

When answering questions requiring numerical work, the results are to be reported in a narrative summary, in your own words. Tables and Figures may be included, but must be formatted along with the text. **Do not include in this summary printouts of computer code.** Analysis of variance/deviance tables, tables of coefficients and their estimated standard errors, and other output should be formatted separately and reported only to the relevant number of significant digits. All computer code used to obtain the results summarized in the response should be provided as an appendix.

1. Exercise 10.2.5, Davison (p. 479)
2. Exercise 10.3.8, Davison (p. 487)
3. Exercise 10.4.1, Davison (p. 497)
4. The data in Table 1 below is taken from *Applied Statistics* by Cox & Snell (p.176). This shows the numbers of subjects reporting “breathlessness” and “wheeze”, categorized by age group. The subjects are a sample of 18,282 coalminers known to be smokers, but with no Xray indication of lung disease.

Table 1: Set 11 from Cox & Snell (1981). Numbers of coalminers responding to breathlessness and wheeze according to age group.

Breathlessness		Yes		No		Total
Wheeze		Yes	No	Yes	No	
Age Group	20–24	9	7	95	1841	1952
	25–29	23	9	105	1654	1791
	30–34	54	19	177	1863	2113
	35–39	121	48	257	2357	2783
	40–44	169	54	273	1778	2274
	45–49	269	88	324	1712	2393
	50–54	404	117	245	1324	2090
	55–59	406	152	225	967	1750
	60–64	372	106	132	526	1136
Total		1827	600	1833	14022	18282

- (a) Consider first the incidence of wheeze among the group “breathlessness = yes”. In age group i , assume that

$$y_i \sim \text{Binomial}(n_i, p_i) \tag{1}$$

where y_i is the number in age group i reporting “wheeze = yes” and n_i is the total number in age group i . Assume that

$$\log \frac{p_i}{1 - p_i} = \beta_0 + \beta_1 x_i$$

where x_i is the midpoint of the i th age interval. Fit this model and provide a summary of the coefficient estimates and their standard error. Give an interpretation of the estimate of β_1 . Is there evidence that the incidence of wheeze increases with age?

- (b) Carry out the same analysis in the “breathlessness=no” group, and answer the same questions as in part (a).
 - (c) Is the binomial model a satisfactory choice for this data? Why or why not?
 - (d) Is there evidence of an interaction of breathlessness and wheeze, or any evidence of an interaction of of breathlessness and wheeze that changes with age?
5. In the Franzoni et al. paper, the authors analysed the number of papers submitted to *Science*, using Poisson regression, since the dependent variable is a count. Fit the model with number of papers submitted as a response, and with country fixed effects, year fixed effects, herd, and a dummy variable for institutional incentive, as covariates, using the Poisson model. In qualitative terms and non-technical language, compare the conclusion about the importance of incentives obtained from this model with the conclusion about the importance of incentives using a Normal model with the same covariates, allowing for the different interpretation of the coefficients.