Table 9.1 *Deaths from coronary heart disease after 10 years among British male doctors categorized by age and smoking status in 1951.*

Age group	Smokers		Non-smokers	
	Deaths	Person-years	Deaths	Person-years
35 – 44	32	52407	2	18790
45 – 54	104	43248	12	10673
55 – 64	206	28612	28	5710
65 – 74	186	12663	28	2585
75 – 84	102	5317	31	1462

Deaths per 100,000 person years

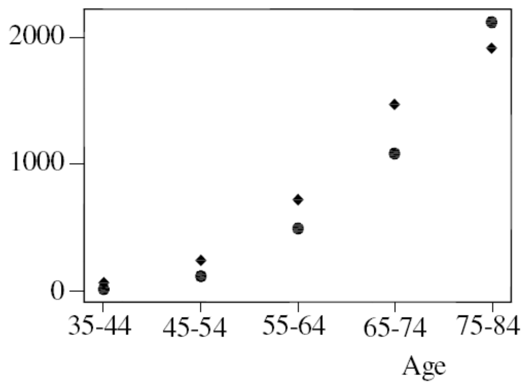


Figure 9.1 *Death rates from coronary heart disease per 100,000 person-years for smokers (diamonds) and non-smokers (dots).*

$$\log(deaths_i) = \log(population_i) + \beta_1 + \beta_2 smoke_i + \beta_3 agecat_i + \beta_4 agesq_i + \beta_5 smkage_i \quad (9.9)$$

Table 9.2 *Parameter estimates obtained by fitting model (9.9) to the data in Table 9.1.*

Term	*agecat*	*agesq*	*smoke*	*smkage*
$\widehat{\beta}$	2.376	-0.198	1.441	-0.308
$s.e.(\widehat{\beta})$	0.208	0.027	0.372	0.097
Wald statistic	11.43	-7.22	3.87	-3.17
p-value	<0.001	<0.001	<0.001	0.002
Rate ratio	10.77	0.82	4.22	0.74
95% confidence interval	7.2, 16.2	0.78, 0.87	2.04, 8.76	0.61, 0.89

Table 9.3 *Observed and estimated expected numbers of deaths and residuals for the model described in Table 9.2.*

Age category	Smoking category	Observed deaths	Expected deaths	Pearson residual	Deviance residual
1	1	32	29.58	0.444	0.438
2	1	104	106.81	-0.272	-0.273
3	1	206	208.20	-0.152	-0.153
4	1	186	182.83	0.235	0.234
5	1	102	102.58	-0.057	-0.057
1	0	2	3.41	-0.766	-0.830
2	0	12	11.54	0.135	0.134
3	0	28	27.74	0.655	0.641
4	0	28	30.23	-0.405	-0.411
5	0	31	31.07	-0.013	-0.013
sum of squares*				1.550	1.635

P

```
> data91 <- read.table(file="Table9.1.txt",header=T)
> data91 <- data.frame(data91)
> attach(data91)
> data91$agesq <- age^2
> data91
   age smoking deaths personyears agesq
1    1       1     32       52407     1
2    2       1    104       43248     4
3    3       1    206       28612     9
4    4       1    186       12663    16
5    5       1    102        5317    25
6    1       0      2       18790     1
7    2       0     12       10673     4
8    3       0     28        5710     9
9    4       0     28        2585    16
10   5       0     31        1462    25
> M2 <- glm(deaths ~ age + agesq + smoking + age*smoking,offset=log(personyears),family=poisson, data=data91)
> summary(M2)

Call:
glm(formula = deaths ~ age + agesq + smoking + age * smoking,
    family = poisson, data = data91, offset = log(personyears))

Deviance Residuals:
       1         2         3         4         5         6         7         8
 0.43820  -0.27329  -0.15265   0.23393  -0.05700  -0.83049   0.13404   0.64107
       9        10
-0.41058  -0.01275

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -10.79176    0.45008 -23.978  < 2e-16 ***
age           2.37648    0.20795  11.428  < 2e-16 ***
agesq        -0.19768    0.02737  -7.223 5.08e-13 ***
smoking       1.44097    0.37220   3.872 0.000108 ***
age:smoking  -0.30755    0.09704  -3.169 0.001528 **
---
```



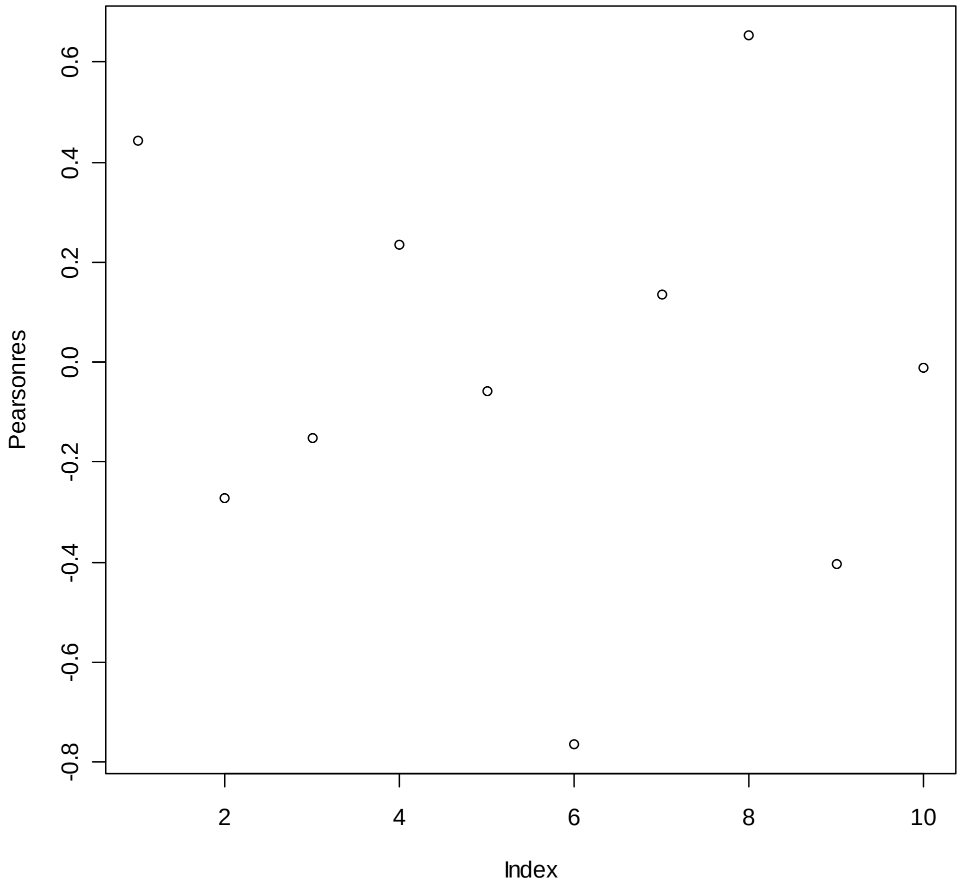
```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 935.0673  on 9  degrees of freedom
Residual deviance:   1.6354  on 5  degrees of freedom
AIC: 66.703

Number of Fisher Scoring iterations: 4

> expected <- predict(M2,type="response")
> Pearsonres <- (deaths-expected)/sqrt(expected)
> Devianceres <- sign(deaths-expected)*sqrt(2*(deaths*log(deaths/expected)-(deaths-expected)))
> cbind(age,smoking,deaths,expected,Pearsonres,Devianceres)
   age smoking deaths   expected  Pearsonres Devianceres
1    1       1     32  29.584734  0.44404929  0.43820403
2    2       1    104 106.811960 -0.27208163 -0.27328873
3    3       1    206 208.198646 -0.15237591 -0.15264528
4    4       1    186 182.827893  0.23459923  0.23392570
5    5       1    102 102.576767 -0.05694769 -0.05700118
6    1       0      2   3.414801 -0.76561908 -0.83049031
7    2       0     12  11.541629  0.13492231  0.13404370
8    3       0     28  24.743377  0.65469354  0.64106682
9    4       0     28  30.229155 -0.40544060 -0.41058325
10   5       0     31  31.071038 -0.01274427 -0.01274913
>
> sum(Devianceres^2)
[1] 1.63537
> sum(Pearsonres^2)
[1] 1.550251
> plot(Pearsonres)
```



9.3.1 Example: Cross-sectional study of malignant melanoma

Table 9.4 *Malignant melanoma: frequencies for tumor type and site (Roberts et al., 1981).* "Industrial statistics": Independence test (total number 400 given)

Tumor type	Site: Head & neck	Trunk	Extremities	Total
Hutchinson's melanotic freckle	22	2	10	34
Superficial spreading melanoma	16	54	115	185
Nodular	19	33	73	125
Indeterminate	11	17	28	56
Total	68	106	226	400

9.3.2 Example: Randomized controlled trial of influenza vaccine

Prospective study. "Industrial statistics": Homogeneity test (row totals fixed)

Table 9.6 *Flu vaccine trial.*

	Response: Small	Moderate	Large	Total
Placebo	25	8	5	38
Vaccine	6	18	11	35

9.3.3 Example: Case control study of gastric and duodenal ulcers and aspirin use

Table 9.7 *Gastric and duodenal ulcers and aspirin use: frequencies (Duggan et al., 1986).* `Four row totals are fixed.`

	Aspirin use		
	Non-user	User	Total
Gastric ulcer			
Control	62	6	68
Cases	39	25	64
Duodenal ulcer			
Control	53	8	61
Cases	49	8	57

This is a $2 \times 2 \times 2$ contingency table. Some questions of interest are:

1. Is gastric ulcer associated with aspirin use?
2. Is duodenal ulcer associated with aspirin use?
3. Is any association with aspirin use the same for both ulcer sites?