28
⊠ Additivity and synergism

When discussing the way two exposures combine to influence the risk of disease the word interaction is used to refer to departures from either multiplicative or additive models. In general these models have no biological basis and interaction is therefore a purely statistical concept. The interaction parameters are chosen solely to test hypotheses and are not useful for describing the data when there is interaction. The word *synergism* is often used, in a similar sense, to refer to departures from a biological model for the independent action of two exposures. When the joint effect of two exposures is greater than would be expected from the separate effects, according to such a model, the exposures are said to display positive synergism. Synergism is therefore a particular kind of interaction but precisely what kind depends on the biological model for independent action.

Epidemiologists often use the word synergism without specifying precisely what they mean by independent action. In other words they use it in a statistical sense. When used in this way synergism is generally measured as a departure from an additive model. This suggests an ill-defined biological model which predicts that the rate for the joint effect of two exposures is the sum of the rates for the separate effects. An example of such a model is shown in Fig. 28.1 which refers to a situation where disease is caused by one or other of two *precipitating* events. Exposure A influences the chance of the first event occurring, while exposure B influences the chance of the second event occurring. When A and B act independently their effects on the rate will be additive because

$$\text{Rate(Event 1 } \textit{or}\text{ 2)} = \text{Rate(Event 1)} + \text{Rate(Event 2)}.$$

In cases like this it makes sense to fit an additive model so that departures from this model can be measured and used to test whether the two exposures act independently. In this chapter we consider some of the special problems which arise when using additive regression models.

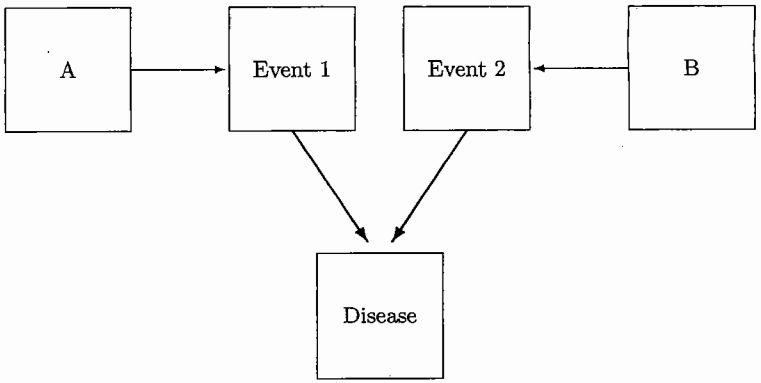

Fig. 28.1. Two precipitating events for disease.

28.1 Fitting additive models

With additive models effects are measured as differences between rates (or odds) parameters rather than as ratios. The use of stratification to control the additive effects of an exposure for confounding would be based on the assumption that the difference between the rate parameters for the different levels of exposure is constant over the strata. Formulating the same problem in terms of regression models the effects of an exposure controlled for a confounder are found by fitting the additive model for the rate,

$$\text{Rate} = \text{Corner} + \text{Exposure} + \text{Confounder}.$$

The assumption that the additive effect of the exposure is the same for all strata formed by the confounder is expressed by the fact that the model is additive, with no interaction terms.

Additive models are fitted to data by choosing parameters to maximize the log likelihood in the same way as for multiplicative models, but the calculations are different and require different computer programs. Similarly log likelihood ratios are used to test hypotheses in the same way as for multiplicative models. In practice additive models can be more troublesome to fit than multiplicative models because the most likely parameter values do not *necessarily* predict rates which are greater than zero. It is then rather difficult to know what to do. Should one treat this as evidence that the additive model is a poor fit, or should one find most likely values subject to the constraint that they predict positive rates? Generally the latter policy is followed, but it can be difficult to implement.*

*This problem does not arise with multiplicative models because these are fitted as additive models for the log rate and the log rate is not constrained to be positive.

28.2 Discriminating between additive and multiplicative models

When there are rival biological grounds for choosing an additive model and a multiplicative model the investigator will wish to discriminate between the two models by seeing which fits the data best. The deviances for the two models provide an informal way of looking at this but they cannot be compared in a formal test because the additive and multiplicative models are not nested. The solution to this technical problem is to find an *extended model* which contains both additive and multiplicative models as special cases. One such model is

$$\frac{(\text{Rate})^{\rho} - 1}{\rho} = \text{Corner} + \text{A} + \text{B},$$

where ρ is a parameter yet to be determined. In this model A and B refer to parameters which measure differences in the value of

$$\frac{(\text{Rate})^{\rho} - 1}{\rho}.$$

As ρ approaches 1 the model reduces to

$$\text{Rate} - 1.0 = \text{Corner} + \text{A} + \text{B}$$

in which the A and B parameters measure differences in the rate. As ρ approaches zero, the left-hand side of the model approaches the log of the rate [†], so the model reduces to

$$\log(\text{Rate}) = \text{Corner} + \text{A} + \text{B},$$

in which the A and B parameters measure differences in the log rate. The two extremes of the extended model therefore correspond to an additive model ($\rho = 0$) and a multiplicative model ($\rho = 1$). When this extended model is fitted for a range of values for ρ, including $\rho = 1$ and $\rho = 0$, a comparison of the log likelihoods for the different values of ρ will indicate which is the most likely value for ρ and whether the additive or multiplicative model is preferred. It may turn out, of course, that both models provide an adequate fit, or that neither model is acceptable. We do not advocate the use of the model with values of ρ other than zero or one, because effect parameters measured as differences in the value of

$$\frac{(\text{Rate})^{\rho} - 1}{\rho}$$

would be hard to interpret. The sole purpose of the extended model is to provide a framework in which to choose between additive and multiplicative models.

Using the extended model to discriminate between multiplicative and additive models involves fitting a non-standard regression model for each of a range of values of ρ. Even with software which allows non-standard models this can be quite a lot of work.

[†]This follows because, for small ρ,

$$R^{\rho} = [\exp(\log(R))]^{\rho} = \exp[\rho \log(R)] \approx 1 + \rho \log(R).$$

28.3 Additive models with case-control studies

There are some special problems which arise when trying to fit additive models to data from case-control studies. To illustrate these we shall consider a case-control study of the joint effect of two exposures A and B in which the ratio of sampling probabilities is

$$K = \frac{\text{Probability of selecting a failure as a case}}{\text{Probability of selecting a survivor as a control}}.$$

We showed in Chapter 23 that parameters which are defined as ratios of the odds of being a case are also ratios of the corresponding odds of failure in the study base. Unfortunately this does not apply to additive models. Parameters which are defined as differences in the odds of being a case are K times the corresponding differences in the odds of being a failure in the study base. The factor K, which relates the odds of being a case to the odds of faliure, cancels in ratios but not in differences. It follows that fitting an additive model to case-control data tells us nothing about the additive effects on the odds of failure in the study base except in those rare cases where the value of K is known. It is still possible, of course, to test hypotheses about zero parameter values since a zero additive effect on the odds of being a case corresponds to a zero additive effect on the odds of being a failure in the study base.

Although it is not possible to estimate the additive effects of A and B on the odds of failure in the study base it is still possible to estimate the ratio of these effects to the corner. This is less satisfactory than estimating differences in the odds themselves, but better than nothing. These new parameters are estimated by fitting the model

$$\text{Odds} = \text{Corner} \times (1.0 + \text{A} + \text{B}).$$

When the model is written in this way the corner parameter is still the odds of being a case when A and B are at level zero, but the A and B parameters are now differences in the ratio

$$\frac{\text{Odds}}{\text{Corner}}.$$

Table 28.1. Estrogen replacement, weight, and endometrial cancer

Weight (kg)	Estrogen replacement			
	No		Yes	
	Cases	Controls	Cases	Controls
< 57	12	183	20	61
57–75	45	378	37	113
> 75	42	140	9	23

This model can be fitted to data using likelihood in the same sort of way as for conventional models but special software is required.

Exercise 28.1. Table 28.1 shows results of a case-control study relating endometrial cancer incidence to use of estrogen therapy and body weight. Calculate odds ratios for each category of weight and estrogen use relative to the corner (top left corner cell). Obtain differences in these odds ratios for estrogen replacement yes compared to estrogen replacement no, at each level of weight. Do the data appear consistent with an additive model?

When a case-control study is stratified by age at time of diagnosis, and controls are sampled separately in each age stratum, there will be a different value of K for each stratum. To make sure the A and B parameters do not depend on these K's the parameters must now be defined as differences in the value of

$$\frac{\text{Odds}}{\text{Age specific corner}},$$

where the age specific corners are the odds in each age stratum when A and B are both at level 0. The A and B parameters will then equal the corresponding differences in the ratio of the odds of failure to the age specific corners in the study base.

Assuming that the new A and B parameters are constant over age strata, their common value can be estimated by fitting the model

$$\text{Odds} = \text{Corner} \times \text{Age} \times (1.0 + \text{A} + \text{B}).$$

where age is a categorical variable with one level for each age stratum. The Corner × Age part of the model corresponds to fitting separate corner parameters for each age stratum. This model again requires special software.

28.4 Discriminating between models using case-control studies

The extended model containing the extra parameter ρ can also be used to compare the fit of a multiplicative model with an additive model using

data from a case-control study. The two models we wish to compare are

$$\text{Odds} = \text{Corner} \times \text{A} \times \text{B},$$

in which A and B parameters are ratios of odds, and

$$\text{Odds} = \text{Corner} \times (1.0 + \text{A} + \text{B}),$$

in which the A and B parameters are differences in the ratios of odds to the corner. The multiplicative model can also be written in the form

$$\log(\text{Odds}) = \text{Corner} + \text{A} + \text{B},$$

in which the A and B parameters are defined as differences in log odds. The extended model is now

$$\frac{(\text{Odds}/\text{Corner})^{\rho} - 1.0}{\rho} = \text{A} + \text{B}.$$

As ρ approaches 0 this model approaches

$$\log(\text{Odds}/\text{Corner}) = \text{A} + \text{B},$$

which simplifies to

$$\log(\text{Odds}) = \log(\text{Corner}) + \text{A} + \text{B}.$$

This is the multiplicative model written in log form, apart from the fact that because the corner parameter is on the original scale in the extended model it appears as log(Corner). As ρ approaches 1, the extended model approaches

$$\text{Odds} = \text{Corner} \times (1.0 + \text{A} + \text{B}),$$

which is the additive model.

The procedure for comparing the fit of a multiplicative and an additive model is illustrated by fitting the extended model to the data in Table 28.1 for a range of values of ρ. To actually do this involved fitting a non-standard model for each of these values. The resulting log likelihood ratios are shown in Fig. 28.2. At $\rho = 0$ the log likelihood ratio is -2.774 and at $\rho = 1$ it is -0.408. To test for the adequacy of the multiplicative model we take $\rho = 0$ as the null value. Minus twice the log likelihood ratio for $\rho = 0$ is 5.548 ($p \approx 0.02$), so the data do not support this model. To test for the adequacy of the additive model we take $\rho = 1$ for the null value. Minus twice the log likelihood ratio for $\rho = 1$ is 0.816 ($p > 0.10$) so the data are consistent with the additive model.

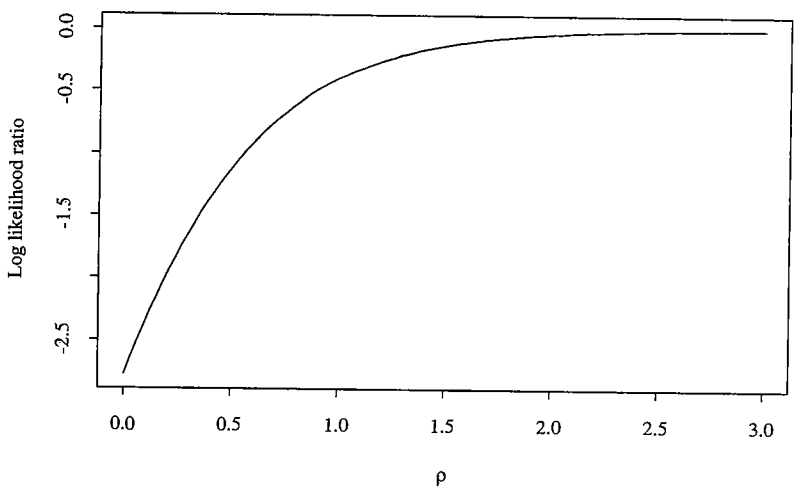

Fig. 28.2. The log likelihood ratio for ρ.

The most frequent outcome when comparing the fit of multiplicative and additive models is that both provide an acceptable description of the data. This has been taken by some epidemiologists as a serious flaw in the modern modelling approach to statistical analysis, since additive and multiplicative models have radically different public health implications (notably in relation to the targeting of interventions). This difficulty is indeed serious, but it is attributable more to an attempt to extrapolate beyond the data than to any shortcomings in statistical methodology.

A good example of this arises in attempts to study the implication of different dose-response relationships for the carcinogenic effect of ionizing radiation. The public health problem (if there is one) is one of relatively large populations exposed to low doses, but the available epidemiological studies have concentrated upon high exposure groups — A-bomb survivors, irradiated patient groups and so on. Additive and multiplicative dose-response models make similar predictions at high doses so these studies are poorly discriminated. However, they make very different predictions for subjects receiving low dose exposure. If data were available for subjects receiving low dose exposure the two models would be easily discriminated; the problem lies in trying to discriminate between them using data from a range of dose levels for which the two models make the same predictions.

Exercise 28.2. We plan to reduce the total burden of disease in a community by attempting to eliminate exposure A but another explanatory variable, B, is also known to be important. Should the intervention be targeted on individuals whose exposure to B is greatest? Consider how the answer to this question depends on whether the effects of A and B on the rate are additive or multiplicative.

Solutions to the exercises

28.1 The odds ratios are shown below.

Weight	Estrogen replacement		
(kg)	No	Yes	Difference
< 57	1.00	5.00	4.00
57–75	1.82	4.99	3.17
> 75	4.58	5.97	1.39

The additive model does not appear to fit particularly well as the differences between the odds ratios for the two estrogen groups seems to fall with increasing weight. Further examination of the table suggests the possibility that there is only a relationship with weight when there is no estrogen replacement.

28.2 Consider a population classified according to the two factors A and B. When these act additively or multiplicatively, the rates follow one of the following patterns:

	Additive model			Multiplicative model		
	A		Potential	A		Potential
B	No	Yes	reduction	No	Yes	reduction
No	1	3	2	1	3	2
Yes	3	5	2	3	9	6

When the multiplicative model holds the reduction in rates by eliminating exposure A is greater in the B-Yes group than in the B-No group. It would therefore be cost effective to target intervention at the high- risk section of the population. When the additive model holds this is no longer the case — there is an equal potential reduction in both sections of the population, and targeted intervention makes little sense.