20
Tests for trend

Up to this point we have dealt exclusively with comparisons of exposed and unexposed groups. Although it is possible that the action of an exposure is 'all or nothing', coming into play only when a threshold dose is exceeded, it is more common to find a dose-response relationship, with increasing dose leading to increasing disease rates throughout the range of exposure. This chapter introduces analyses which take account of the level or *dose* of exposure.

20.1 Dose-response models for cohort studies

The simplest model for dose-response relationship assumes that the effect of a one-unit increase in dose is to multiply the rate (or odds) by θ, where θ is constant across the entire range of exposure. Thus the effect of each increment of dose on the log rate or odds is to add an amount $\beta = \log(\theta)$. This model is called the *log-linear model* and is illustrated in Fig. 20.1. The dose level is denoted by z. The rate at dose $z = 0$ is given by $\log(\lambda_0) = \alpha$, at $z = 1$ by $\log(\lambda_1) = \alpha + \beta$, at $z = 2$ by $\log(\lambda_2) = \alpha + 2\beta$, and so on.

In principle, log-linear models present no new problems. The model describes the rate at different doses z in terms of two parameters α and β. The first of these describes the log rate in unexposed persons and will normally be a nuisance parameter; the second is the parameter β, which describes the effect of increasing exposure. The contribution to the log likelihood from D_z events in Y_z person-years of observation at dose z is

$$D_z \log(\lambda_z) - Y_z \lambda_z$$

and the total log likelihood is the sum of such terms over all levels of exposure observed. This is a function of both α and β but, as before, we can obtain a profile likelihood for the parameter of interest, β, by replacing α by its most likely value for each value of β. This profile likelihood is given by the expression:

$$\sum D_z \log\left(\frac{Y_z \exp(\beta z)}{\sum Y_z \exp(\beta z)}\right),$$

where both summations are over dose levels z. Exactly the same log likeli-

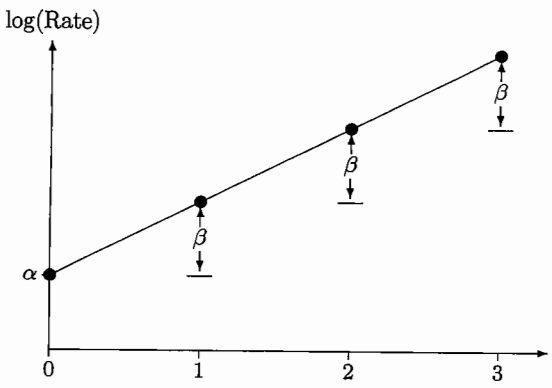

Fig. 20.1. Log-linear trend.

hood is obtained using the retrospective conditional argument based on the probability that the cases split between exposure categories in the ratios observed.

To find the most likely value of the parameter β requires computer programs for *Poisson regression*, whose use will be discussed in Part II. However, the likelihood can be used to obtain some simpler analytical procedures. Most importantly, a statistical test for the significance of a dose-response effect can be derived by calculating the gradient of the log likelihood at $\beta = 0$. This leads to the score

$$U = D\left(\frac{\sum D_z z}{\sum D_z} - \frac{\sum Y_z z}{\sum Y_z}\right)$$

where summation is over exposure doses z and, as usual, $D = \sum D_z$. The first term within the brackets is the mean exposure for *cases*, while the second is the mean exposure in the entire cohort, using the person-time observation as weights. The weighting ensures that a subject observed for twice as long contributes twice as much to the mean; this is necessary since he or she has twice the chance of becoming a case.

Denoting means of z by $\bar{z}$, the score may be written

$$U = D\,(\bar{z}_{\text{Cases}} - \bar{z}_{\text{Cohort}})\,.$$

The score variance, obtained from the curvature of the log likelihood curve

Table 20.1. Observed and expected deaths from bladder cancer in workers in the nuclear industry

Dose code, z	0	1	2	3	4	5	6
Dose (mSv):	< 10	10–	20–	50–	100–	200–	≥ 400
Observed, D_z	3	2	1	1	3	2	2
Expected, E_z	6.2	1.0	2.2	1.8	1.5	1.0	0.4

at $\beta = 0$, is

$$V = D\left[\frac{\sum Y_z(z)^2}{\sum Y_z} - (\bar{z}_{\text{Cohort}})^2\right].$$

This expression is D times the *variance* of the exposure doses z within the cohort (again weighting by person-time of observation). The calculation of weighted means and variances is easily carried out on scientific calculators which include special keys for these operations.

The same argument applies in the construction of tests for trend in SMR's except that instead of the person-time Y_z we now use E_z, the expected numbers of events obtained by application of age-specific reference rates. The use of this test is illustrated in the following example.

RADIATION AND BLADDER CANCER

Table 20.1 shows observed deaths from carcinoma of the bladder in a cohort of radiation workers, classified according to the radiation dose received. Also shown are the numbers of deaths expected in each category on the basis of England and Wales rates.* The mean dose code for the bladder cancer cases is:

$$\frac{3\times 0 + 2\times 1 + 1\times 2 + \ldots + 2\times 6}{14} = 2.93$$

The expected mean is obtained by using the expected numbers of cases as weights, is

$$\frac{6.2\times 0 + 1.0\times 1 + 2.2\times 2 + \ldots + 0.4\times 6}{14.1} = 1.72$$

so the score is

$$U = 14(2.93 - 1.72) = 16.9.$$

The weighted variance of the dose may be calculated using the appropriate calculator key, or from

$$\frac{6.2\times (0)^2 + 1.0\times (1)^2 + 2.2\times (2)^2 + \ldots + 0.4\times (6)^2}{14.1} - (1.72)^2 = 3.31,$$

*From Smith, P.G. and Douglas, A.J. (1986) *British Medical Journal*, **293**, 845–854.

so the score variance is $V = 14 \times 3.31 = 46.4$. The score test is therefore $(16.9)^2/46.4 = 6.16$, which corresponds to a p-value of 0.013. Although in this example, radiation dose was grouped into a few discrete categories, this is not a requirement of the analysis. Dose could be recorded more exactly so that no two individuals share the same dose. Observed and expected mean doses are calculated in the same way.

When the exposure dose is roughly normally distributed within cases, the log likelihood is nearly quadratic and an approximation to the most likely value of β is provided by

$$\frac{U}{V} = \frac{\text{Mean dose (cases)} - \text{Mean dose (cohort)}}{\text{Variance of dose (cohort)}}.$$

The standard deviation of this estimate is approximately $\sqrt{1/V}$.

Exercise 20.1. (a) Calculate a rough estimate of β for the bladder cancer data. (The maximum likelihood estimate is 0.328.)
(b) What is the interpretation of β? How may the effect be expressed in terms of *rate ratios*?
(c) How would the interpretation of the analysis be changed if the calculations had been carried out using the actual radiation dose as z rather than the 0–6 code?

20.2 Stratified analysis of cohort data

The extension of these ideas to stratified analysis involves only a slight extension of the model. Use of either a profile or conditional approach leads to a log likelihood function for β which is simply a sum over strata of contributions of the same form as in the previous section. In consequence, the score and score variances at $\beta = 0$ are simply sums of contributions from each stratum:

$$\begin{aligned} U &= \sum U^t \\ &= \sum D^t \left(\bar{z}^t_{\text{Cases}} - \bar{z}^t_{\text{Cohort}}\right), \end{aligned}$$

and

$$V = \sum V^t.$$

ENERGY INTAKE AND IHD

An example of the use of this method is shown in Table 20.2. The table is calculated from the same data on energy intake and ischaemic heart disease which has been encountered in previous chapters, and compares observed and expected mean energy intake of heart disease cases. The study cohort was drawn from three rather different occupational groups, bank workers, London bus drivers, and London bus conductors. To control

Table 20.2. Mean energy intake (kcal/day) of IHD cases

	Bank staff		Drivers		Conductors	
Age	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
40–49	2769	3015	2918	2853	–	–
	(4)		(2)		(0)	
50–59	2514	2894	2808	2838	2515	2845
	(8)		(4)		(5)	
60–69	2725	2846	2458	2833	2718	2828
	(7)		(6)		(9)	

for confounding by age and occupation, 9 strata are required. Table 20.2 shows the comparisons of means for the 9 strata formed by crossing the three occupational groups by three age bands. The numbers of cases are shown in parentheses.

The most striking feature of this table is the consistency of the finding that energy intake is lower in cases than would be expected under the null hypothesis. This is confirmed by the overall significance test for which

$$\begin{aligned} U &= 4 \times (2\,769 - 3\,015) + \cdots + 9 \times (2\,718 - 2\,828) \\ &= -9\,765 \\ V &= 8\,446\,000, \end{aligned}$$

so that the score test is $(-9\,765)^2/8\,446\,000 = 11.29$ and $p < 0.001$ (detailed workings for V are not shown).

The use of U and V to obtain a rough estimate of β is exactly the same as in the unstratified case.

Exercise 20.2. Calculate an approximate estimate of β for the energy intake data, using the values of U, V given above. Calculate the change in log rate predicted for a 500 kcal change in energy intake and express this as a rate ratio.

20.3 Dose-response relationships in case-control studies

The extension of these methods to deal with case-control studies requires only the change to an appropriate likelihood. In Chapter 17 we showed that this is the likelihood based upon the split of the N_z subjects observed with exposure level z as D_z cases and H_z controls. If the odds predicted by the model for such a split are ω_z, the log likelihood is

$$\sum \left[D_z \log(\omega_z) - N_z \log(1 + \omega_z)\right].$$

The idea that the rate ratio for each dose increment is constant translates, in the case-control study, to a constant *odds ratio* for each one unit change

Table 20.3. Screening histories in breast cancer deaths and controls

	Negative screens				
	0	1	2	3	Total
Cases	29	22	3	3	57
Controls	99	122	40	24	285
Subjects	128	144	43	27	442

in dose. Thus the model for the log odds takes the same form as Fig. 20.1:

$$\log(\omega_z) = \alpha + \beta z.$$

This is a *logistic regression* model. Computer programs for estimating β are widely available and their use will be discussed in Part II, but a score test of the null hypothesis $\beta = 0$ requires only simple tabulations and a hand calculator. The nuisance parameter, α, is removed either by a profile likelihood approach, or by a conditional argument leading to the hypergeometric likelihood. In either case, the score test given by the gradient of the log likelihood curve turns out to be:

$$\begin{aligned} U &= \frac{DH}{N}\left(\frac{\sum D_z z}{D} - \frac{\sum H_z z}{H}\right). \\ &= \frac{DH}{N}\left(\bar{z}_{\text{Cases}} - \bar{z}_{\text{Controls}}\right) \end{aligned}$$

The score variance is obtained from the curvature of the log likelihood and, as in section 17.3, the profile and the conditional approaches lead to slightly different expressions. For the conditional approach,

$$V = \frac{DH}{N}\frac{\sum N_z(z)^2 - N(\bar{z})^2}{(N-1)},$$

where $\bar{z}$ is the overall mean dose $(\sum N_z z)/N$. Apart from the factor DH/N, this is the usual estimate of the variance of dose in the study when cases and controls are combined. The profile likelihood argument leads to the same expression, but with $(N-1)$ replaced by N.

Exercise 20.3. In Chapter 19, a case control study of the efficacy of a radiographic breast cancer screening programme was discussed. Table 20.3 shows data drawn from a similar study concerning the number of times women had been screened (with negative result).†
(a) By calculating case/control ratios, examine the data for evidence of decreasing risk with increasing numbers of negative screens.
(b) The mean number of screens for cases is 0.649, and for controls is 0.961. The overall variance of the number of screens is 0.810. Calculate the score and score variance and the corresponding chi-squared value.

†From Palli, D. *et al.* (1986) *International Journal of Epidemiology*, **38**, 501–504.

Extension of these results to stratified and matched case-control studies follows along familiar lines. Each stratum (or case-control set) provides its own contribution to the score:

$$U^t = \frac{D^t H^t}{N^t}\left(\bar{z}^t_{\text{Cases}} - \bar{z}^t_{\text{Controls}}\right).$$

The overall score is the sum of these contributions and the score variance (using the hypergeometric conditional argument) is the sum contributions:

$$V^t = \frac{D^t H^t}{N^t}\frac{\sum_z N^t_z(z)^2 - N^t(\bar{z}^t)^2}{N^t - 1}.$$

This stratified version of the score test for $\beta = 0$ is often called the *Mantel extension* test.

Under the log-linear model, if the dose is normally distributed in controls then it will also be normally distributed in cases, but with a different mean value. In those circumstances, an estimate of β will be provided by U/V as in earlier sections.

When there are only two dose levels ($z = 0$ and $z = 1$), it can be shown that the tests set out in this chapter are identical to those discussed in previous chapters. It follows from this equivalence that all the score tests discussed in this book may be thought of as comparisons of mean exposures. This insight makes possible the use of standard computer programs for summary tabulations of large bodies of data. This is particularly valuable for preliminary analysis and for demonstrating the consistency of a finding over subgroups.

Exercise 20.4. If you are undeterred by algebra, you might like to try and prove this equivalence.

Solutions to the exercises

20.1 The rough estimate of β is $16.9/46.4 = 0.36$. This is the log of the rate ratio for one unit change in dose score. The rate ratio is $\exp 0.36 = 1.4$. The dose code is constructed so that one unit change in z represents a doubling of the radiation dose, so that the approximately fitted model suggests that doubling the radiation dose multiplies the bladder cancer rate by approximately 1.4. If the analysis had been carried out by calculating means of radiation dose itself rather than mean dose code, the implied model would have been rather different — that the *addition* of a given radiation dose would multiply the rate by some constant amount.

20.2 The rough estimate of β is $-9\,765/8\,446\,000 = -1.16 \times 10^{-3}$. This

is the change in the log rate for one unit change in energy intake. For 500 kcal change, the change in log rate is $-1.16 \times 10^{-3} \times 500 = -0.58$. This corresponds to a rate ratio of $\exp -0.58 = 0.56$. The study therefore indicates that an increase of 500 kcal in daily energy intake is associated with an approximate halving of the incidence rate of IHD.

20.3 The case/control ratios for 0, 1, 2 and 3 previous negative screens are 0.29, 0.18, 0.08 and 0.13 respectively, suggesting that mortality rates from breast cancer fall with increasing numbers of previous negative screens. The score is

$$U = \frac{57 \times 285}{342}(0.649 - 0.961) = -14.82$$

and the score variance is

$$V = \frac{57 \times 285}{342} \times 0.810 = 38.47,$$

so that the score test is $(-14.82)^2/38.47 = 5.71$, corresponding to a p-value of 0.017. The use of this test in this case is debatable, since it is not by any means clear that a simple linear or log-linear dose-response relationship should apply. The true relationship between screening history and subsequent mortality depends in a complex way upon the sensitivity of the test, the speed of growth of tumours, the relationship between prognosis and tumour stage at start of treatment, together with the time interval between screens. Most of the evidence for trend comes from the higher case/control ratios in the *never* screened group, rather than from a gradient with increasing number of screens. We must be careful not to interpret a significant trend test as indicating evidence for dose-response as such.

20.4 For cohort studies, the equivalence follows from the fact that $\bar{z}_{\text{Cases}}$ is the proportion of cases exposed, D_1/D. Similarly $\bar{z}_{\text{Cohort}}$ is the proportion of person-time exposed, Y_1/Y. The variance of a binary z in the cohort is

$$\frac{Y_1}{Y} - \left(\frac{Y_1}{Y}\right)^2 = \frac{Y_0 Y_1}{(Y)^2}$$

and substitution of these expressions into the formulas given in section 20.1 gives the same test as Chapter 13.

For case-control studies, the means of z in cases and in controls are the corresponding proportions exposed, D_1/D and H_1/H. The variance of z in the study is

$$\frac{N_1 - N(N_1/N)^2}{N-1} = \frac{N_0 N_1}{N(N-1)}.$$

Substitution of these values into the formulas of section 20.3 gives the test discussed in Chapter 17.