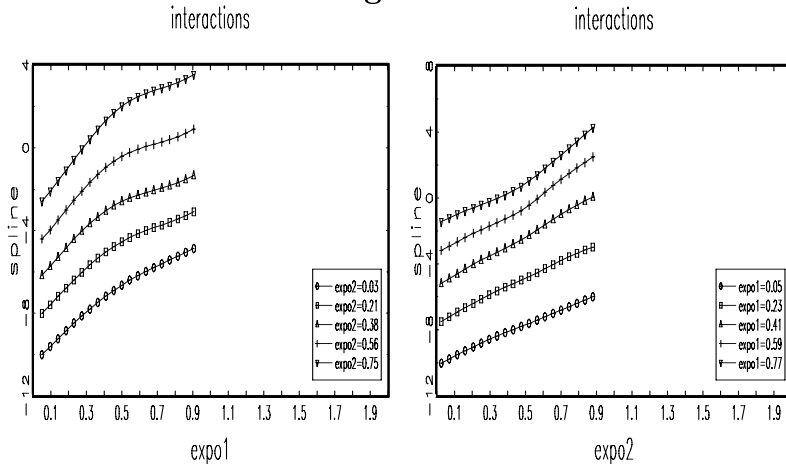
Though we know with simulated data that the smoothing parameter γ^* for this data is not the optimal one we stick to this choice as if we had real data and proceed evaluating the detection of interactions. Interactions can be best visualized displaying “orbits” of h or its estimate and looking for parallelism. Fig. 4.1.5 displays the curves for the true function h , Fig. 4.1.6 those for the smoothing spline based on robustified cross-validation. (The curves are shifted in order avoid overlay, for the true and estimated function to the same amount. Thus the values on the z -axis in Fig. 4.1.5 do not agree with those on the z -axis of Fig. 4.1.1.)

Fig. 4.1.5



h (+ a shift) as a function of one exposure fixing the other exposure at one of five values.

Fig. 4.1.6



The smoothing spline (+ a shift) as a function of one exposure fixing the other exposure at one of five values.

The reproduction of the curvature in the qualitative sense of convexity/concavity is satisfying. The details of acceleration and therefore interac-

tion, however, are lost due to oversmoothing.

4.1.2 Likelihood analysis for logistic regression models

It is instructive to compare the preceding non-parametric analysis with a standard likelihood analysis based on second-order logistic regression models, when the risk function takes the form

$$\begin{aligned}\tilde{h}(x_1, x_2) &= \beta^T w(x_1, x_2) && \text{with} \\ w(x_1, x_2) &= (x_1, x_2, x_1^2, x_2^2, x_1x_2, x_1, x_2, x_1^2x_2, x_1^2x_2^2)\end{aligned}$$

and hence $\tilde{h}(0, 0) = 0$. For suitable β the function \tilde{h} represents a second-order Taylor expansion of h at the point $(0, 0)$ and may be viewed as an approximation to the true risk function h . The above model provides a satisfactory fit for the data. Focussing on the interaction between the two exposures, only two out of four interaction terms turned out to be significant (at the 5% level). Excluding the non-significant interactions $x_1^2x_2$ and $x_1^2x_2^2$ increases the deviance by only 2.60 with 2 d.f. ($p = 27\%$) and leads to a submodel where both remaining interaction terms are significant (at the 5% level). The parameters for this submodel are given in table 4.1.1. And the estimate of the corresponding function

$$\tilde{h}(x_1, x_2) = \beta_{10}x_1 + \beta_{20}x_1^2 + \beta_{01}x_2 + \beta_{02}x_2^2 + \beta_{11}x_1x_2 + \beta_{12}x_1x_2^2$$

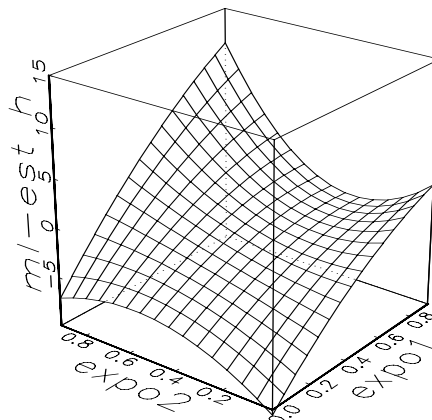
as shown in Fig. 4.1.7 qualitatively resembles the true function h with respect to curvature and convexity/concavity but details are lost due to the restriction to a quadratic function. The same impression is given by the estimated orbits shown in Fig. 4.1.8.

Covariable	Parameter	S.E.	Parameter/S.E.	P-Level
expo1	+21.437113	8.319573	+2.576708	0.9975%
expo1 ²	-7.749918	4.600877	-1.684444	9.2096%
expo2	+17.508224	11.756871	+1.489191	13.6437%
expo2 ²	-14.985954	10.441201	-1.435271	15.1210%
expo1*expo2	-43.271096	20.528977	-2.107806	3.5048%
expo1*expo2 ²	+54.015306	23.167612	+2.331501	1.9727%

Table 4.1.1

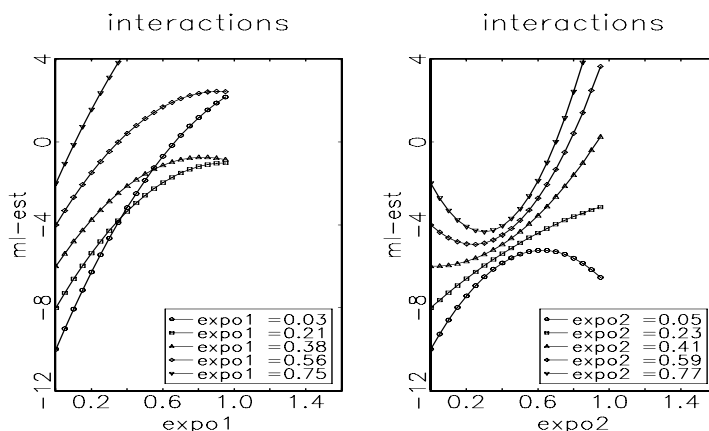
ML-estimates of parameters with standard errors (S.E.) and two-sided p-level for a submodel of the second order logistic regression model (simulated data)

Fig. 4.1.7



ML-estimate \tilde{h} of the relative risk function h based on a quadratic logistic regression model with parameters given in table 4.1.1.

Fig. 4.1.8



ML-estimate \hat{h} (+ a shift) based on table 4.1.1 as a function of one exposure fixing the other exposure at one of five values.

To summarize, the ML analysis based on quadratic logistic regression detects the interaction and its qualitative aspect, but not the quantitative part, due to misspecification of the model. More satisfactory fits may possibly be obtained using a more flexible class of logistic regression models, e.g. cubic models, or/and (log-) transformations of one or both exposures.

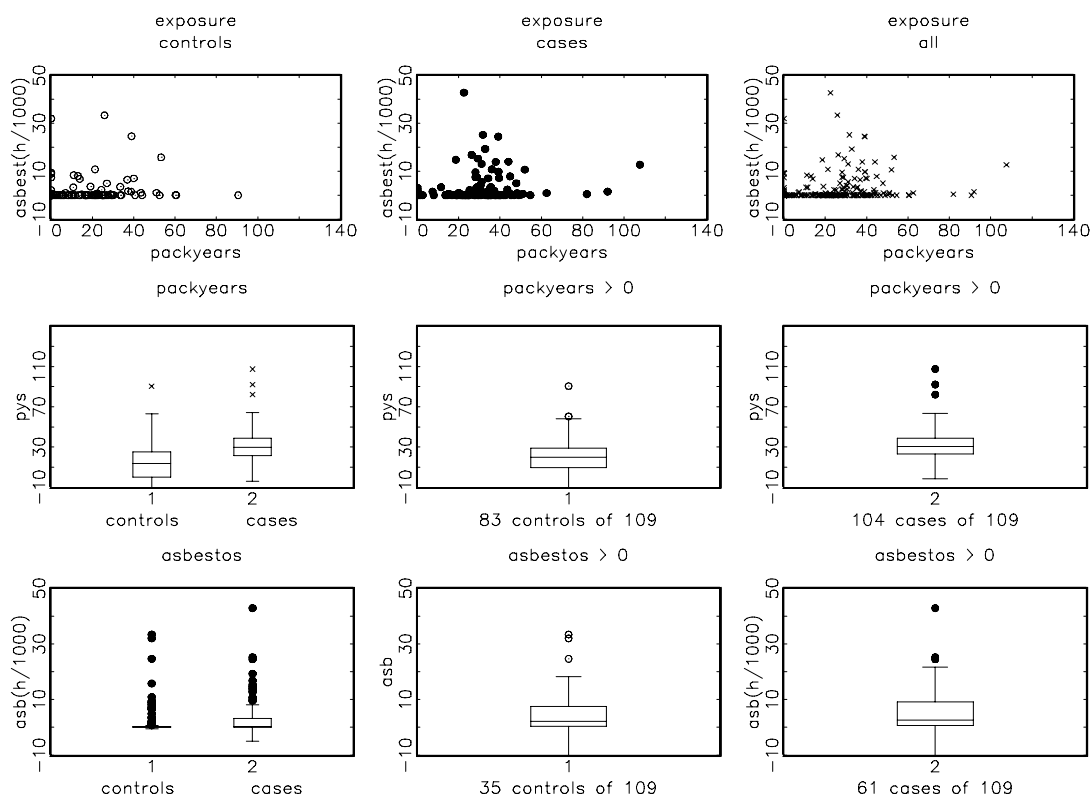
4.2 Real data

Our real data set is part of a more comprehensive case-control study on occupational risk factors for lung cancer. A report on this study is available in German (Joeckel et al., 1995).

Our subanalyses comprise data of 109 pairs of men, less than 50 years old, matched w.r.t. age (in classes of 5 years) and home region. Exposures of interest are smoking, measured in packyears, and exposure to asbestos, measured in life hours and rescaled dividing by 1000. Behind there was a rather detailed questionnaire allowing for in depth investigation of occupational exposures. We are interested in the relative risk function of lung cancer and in particular in the interactions of smoking and exposure to asbestos in constituting the risk of lung cancer.

A first description of the distribution of exposures is given in Fig. 4.2.1. There are a few extreme exposures, packyears > 70 and asbestos > 30 . Thus the estimate of the function beyond these values is supported insufficiently by data. This should be reflected by an increased variance of the estimates. We did perform the analysis excluding these points, too, but the estimated function did not change in that part of the plane where the bulk of data is located. Therefore the analysis we report here is based on all of the data.

Fig. 4.2.1



1. line: Scatterplots of packyears vs asbestos for controls, cases and overall
2. line: Boxplots displaying the distribution of packyears in cases and controls
3. line: Boxplots displaying the distribution of asbestos in cases and controls

4.2.1 Smoothing

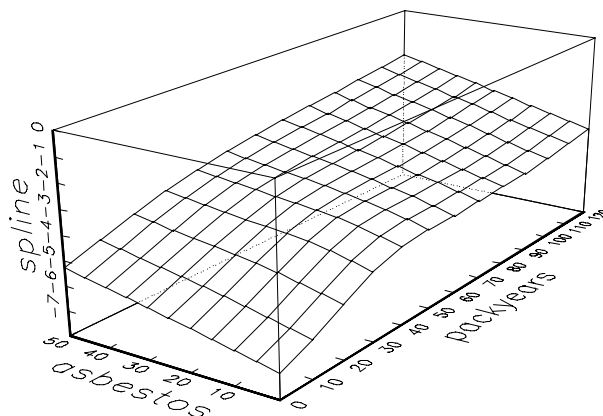
Robustified cross-validation yields the smoothing parameter γ^* with

$$\ln \gamma^* = -7.55 .$$

Based on alternative criteria the smoothing parameter is estimated quite coherently ranging from $\ln \gamma = -7.83$ yielding the smoothest fit to $\ln \gamma = -6.93$ yielding the least smooth fit. The smoothing spline using γ^* and fixing the function with the arbitrary value of -7.0 in $(0,0)$ is displayed in Fig. 4.2.2. The smoothing splines for alternative smoothing parameters look very similar, less smoothing becoming apparent mainly in less linear extension where the data are scarce (packyears > 60).

The most influential point (matched pair) in the data yielding the largest discrepancy in estimating θ when leaving out this point is the one with extreme packyears of 82 for the case compared to 34 for the control and negligible exposure to asbestos for both. Similarly the second most influential pair exhibits 90.5 packyears for the control compared to 11.4 for the case again with no relevant difference in exposure to asbestos. Leaving out these points makes the fitted spline show a “valley” along the packyears-axis at the left out value. Leaving out these points increases the uncertainty about extension of the spline where there are only very few data. This is of no relevance to the main features of the curve.

Fig. 4.2.2

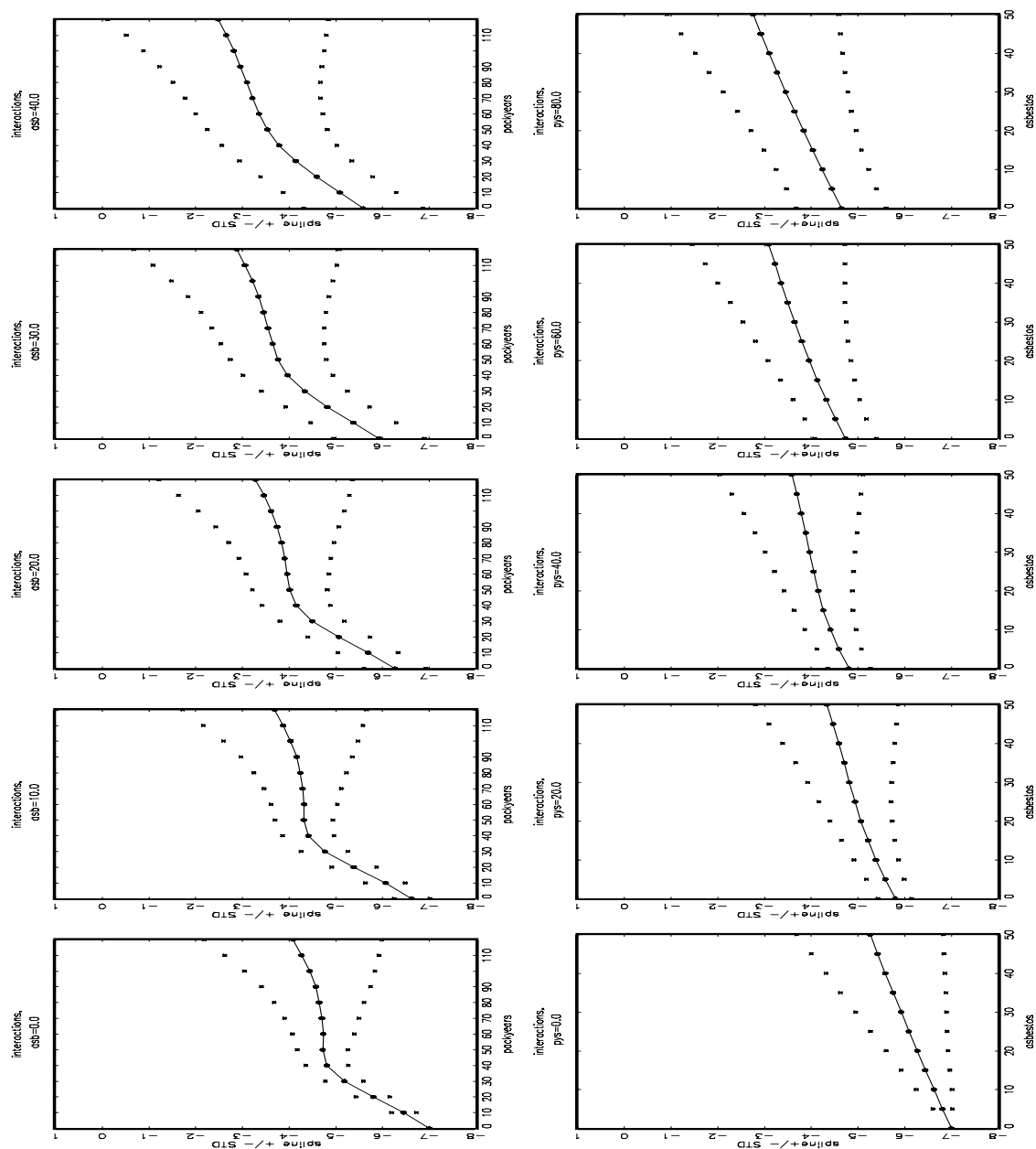


The smoothing spline based on the smoothing parameter obtained from robustified cross-validation.

As to interactions there is an indication that with increasing level of the one exposure the increase in relative risk induced by the other one is slowed down a bit. The smoothing spline along several orbits of packyears and asbestos is displayed in Fig. 4.2.3. In terms of odds ratios the same behaviour is

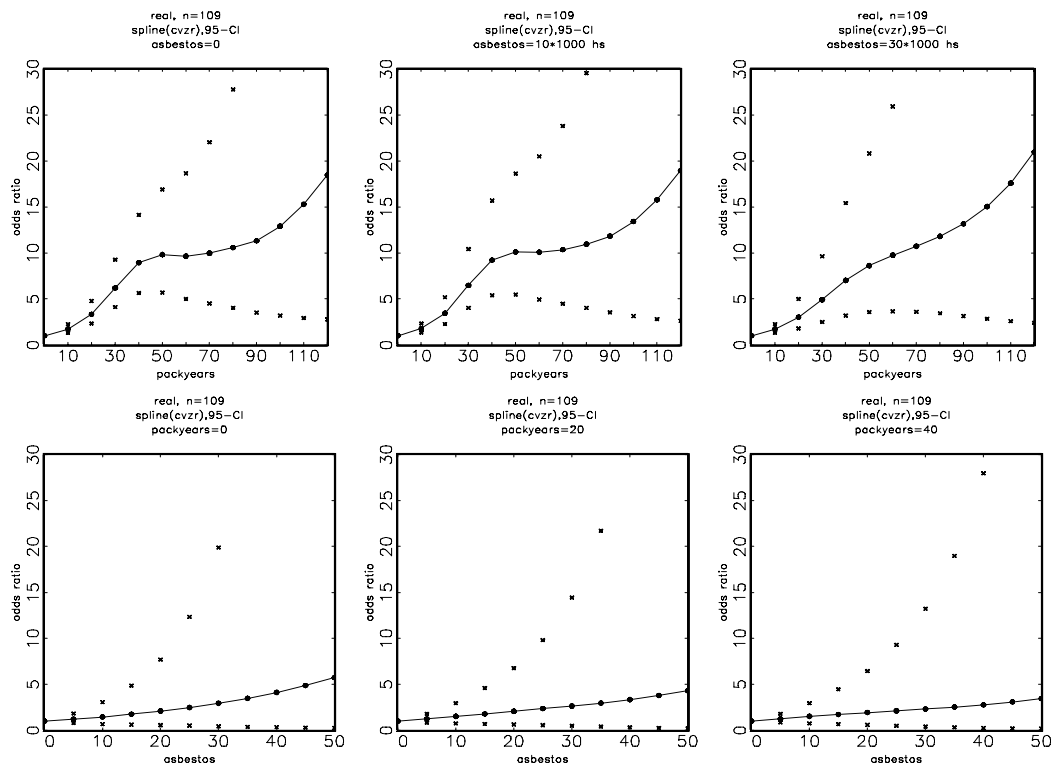
seen in Fig 4.2.4. At a low level of asbestos (for case and control) the odds ratio increases from about 1 for non-smoking to about 10 for smoking at the level of 50 packyears. At a higher level of asbestos an increase from 1 to 8 only occurs. Similar statements hold for the increase of the odds ratio as a function of exposure to asbestos.

Fig 4.2.3



The smoothing spline \pm posterior standard deviation based on robustified cross-validation as a function of one exposure fixing the other one at one of five levels.

Fig 4.2.4

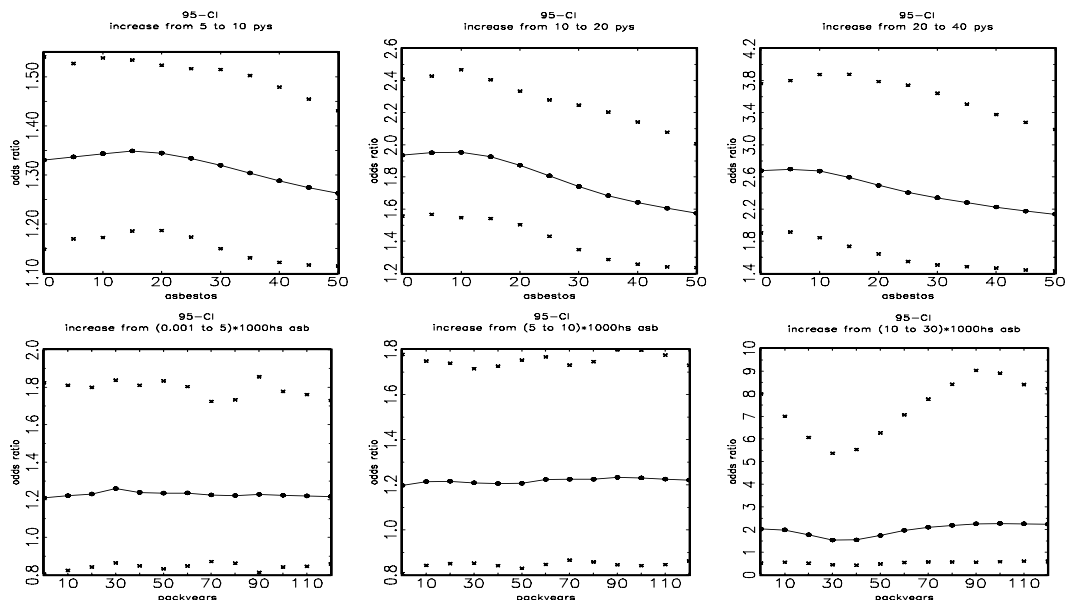


Odds ratios and (exponentially transformed posterior) 95%-central intervals calculated from the smoothing spline based on robustified cross-validation, increasing one exposure and keeping the other one constant at one of three levels.

Another way of looking at these – not very pronounced – interactions is suggested in Fig 4.2.5. Here the odds ratios resulting from a fixed increase in one exposure are studied as a function of the level of the other one. In the simple loglinear model without interactions this would be a constant function, the constant being “the” odds ratio. Up to a level of 10.000 life hours exposure to asbestos the increase in odds ratios caused by an increase in packyears is larger than at levels of exposures to asbestos of more than 10.000 life hours. An effect for the increase in exposure to asbestos is not distinct.

Figures 4.2.4 and 4.2.5 demonstrate the usefulness of the graphical exploration but it is also obvious that for summarizing the data analysis this approach is not very handsome. Therefore it may be desirable to make use of the exploration as suggestion for improved parametric model building. All computations for smoothing were performed using GAUSS.

Fig 4.2.5



Odds ratios and (exponentially transformed posterior) 95%-central intervals calculated from the smoothing spline based on robustified cross-validation, fixing the increase in one exposure for one of three situations and displaying as a function of the level of the other exposure.

The functions in the above plots are not reliable in the regions of large exposures where the data are sparse. Therefore we show the following displays illustrating the likelihood analysis only in the region where the data is concentrated, e.g. asbestos ≤ 20 and packyears ≤ 60 .

4.2.2 Likelihood analyses for logistic regression models

Proceeding as in 4.1.2 for simulated data, we first fitted the second order logistic regression model, whose estimates are given in table 4.2.1. Focussing on the interactions only the interaction ASB.PY between asbestos and smoking is slightly insignificant on the 5%-level (based on Wald's test). However, dropping this interaction increases the deviance by 4.11 with 1 d.f. significantly ($p = 4.3\%$), and we decided to keep this interaction in the model. The corresponding estimate \hat{h} of the relative risk function h is shown in Fig. 4.2.6 and along several orbits of asbestos and packyears in Fig. 4.2.7. The steep drop of \hat{h} for large exposures of asbestos and small or large packyears is not well supported by the data and mainly results from the model based quadratic extrapolation from the bulk of the data and hence differs from the spline (Fig. 4.2.2) which extrapolates linearly in

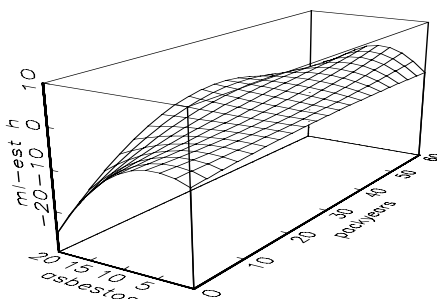
these areas.

Covariable	Parameter	S.E.	Parameter/S.E.	P-Level
Asb	+0.792769	0.455780	+1.739369	8,1970%
Asb ²	-0.085121	0.044766	-1.901473	5.7240%
PY	+0.116182	0.029600	+3.924999	0.0087%
PY ²	-0.001296	0.000443	-2.927221	0.3420%
Asb *PY	-0.067119	0.036014	-1.863682	6.2366%
Asb *PY ²	+0.001371	0.000660	+2.077063	3.7796%
Asb ² *PY	+0.006352	0.003244	+1.958040	5.0225%
Asb ² *PY ²	-0.000113	0.000055	-2.067496	3.8687%

Table 4.2.1

ML-estimates of parameters with standard errors (S.E.) and two-sided p-level for second order logistic regression model (real data)

Fig. 4.2.6



ML-estimate \hat{h} of the relative risk function h based on a quadratic logistic regression model with parameters given in table 4.2.1.

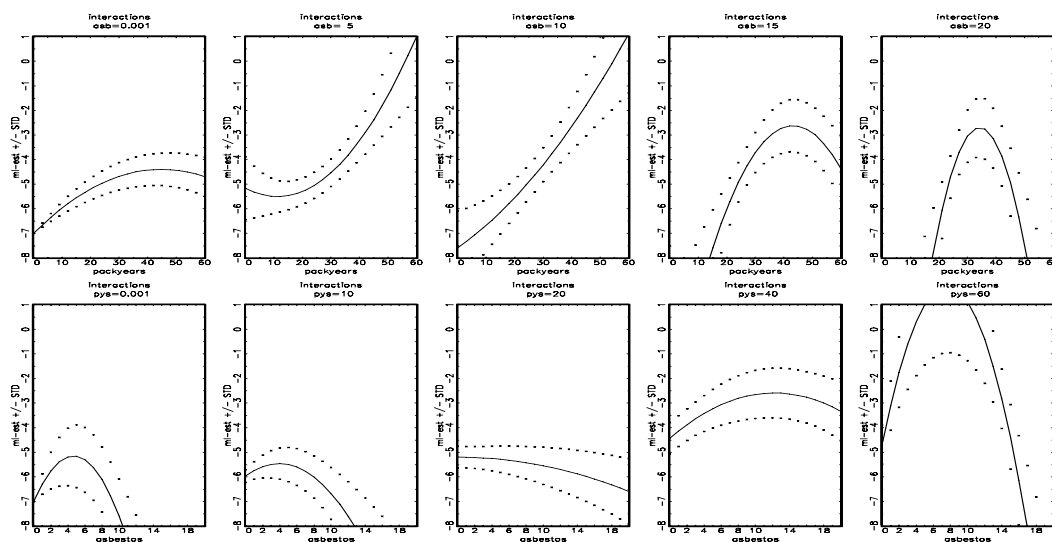
The ML-estimates of the odds ratios with 95% confidence intervals (based on the estimated asymptotic normal distribution of the parameter vector β) are displayed in Fig. 4.2.8–4.2.9.

Keeping first asbestos fixed and considering the odds ratio as a function of packyears in Fig. 4.2.8 we essentially get increasing functions. The ascent of the functions accelerate with increasing levels of asbestos and the curvature changes too thus confirming a nonlinear interaction of asbestos and packyears in the logistic model.

For fixed packyears the odds ratio functions in Fig. 4.2.8 indicate different risks depending on the level of packyears, however a constant risk is consistent with the confidence bands too. And the same picture emerges in

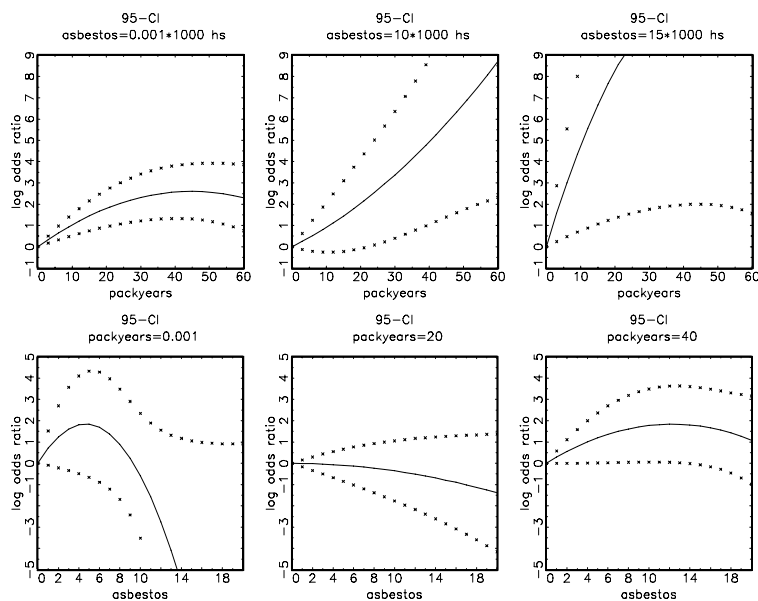
Fig. 4.2.9 for fixed increases of asbestos. The odds ratio as a function of asbestos in Fig. 4.2.9 however yield steeper functions for smaller increases of packyears although again the confidence bands are fairly wide.

Fig. 4.2.7



ML-estimate $\hat{h} \pm$ standard deviation based on table 4.2.1 along several orbits of asbestos and packyears.

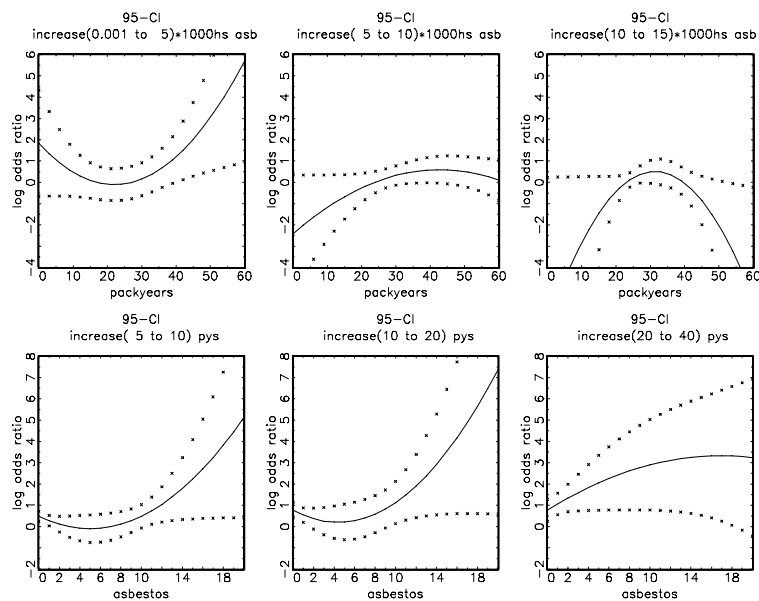
Fig. 4.2.8



ML-estimate of odds ratios with 95% confidence intervalls as a function of one exposure, keeping the other one constant at one of three levels.

Comparing our analysis for the logistic regression and the nonparametric model we arrive at compatible conclusions only within the range of exposures supported by the bulk of the data and taking into account the large variances of the estimates. While the parametric model on the one hand shows a significant non-linear interaction its restriction to quadratic functions produces inconsistent extrapolations outside the range supported by the data. The hypothesis concerning interactions suggested by the explorative analysis with splines is not confirmed by the ML-analysis. The influence of one of the two risk factors given a fixed level of the other one is at best constant, according to the estimated function however increasing, thus strengthening the risk potential.

Fig. 4.2.9



ML-estimate of odds ratios with 95% confidence intervals as a function of one exposure, keeping the increase of the other one constant for one of three situations.

5. Discussion

Nonparametric risk functions are of interest particularly for exploratory purposes. Indications of interactions between risk factors for lung cancer have been observed before. Our exploratory data analysis confirms that such interactions exist but represent a weak structure in the data. In order to refine inferences about interactions using smoothing with splines more

data in terms of repetitions would be needed. More data in terms of a finer grid introduce more parameters (function values) to be estimated and therefore are not that efficient. Repeated observations of combinations of risk factors, however, increase the precision in estimating the corresponding function values more effectively. Although the parametric ML-analysis based on second order models confirms the existence of an interaction between both risk factors, the estimated odds ratio functions from various perspectives differ – and in parts substantially – from the corresponding spline-based counterparts. Hence no consistent picture of the interaction emerges from the two approaches, possibly because of the fairly low sample size.

We formally condition the sampling distribution of a matched case-control-study on the *order statistic* of the observed risk factors: this conditioning is also necessary to ensure consistent estimates in the case of parametric functions h unless the matching variable takes only finitely many values (cf. Breslow and Day (1980), Ch. 7.1 and Osius (1997b) for a more general derivation). A small price has to be paid however, because the distribution of the rank statistic (U, V) does also contain some information about the odds-ratio which is lost in conditioning. Our conditioning approach using the order statistic naturally extends (Osius 1997b) to more general matching patterns (instead of 1 : 1) as well as to matched studies using a finite number of disease categories (instead of 1 = case and 0 = control). Our discussion of the likelihood may be of interest for Bayesian analyses of case-control studies more generally. See for example (Richardson and Gilks, 1993).

A crucial step in estimating the risk function by a smoothing spline is the determination of the smoothing parameter. We recommend to use several criteria and to check for consistency among these. A simulation study (v.d. Linde, 1997a) with binary data indicates that CV_Q (introduced in 3.2) performs very well in comparison to other competitive cross-validatory criteria. According to our experience global minima of the cross-validatory criteria do not always exist respectively occur sometimes at the margin of the pre-specified region for minimization. Sometimes also reasonable smoothing parameter are reflected by a local minimum only.

The empirical Bayesian approach to smoothing as described is based on an approximation of the likelihood function by a second order Taylor expansion. This may affect the posterior variances used to give an idea about the precision of estimates. Therefore it may be desirable to try a fully Bayesian approach with the exact likelihood and a hyperprior for the smoothing parameter instead and to obtain the posterior distributions by simulation. We tried this approach for one-dimensional risk functions and simulated data, using the Gibbs sampling as implemented in BUGS. (Spiegelhalter, et.al., 1996.) A vague hyperprior then gives a smoothing parameter estimate often consistent with that obtained from robustified cross-validation. The poste-

rior variances for the functional estimates are indeed slightly larger then. However, we often observed lack of convergence in the Gibbs sampling w.r.t. to the smoothing parameter. Convergence for the estimates of the values of the risk function was acceptable. The results are reported in (v.d. Linde, 1997b).

We conclude that, although the analysis is not immediate computationally, the great flexibility in modeling using nonparametric regression functions is worth being explored and exploited in epidemiological applications.

6. Acknowledgements

We thank the Bremen Institute for Prevention Research and Social Medicine (BIPS) for providing the real data. This work was supported by the German National Science Foundation (DFG), grant, OS 144-1.

7. References

- Breslow, N.E. and Day, N.E. (1980), *Statistical Methods in Cancer Research, Volume I: The Analysis of Case-Control Studies*. International Agency for Research on Cancer, Lyon.
- Breslow, N.E. and Storer, B.E. (1985), *General Relative Risk Functions for Case-Control Studies*, Amer. J. Epid. 122, 149-162.
- Clayton, D. and Hills, M. (1993), *Statistical Models in Epidemiology*. Oxford University Press, Oxford.
- Fahrmeir, L. and Tutz, G. (1994), *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York: Springer.
- Green, P.J. and Silverman, B.W. (1994), *Nonparametric Regression and Generalized Linear Models*. Chapman & Hall.
- Gu, Ch. (1990), *Adaptive spline smoothing in non-Gaussian regression models*. J. Amer. Statist. Ass. 85, 801-807.
- Gu, Ch. (1992a), *Penalized Likelihood Regression: A Bayesian Analysis*. Statistica Sinica 2, 255-264.
- Gu, Ch. (1992b), *Cross-Validating non-Gaussian Data*. J. Comp. Graph. Statist. 1, 169-179.
- Joeckel, K.-H., Ahrens, W., Jahn, I., Pohlabeln, H., Bolm-Andorff, U. (1995), *Untersuchungen zu Lungenkrebs und Risiken am Arbeitsplatz (Schlussbericht)*, Schriftenreihe der Bundesanstalt für Arbeitsmedizin (Fb 01 HK 546), Berlin.
- van der Linde (1997a), *Estimating the Smoothing Parameter in Generalized Spline-Based Regression I*, Experiences with Cross-Validatory Criteria in Simulated Examples with Binary Data and Small Sample Size, Preprint, University of Bremen.
- van der Linde (1997b), *Estimating the Smoothing Parameter in Generalized Spline-Based Regression II*, Empirical and Fully Bayesian Approaches. Experiences with Gibbs sampling in Simulated Examples, Preprint, University of Bremen.
- Moolgavkar, S.H. and Venzon, D.J. (1987), *General Relative Risk Regression Models for Epidemiologic Studies*, Amer. J. Epid. 126, 949-961.
- O'Sullivan, F., Yandell, B.S. and Raynor, W.J. (1986), *Automatic Smoothing of Regression Functions in Generalized Linear Models*. J. Amer. Statist. Ass. 81, 96-103.
- O'Sullivan, F. (1988), *Nonparametric Estimation of Relative Risk Using Splines and Cross-Validation*, SIAM. J. Sci. Stat. Comput. 9, 531-542.
- Osius, G. and Rojek, D. (1992), *Normal Goodness-of-Fit tests for Multinomial Models With Large Degrees of Freedom*. J. Amer. Statist. Ass. 87, 1145-1152.

- Osius, G. (1994), *Evaluating The Significance Level of Goodness-of-Fit Statistics For Large Discrete Data* In: Dirschedl, P. & Ostermann, R. (Eds.), *Computational Statistics*, Physica-Verlag Heidelberg, p. 393-417.
- Osius, G. (1997a), *A characterization of association for two random vectors in terms of odds ratios*. In preparation.
- Osius, G. (1997b), *A formal derivation of the conditional likelihood for matched case-control data*. In preparation.
- Prentice, R.L. and Pyke, R. (1979), *Logistic Disease Incidence Models and Case-Control Studies*. *Biometrika* 66, 403-411.
- Richardson, S. and Gilks, W.R. (1993). *Conditional Independence Models for Epidemiological Studies with Covariate Measurement Error*, *Statistics in Medicine* 12, 1703-1722.
- Robinson, T. and Moyeed, R. (1989), *Making Robust the Cross-Validatory Choice of Smoothing Parameter in Spline Smoothing Regression*. *Comm. Statist. - Theor. Meth.* 18, 523-539.
- Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1996) BUGS 0.5*, *Bayesian Inference Using Gibbs Sampling, Manual (version ii)*, MRC Biostatistics Unit, Cambridge.
- Thomas, D.C. (1981), *Relative-Risk Models for Survival Time and Matched Case-Control Analysis*. *Biometrics* 37, 673-686.
- Wahba, G. (1990), *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Vertrieb der Hefte 4, 14, 23, 26 durch Universitätsbuchhandlung, Bibliothekstr. 3, D-28359 Bremen. Vertrieb der übrigen Hefte (soweit nicht vergriffen) durch die Autoren oder FB 3 Mathematik/Informatik Universität Bremen, Postfach 330440, D-28334 Bremen.

1. Ulrich Krause (1976): Strukturen in unendlichdimensionalen konvexen Mengen, 74 S.
2. Fritz Colonius, Diederich Hinrichsen (1976): Optimal control of hereditary differential systems. Part I, 66 S.
3. Günter Matthiessen (1976): Theorie der heterogenen Algebren, 88 S.
4. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer (1976): Skript zur Analysis, Band 1 (11. Auflage 2000), 286 S.
5. Wolfgang Schröder (1977): Operator-algebraische Ergodentheorie für Quantensysteme, 59 S.
6. Rolf Röhrig, Michael Unterstein (1977): Analyse multivariabler Systeme mit Hilfe komplexer Matrixfunktionen, 216 S.
7. Horst Herrlich, Hans-Eberhard Porst, Rudolf-Eberhard Hoffmann, Manfred Bernd Wischnewsky (1976): Nordwestdeutsches Kategorienseminar, 193 S.
8. Fritz Colonius, Diederich Hinrichsen (1977): Optimal Control of Hereditary Differential Systems. Part II: Differential State Space Description, 36 S.
9. Ludwig Arnold (1977): Differentialgleichungen und Regelungstheorie, 185 S.
10. Rudolf Lorenz (1977): Iterative Verfahren zur Lösung großer, dünnbesetzter symmetrischer Eigenwertprobleme, 104 S.
11. Konrad Behnen, Hans-Peter Kinder, Gerhard Osius, Rüdiger Schäfer, Jürgen Timm (1977): Dose-Response-Analysis, 206 S.
12. Hans-Friedrich Münzner, Dieter Prätzel-Wolters (1978): Minimalbasen polynomialer Moduln, Strukturindizes und BRUNOVSKY-Transformationen, 53 S.
13. Konrad Behnen (1978): Vorzeichen-Rangtests mit Nullen und Bindungen, 53 S.
14. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer, Eberhard Oeljeklaus (1978): Skript zur Linearen Algebra, Band 1 (13. Auflage 2000), 249 S.
15. Günter Ludyk (1978): Abtastregelung zeitvarianter Einfach- und Mehrfachsysteme, 54 S.
16. Momme Johs Thomsen (1977): Zur Theorie der Fastalgebren, 146 S.
17. Klaus Horneffer, Horst Diehl (1978): Modellrechnungen zur anaeroben Reduktionskinetik des Cytochroms P-450, 34 S.
18. Horst Herrlich, Rudolf-Eberhard Hoffmann, Hans-Eberhard Porst, Manfred Bernd Wischnewsky (1979): Structure of Topological Categories, 252 S.
19. Hans-Friedrich Münzner, Dieter Prätzel-Wolters (1979): Geometric and moduletheoretic approach to linear systems. Part I: Basic categories and functors, 28 S.
20. Hans-Friedrich Münzner, Dieter Prätzel-Wolters (1979): Geometric and moduletheoretic approach to linear systems. Part II: Moduletheoretic characterization and reachability, 28 S.
21. Eckart Beutler, Hans Kaiser, Günter Matthiessen, Jürgen Timm (1979): Biduale Algebren, 165 S.
22. Horst Diehl, Detlef Harbach, Jürgen Timm (1980): Planung und Auswertung von Atomabsorptions-Spektrometrie-Untersuchungen mit der Additionsmethode, 44 S.
23. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer (1981): Skript zur Analysis, Band 2 (7. Auflage 2001), 299 S.
24. Horst Herrlich (1981): Categorical Topology 1971-1981, 105 S.
25. Horst Herrlich, Rudolf-Eberhard Hoffmann, Hans-Eberhard Porst, Manfred Bernd Wischnewsky (1981): Special Topics in Topology and Category Theory, 108 S.

26. H. Wolfgang Fischer, Jens Gamst, Klaus Horneffer (1984): Skript zur Linearen Algebra, Band 2 (7. Auflage 1999), 257 S.
27. Rudolf-Eberhard Hoffmann (1982): Continuous Lattices and Related Topics, 314 S.
28. Horst Herrlich, Rudolf-Eberhard Hoffmann, Hans-Eberhard Porst (1987): Workshop on Category Theory, 169 S.
29. Harald Boehme (1987): Zur Berufspraxis des Diplommathematikers, 16 S.
30. Jürgen Timm (1986): Mathematische Modelle der Dosis-Wirkungsanalyse bei den experimentellen Untersuchungen der Arbeitsgruppe zur karzinogenen Belastung des Menschen durch Luftverunreinigung, 65 S.
31. Dieter Denneberg (1988): Mathematik für Wirtschaftswissenschaftler. I. Lineare Algebra, 97 S.
32. Peter E. Crouch, Diederich Hinrichsen, Anthony J. Pritchard, Dietmar Salamon (1988, previous edition University of Warwick 1981): Introduction to Mathematical Systems Theory, 244 S.
33. Gerhard Osius (1989): Some Results on Convergence of Moments and Convergence in Distribution with Applications in Statistics, 27 S.
34. Dieter Denneberg (1989): Verzerrte Wahrscheinlichkeiten in der Versicherungsmathematik, Quantilsabhängige Prämienprinzipien, 24 S.
35. Eberhard Oeljeklaus (1989): Birational splitting of homogeneous Albanese bundles, 30 S.
36. Gerhard Osius, Dieter Rojek (1989): Normal Goodness-of-Fit Tests for Parametric Multinomial Models with Large Degrees of Freedom, 38 S.
37. Dieter Denneberg (1990): Mathematik zur Wirtschaftswissenschaft. II. Analysis, 59 S.
38. Ulrich Krause, Cornelia Zahlten (1990): Arithmetik in Krull monoids and the cross number of divisor class groups, 29 S.
39. Dieter Denneberg (1990): Subadditive Measure and Integral, 39 S.
40. Ulrich Krause, Peter Ranft (1991): A limit set trichotomy for monotone nonlinear dynamical systems, 31 S.
41. Angelika van der Linde (1992): Statistical analyses with splines: are they well defined? 22 S.
42. Dieter Denneberg (1992): Lectures on non-additive measure and integral (new edition: Non-additive measure and integral. TDLB 27, Kluwer Academic, Dordrecht (1994)), 114 S.
43. Gerhard Osius (1993): Separating Agreement from Association in Log-linear Models for Square Contingency Tables With Applications, 23 S.
44. Hans-Peter Kinder, Friedrich Liese (1995): Bremen-Rostock Statistik Seminar, 5. - 7. März 1992, 110 S.
45. Dieter Denneberg (1995): Extension of a measurable space and linear representation of the Choquet Integral, 30 S.
46. Dieter Denneberg, Michael Grabisch (1996): Shapley value and interaction index, 20 S.
47. Angelika Bunse-Gerstner, Heike Faßbender (1996): A Jacobi-like method for solving algebraic Riccati equations on parallel computers, 24 S.
48. Hans-Eberhard Porst editor (1997): Categorical methods in algebra and topology - a collection of papers in honour of Horst Herrlich, 498 S.
49. Angelika van der Linde, Gerhard Osius (1997): Estimation of nonparametric risk functions In matched case-control studies, 28 S.
50. Angelika van der Linde (1997): Estimating the smoothing parameter in generalized spline-based regression, 46 S.
51. Ursula Müller, Gerhard Osius (1998): Asymptotic normality of goodness-of-fit statistics for sparse Poisson data, 15 S.
52. Ursula Müller (1999): Nonparametric regression for threshold data, 18 S.
53. Gerhard Osius (2000): The association between two random elements – A complete characterization in terms of odds ratios, 32 S.

54. Horst Herrlich, Hans-E. Porst (2000): CatMAT 2000, Proceedings of the Conference: Categorical Methods in Algebra and Topology, 490 S.
55. Gerhard Osius (2001): A formal derivation of the conditional likelihood for matched case-control studies, 30 S.